Performance analysis and strategic planning of dairy supply chain in Indonesia: a comparative study
Avni Suryawati, Arif Ruton,翅出Hafidh Fathulrahman and Emily Murphee

Measuring co-operation of employee engagement across race groups
Nikky Menique Gabel and Ross Martin

Best supply chain management practices and high-performance firms: the case of Gulf manufacturing firms
Nabil Al-Zubaidy, Emad Al-Kurie, Omar Qurieh and Varun Kumar

Interpersonal antecedents to employee aligned action: mediating role of psychological safety
Hijj Ibrahim Al-Subaie and Emad Abou-El-Seoud

A performance analysis of Brazilian public health, TOPRIS and rural networks application
Nachiappan Subramanian and Sakun Boon-itt

The relationship between cost stickiness and financial reporting quality in Tehran Stock Exchange
Hamed Sanjani, Hamid Azizi and Ali Zandi

Determinants of liquidity in Malaysian SMBs: a quantile regression approach
Hassan Al-Dossary and Muna Al-Saghir

How authentic leadership promotes individual performance: mediating role of organizational citizenship behavior and creativity
Yuejie Ren, Xushua Zhan and Yuqi Pan

Four facets of learning in performance measurement
Jason Martin, Melissa E. Daniels, Bill Baljon and Harry Koch

Income diversification and bank performance: evidence from BBRC nations
Novella Binder, Ramon Naveiras and Aditya Dey

Flexible budgeting influence on organizational inertia and flexibility
Jose Carlos Temis-Olazarain, Julio Lagos, Linda Pereira-Ramos, Konrad S Serri-da-Silva, David Diet, Uzcategi Miren-Aribide, Pedro Mendizabal, Sofia Estepa, Maria Esteban and Laureano Guzman

A transformed fixed effect stochastic frontier approach for productivity evaluation in Indian electricity sector
Kishan Vijhahh

The influence of time on employee engagement in the SA business environment
Telephine Lepo, Evonith Pity, Lungisa DlANGO, Lungisa DlANGO, Lungisa DlANGO and Lungisa DlANGO

Pilotizing warehouse performance measures in contemporary supply chains
Desmond Mkhize, Chiyanuku Masiye and Mthombeni Mbambo

Assessing Egyptian construction projects performance using principal component analysis
Mohamed Aly Hefny Mahmoud and Emad Fathy El-Saad

The impacts of leadership support and coworker support on employee creative behavior
Nichol Dzahar and Hermien van Jaarsveld

A valid and applicable measurement method for knowledge worker productivity
Seth Jesse Cobham, Mohammad Reza Sholeh Neale and Hid Jugherieh Sugh

Productivity growth of Indonesian rice production: countries and efforts to improve performance
John Marzono

Critical success factors for Nigerian implementation of Nigerian construction industry
Ahmed Ibrahim Ali and Aminu Koleh

Aligning product variety with supply chain and business strategy: an integrated approach
Mohamed Ali and Hadi Sleem

Measuring resource efficiency and resource effectiveness in manufacturing
Hamed Sharbaty, Salar A and Ali Jarrett

Cost of quality measurement in food manufacturing companies: the Greek case
Engelina Tsech-preview, Konstantinos Vlachos, Yiannis Louridas and Panagiotis Economou

How transformational leadership predicts employees’ affective commitment and performance
Yuejie Ren, Shaowei Song and Er Dufu

Productivity paradox: The impact of office redesign on employee productivity
Brady J. Park and Lisa Greenfield

Can artificial neural networks predict lawyers’ performance rankings?
Suara Ferrone Lizz, Mike Berkevick Enfield and Julie Ferrone Lizz

The impact of corporate finance decisions on market value in emerging markets
Suara Ferrone Lizz and Nadeem Khan

Retail supply chain responsiveness: towards a retail-specific framework and a future research agenda
Eric Sander and Hernando Jaramillo

Human resource management, commitment and performance links in Iran and Turkey
M. Celina Teodoro, Cindy Kope and Ayelet Ragan

Effects of strategic information systems on competitive strategy and performance: the case of mattress retailers
Mehrdad M. Malak

Exploiting the relationships of strategic entrepreneurship and social capital to sustainable supply chain management and organizational performance: spilt across firm-level and industry-level study
Authentic leadership and creativity: the interesting role of psychological meaningfulness, safety and work engagement
Nicole Chittenden and Timothy Rea

The relationship between managerial entrepreneurship, earnings management and firm innovation
Mehrdad M. Malak, Mehmet Yilmaz and Mohammad Dehghan Noori

Proposing an improved economic value model for human resource valuation
Saeed Omrani, Roba Rahim and Re Salar

Efficiency in European football teams using Windows-DEA analysis and evaluation
Laura Sabaté, Daniel Sáez, Fabiola Zambon-Ferraresi and Fernando Lera-López
Economic performance and efficiency determinants of crop-producing farms in Norway

Habtamu Alem
Norsk Institutt for Bioøkonomi, As, Norway

Gudbrand Lien
Inland Norway University of Applied Science, Lillehammer, Norway, and
J. Brian Hardaker
University of New England, Armidale, Australia

Abstract

Purpose – The purpose of this paper is to explore the economic performance of Norwegian crop farms using a stochastic frontier analysis.

Design/methodology/approach – The analysis was based on a translog cost function and unbalanced farm-level panel data for 1991–2013 from 455 Norwegian farms specialized in crop production in eastern and central regions of Norway.

Findings – The results of the analysis show that the mean efficiency was about 78–81 percent. Farm management practices and socioeconomic factors were shown to significantly affect the economic performance of Norwegian crop farms.

Research limitations/implications – Farmers are getting different types of support from the government and the study does not account for the different effects of different kinds of subsidy on cost efficiency. Different subsidies might have different effects on farm performance. To get more informative and useful results, it would be necessary to repeat the analysis with less aggregated data on subsidy payments.

Practical implications – One implication for farmers (and their advisers) is that many of them are less efficient than the estimated benchmark (best performing farms). Thus, those lagging behind the best performing farms need to look at the way they are operating and to seek out ways to save costs or increase crop production. Perhaps there are things for lagging farmers to learn from their more productive farming neighbors. For instance, those farmers not practicing crop rotation might be well advised to try that practice.

Social implications – For both taxpayers and consumers, one implication is that the contributions they pay that go to subsidize farmers appear to bring some benefits in terms of more efficient production that, in turn, increase the supply of some foods so possibly making food prices more affordable.

Originality/value – Unlike previous performance studies in the literature, the authors estimated farm-level economic performance accounting for the contribution of both an important farm management practice and selected socioeconomic factors. Good farm management practices, captured through crop rotation, land tenure, government support and off-farm activities were found to have made a positive and statistically significant contribution to reducing the cost of production on crop-producing farms in the Central and Eastern regions of Norway.

Keywords Benchmarking, Cost function, Farm management, Farm performance

Paper type Research paper

1. Introduction

The United Nation has predicted that current world population of 7.6bn is expected to reach 8.6bn in 2030, 9.8bn in 2050 and 11.2bn in 2100 (United Nations, 2017). Finding a way to meet the growing demand for food represents a major challenge for farmers,
policymakers and agricultural researchers. These global challenges also affect Norway. Increasing the quantity and quality of food in response to growing food demand requires improvement in the performance of the agricultural sector. Agricultural production growth in Norway is a topic of continuing interest to researchers and policymakers who aim to improve the economic efficiency and economic sustainability of primary agriculture. Good farm management and agronomic practices, in combination with efficient input use, are the best ways to improve farm productivity. In this context, agronomic practices are steps farms incorporate into their farm management to produce crop output. These steps include land preparation, time of sowing, crop rotation, use of new crop varieties, pest control, etc. While some practices may be designed to increase output, others may be directed to reducing labor time or costs.

Farms use different farm management/agronomic practices and resource combinations to produce crops. Thus, we can expect differences in performance between farms, with some less efficient than others. Farm inefficiency can be defined as the extent to which farmers are using more resources to produce a given level of output than the resources used by the best practice farmers.

In recent years, several studies have used “benchmarking” techniques to examine the efficiency in the agricultural sector (e.g. Kumbhakar and Lien, 2010; Koesling et al., 2008; Odeck, 2007; Flaten et al., 2010; Lien et al., 2010; Flaten et al., 2014; Sipiläinen et al., 2013). “Benchmarking” is defined as a measurement of the quality of an organization’s policies, products, programs, strategies, etc., in comparison with standard measurements, or with similar measurements of its peers, see for detail Kumbhakar and Lovell (2000), and Coelli et al. (2005). The objectives of benchmarking are to analyze how the more successful organizations achieve their high-performance levels; to determine what and where improvements are called for; and to use this information to improve performance. Measuring such efficiency gaps and identifying the causes can be useful to both farm managers, taxpayers, consumers and to policymakers and planners seeking to help farmers to improve their performance (Singh et al., 2016).

An investigator may not only be interested in the level of economic performance but also might want to know which factors (exogenous variables) affect the level of farm performances (Smith et al., 2006; Parmeter and Kumbhakar, 2014).

The ability of a farm manager to convert inputs into outputs via a given technology is often influenced by “exogenous variables” that characterize the environment in which production takes place (Coelli et al., 2005) (different names have been used in the economics literature for exogenous variables, such as environmental variables, \( z \)-variables and determinants of inefficiency). Thus, the accurate measurement of the economic performance of the crop farms demands an understanding of differences in the working environment. The environmental factors include farm-specific factors, such as management skill, institutional constraints and attitude to risk, or innovations that are unmeasured but can be partially represented by observable variables such as age, experience, participation in farm improvement programs and education. Environmental variables can be expected to provide farmers with various types of opportunities and challenges, which ultimately affect their level of farm performance.

In recent years, several studies have used benchmarking techniques to examine the level and determinants of inefficiency in farming. For instance, Latruffe et al. (2004) analyzed the technical efficiency and its determinants for individual farms specialized in crop and livestock production in Poland. They reported that, on average, livestock farms are more technically efficient than crop farms. Moreover, soil quality and the degree of market integration proved to be important determinants of technical efficiency. Curtiss (2000) investigated technical efficiency and competitiveness for Czech crop production for the years 1996–1998 and reported that there was a high correlation between technical efficiency and competitiveness. Moreover, market conditions, transaction costs in marketing and bargaining power were positive determinants for competitiveness. Zhu and Lansink (2008)
analyzed the impact of the CAP reforms on the technical efficiency of three crop-producing EU countries in the period 1995–2004. They reported that the mean technical efficiency was 75 percent in the Netherlands, 70 percent in Sweden and 59 percent in Germany. The ratio of crop subsidy to the total subsidy had a positive impact on the technical efficiency in the Netherlands and Sweden, but no significant impact in Germany.

Kumbhakar et al. (2014) estimated technical efficiency and its determinants for Norwegian grain farms for the period 2004–2008. They reported that resources were sub-optimally used. Off-farm activity, direct subsidy and entrepreneurial orientation were found to have negative effects on technical efficiency, while farm managers with more experience were likely to be more efficient than those with fewer years of farm experience.

Lien et al. (2018) estimated technical efficiency of the Norwegian crop-producing farms observed from 1993 to 2014 and reported that the mean technical efficiency was 0.82–0.88. Moreover, off-farm work was found to decrease technical efficiency while direct subsidy payments had a positive and statistically significant effect on technical efficiency.

Previous studies of agricultural efficiency have given useful insights into farm performance. However, the contribution of farm management practices such as crop rotation to the economic performance of crop farms remains unclear. This paper contributes to the literature in a number of ways. First, we hypothesize that:

**H1.** In addition to commonly used efficiency determinants (land rent, off-farm activity and direct government support), cost efficiency (CE) depends on the important farm management/agronomy practice of crop rotation.

Second, unlike previous studies, which commonly estimated the technical efficiency of farms:

**H2.** We estimated the economic (cost) efficiency, which accounts for both the technical and the allocative inefficiency of each farm.

A farm achieves technical efficiency when it is able to minimize the use of inputs to produce a given amount of output so that no inputs are wasted. Allocative efficiency is achieved when the farm is able to use its inputs according to their respective relative prices. Measuring such CE gaps and identifying determinants can be useful to farm managers, policymakers, planners and advisers seeking to help farmers to improve their performance.

The rest of the paper is organized as follows: Section 2 discusses the nature of Norwegian agriculture. Section 3 addresses the approach to measuring farm performance, while Section 4 discusses model specification. Section 5 includes a discussion of the data and definitions of variables used in the cost function. Section 6 includes the results and discussion thereof. Section 7 covers conclusions and implications.

### 2. The nature of agricultural production in Norway

Norwegian farms are usually small and family-operated. Only 3.3 percent of the total Norwegian land area is farmland (Statistics Norway, 2013). Owing to the topography of the country, fields are often small, scattered and difficult to cultivate, which contributes to the high costs of agricultural production. With a relatively long winter in most parts of the country and a short growing season (five months on average), growing fodder, mainly grass, has the comparative advantage in most parts of the country. On the other hand, long summer days, with sufficient rainfall, are beneficial for crop production. Moreover, the cool climate limits the spread of pests and diseases (Steinshamn et al., 2016).

The primary objectives of the Norwegian agricultural and food policies, as set out in the White paper No. 11 (2016–2017) are: long-term food self-sufficiency along with protection of the environment; creating more added value; and maintenance of small-scale farming in all regions. To achieve these objectives the government supports the farmers. The two main support instruments used by the government are border protection measures to limit or exclude
competing imports, and budgetary payments (subsidies). The budgetary support includes market price support, special tax rules for agriculture, area payments, investment grants and grants for research and extension services. Market price support for most commodities, in the form of wholesale target prices, is provided. These target prices and most budgetary payments are negotiated on an annual July/June basis between the government and representatives of the two farmers’ unions (the Norwegian Farmers and Smallholders Union and the Norwegian Farmers Union), resulting in an Agricultural Agreement. Despite the various support measures, the value of support payments has been decreasing in real terms.

As in most developed countries, farming has become highly mechanized and the number of farms has been declining, with production becoming concentrated on fewer farms. According to Statistics Norway, the number of farms was 96,000 in 1991 and had declined to 42,000 in 2015. Moreover, 2.3 percent fewer farms were registered in 2016 compared to 2015. The number of farms growing only crops decreased by 29 percent over the years 2006–2016 (Statistics Norway, 2013). However, according to the NIBIO (2016) farm account survey report, the average size of farm holdings in 2015 was 37 ha, which was an increase of 8 percent compared to 2014. Moreover, the area of rented farmland has been increasing over time, and the average area of rented land per farm reported in 2015 was 17 ha (NIBIO, 2016).

The structure of farms in Norway has been regulated by the Norwegian Concession Act. Norwegian farmers face extensive farm policies which have effects on their choices of the size and scale of the farm. For example, the Norwegian Agricultural Authority manages the quantities produced of milk, meats, vegetables, potatoes, fruits and berries (Knutsen, 2007). Limited access to land and capital restricts productivity changes, as does policy regulation in the form of quantitative restrictions on milk supply.

Norwegian agriculture is so heavily subsidized that, without support, it would not be competitive with imports. There is a threat that Norway may be obliged by international pressures to cut back on border protection and on output-related subsidies. If that happened, it would force a dramatic and painful shift toward a more competitive agriculture. Therefore, there is a case to be made to take urgent steps to improve the productivity and management of farming.

3. Approaches to measuring farm performance

There are two main benchmarking methods in the literature to measuring the performance of farms: a parametric (econometric) approach, such as a stochastic frontier analysis (SFA); and a non-parametric approach, such as data envelopment analysis (DEA). In both methods, the basis for performance measurement is the radial contraction/expansion connecting inefficient observed points representing individual firms with reference points on the efficient frontier (There are other approaches too, for instance, Bayesian stochastic frontier (SF) (Koop and Steel, 2001), semi-parametric (Simar and Wilson, 2007) and stochastic DEA (Huang and Li, 2001), but these are not commonly used in empirical studies).

Based on a sample of producers, both the two main approaches involve estimating a “best practice” frontier for a specific industry or sample of firms. Each approach has its pros and cons. For details, see Coelli et al. (2005), Parmeter and Kumbhakar (2014) and Kumbhakar et al. (2015). The treatment of measurement error is the critical distinction between parametric and non-parametric approaches. SFA models can accommodate stochastic noise, such as measurement errors due to weather, disease and pest infestation that are likely to be significant in farming. The DEA approach is sensitive to outliers since the measurement error is ignored (e.g. Coelli et al., 2005; Barnes et al., 2009). Since farms in our study are sensitive to external random shocks, we have chosen the SFA approach to evaluate the CE scores and determinants of inefficiency.

Depending on the nature of the data set at hand, there are two classes of SFA models. If we have only one observation per farm, then a cross-sectional model has to be chosen to
estimate the performance of each individual farm. A data set that consists of different farms that are observed at different time periods is called a panel data set. A panel data set contains more information, and therefore allows us to separate unobserved heterogeneity (farm-specific effects) from inefficiency.

Since the introduction of the SFA model by Aigner et al. (1977) and Meeusen and Van den Broeck (1977), there has been considerable research to extend the basic model, creating a new subfield in econometrics. Different SFA models have been developed based on different assumptions about the temporal behavior of the inefficiency. Reviews of much of this research are provided by Kumbhakar and Lovell (2000), Coelli et al. (2005), Greene (2008) and Kumbhakar et al. (2015).

In addition to estimating CE for each crop farm, it is useful to learn about factors that affect cost inefficiency between and within farms (Lien et al., 2018). The first SFA models that dealt with modeling the impact of exogenous variables (z-variables) on the level of inefficiency between and within farms are those of Kumbhakar et al. (1991) and Reifsneider and Stevenson (1991). In these cross-sectional models, the variance of the inefficiency term was specified as a function of a set of exogenous variables (z-variables). Later, Wang (2002) introduced a further generalization in which both the mean and variance of the inefficiency are functions of exogenous variables. For detailed reviews see Parmeter and Kumbhakar (2014) and Lien et al. (2018).

Among panel data models, the inefficiency specification used by Battese and Coelli (1995), commonly known as BC95, is frequently used in the empirical analysis of performance studies (Parmeter and Kumbhakar, 2014). Using the BC95 model, a researcher not only estimates the efficiency score, but also can investigate determinants of firm inefficiency (exogenous variables) in a single-step procedure. It is also possible to estimate the determinants of inefficiency in two steps, but that method creates biased results (Wang and Schmidt, 2002). In the single-step procedure, the parameters of inefficiency and the determinants are estimated together via maximum likelihood.

The general BC95 model for our panel data in a cost function form can be specified as:

\[
\ln C_{it} = \beta_0 + f(y_{it}, w_{it}; \beta) + u_{it}(\varepsilon_{it}; \delta) + v_{it},
\]

where \( \ln C_{it} \) is the logarithm of actual costs incurred by farm \( i \) in time \( t \); \( f(y_{it}, w_{it}; \beta) \) the chosen function form (e.g. Cobb-Douglas, Translog); \( y_{it} \) a vector of outputs; \( w_{it} \) the vector of input prices; and \( \beta \), the vector of parameters to be estimated. The component \( v_{it} \) is a symmetric disturbance (error term) capturing the effects of noises that are beyond the control of the farmers. The error term has both positive and negative effects, and satisfies the classical assumptions, i.e., \( v_{it} \sim N(0, \sigma_v^2) \), \( v_{it} \perp u_{it} \). Further, \( u_{it} \) is a one-sided non-negative term, accounting for inefficiency. In the BC95 model, \( u_{it} \) is obtained by truncating (at zero) the normal distribution, i.e., \( u_{it} \sim N^+(\mu_{it}, \sigma_{u_{it}}^2) \) and \( \mu_{it} = z_{it}\delta \), \( z_{it} \) is a vector that includes exogenous variables associated with variability in the efficiency score and \( \delta \) is an unknown parameter to be estimated.

The CE is the ratio of the minimum cost of each farm \( \exp(f(y_{it}, w_{it}; \beta)+v_{it}) \) to its actual cost \( \exp(f(y_{it}, w_{it}; \beta)+u_{it}+v_{it}) \), i.e. \( \text{CE} = \exp(-u_{it}) \). CE has a value between 0 and 1, with 1 defined a cost-efficient farm. Since only the sum of two error terms \( (\varepsilon_{it} = u_{it}+v_{it}) \) can be observed, the farm’s cost inefficiency index can be estimated using the conditional mean of the inefficiency term, as proposed by Jondrow et al. (1982), i.e. \( E[u_{it}]/v_{it}+u_{it}] \).

4. Empirical model specification
Consistent with the farm efficiency literature (e.g. Christensen and Greene, 1976), we estimated a transcendental logarithmic (translog) cost function incorporating Hicks-neutral technology change. We used panel data but, to simplify the notation, we have dropped the
subscripts \(i\) and \(t\), where \(i\) would denote firm \(i = 1, \ldots, n\) and \(t\) would denote time, \(t = 1, \ldots, T\).

Our specification of the cost function \(C\), including the prices of inputs \((w_j)_{j=1, \ldots, J}\), outputs \((y_k)_{k=1, \ldots, K}\), and a local wage distribution index \((r_m)_{m=1, \ldots, M}\), is:

\[
\ln C = \beta_0 + \sum_{j=2}^J \alpha_j \ln w_j + \sum_{k=1}^K \beta_k \ln y_k + \sum_{m=1}^M \phi_m \ln r_m + \delta t \\
+ \sum_{k=1}^K \sum_{j=1}^J \alpha_{kj} \ln y_k \ln w_j + \frac{1}{2} \sum_{m=1}^M \sum_{m=1}^M \phi_{mm} \ln r_m \ln r_m \\
+ \frac{1}{2} \sum_{k=1}^K \sum_{k=1}^K \beta_{kk} \ln y_k \ln y_k + \sum_{j=2}^J \sum_{k=1}^K \beta_{kj} \ln w_j \ln w_j + \delta t^2 \\
+ \sum_{k=1}^K \sum_{m=1}^M \beta_{km} \ln y_k \ln r_m + \sum_{j=2}^J \sum_{m=1}^M \alpha_{jm} \ln w_j \ln r_m \\
+ \sum_{j=2}^J \alpha_{jt} \ln w_j + \sum_{k=1}^K \beta_{kt} \ln y_k + \sum_{m=1}^M \phi_{mt} \ln r_m + \epsilon. \tag{2}
\]

The error terms \(\epsilon\) splits into two components, i.e., \(\epsilon = \nu_i + \eta_t\). As discussed in section 3 component \(u_i\) captures cost inefficiency with \(u_i \sim N^+ (\mu_i, \sigma^2_{ui})\). \(v_i\) is the error term capturing random shocks and assumed to be symmetric and to satisfy the classical assumptions i.e., \(v_i \sim N(0, \sigma^2_v)\). All Greek letters are parameters to be estimated. The trend variable, \(t\), is included to capture technology change and starts with \(t = 1\) for 1991 and increases by one annually.

Economic theory requires imposition of price homogeneity and symmetry restrictions on the parameters. Symmetric restrictions require \(\sum_{j=1}^J \alpha_j = 1\), and \(\sum_j \beta_{kk} = \sum_j \beta_{kj} = \sum_m \phi_{mm} = 0\). An easy way to impose price homogeneity is to divide the all inputs prices and total cost by an arbitrary chosen input price. Thus, in Equation (2) the left-hand side is re-defined as \(\ln C = \ln (C/\ln w_i)\), and all input prices are re-defined as \(\ln w_j = \ln (w_j/\ln w_i)\), i.e., we divided all input prices and the total cost by wages before estimating the translog cost function. Given the translog specification in (2), the farm-specific cost inefficiency and marginal effects of the exogenous variables are calculated following the procedures of Jondrow et al. (1982) and of Wang (2002), respectively.

As discussed above in Section 3, we included exogenous determinants of farm CE in our model. The choice of variables in the final model was based on two criteria. First, we considered data availability. Many variables that could affect crop management, such as skill, education level, soil type and slope or aspect of the farmland, were not available in our data set, so could not be included. Second, we considered the literature available on the subject, for example, Latruffe et al. (2004), Bozoğlu, and Ceyhan (2007) and Lien et al. (2018). As a result, we chose the following variables:

1. **Crop rotation** (\(z_1\)) – reflecting the impact of rotation system on crop and forage production, measured as the ratio of non-cereal crops such as root and legume crops to the total cropped area. We expected the farm managers’ decisions to rotate the type of crop grown on the land would make a positive contribution to the performance of their farms. Our expectation is in line with other research findings that crop rotation reduces agriculture’s dependence on external inputs through internal nutrient recycling, maintenance of the long-term productivity of the land (Gebremedhin and Schwab, 1998). Crop rotation can improve the fertility of the soil.
(Reckling et al., 2015) and can increase yields to a higher degree than pesticide intensity and tillage use (Deike et al., 2008).

(2) Land tenure ($z_2$) – the proportion of the total farmland that is rented. Farm managers’ decisions to rent land or not depend on the price of land and other factors. We expected a positive contribution to this variable because we hypothesize that well-performing farms more commonly rent extra land. Deininger and Byerlee (2011) support our hypothesis and stressed that land rental markets can help transfers of land from producers with low levels of productivity and low comparative advantage in agriculture to more efficient farmers.

(3) Off-farm activity ($z_3$) – the ratio of time of owner plus partner allocated to off-farm activity to the total time assigned to the farm. We expected a positive contribution from this variable on the basis that off-farm experience and income are likely to promote better farm management. Our hypothesis is supported by the literature. Off-farm income has been found to relax cash constraints and to allow farmers to spend significantly more on improved farming technologies (e.g. Pfeiffer et al., 2009; Stampini and Davis, 2009).

(4) Government support ($z_4$) – the ratio of government assistance to the total agricultural income. We hypothesized a positive contribution from this variable because government support can motivate farm managers to invest in new technology and may facilitate such investment by easing cash flow constraints. In line with our hypothesis, there is evidence that government support helps to promote better use of economic resources (e.g. Ferjani, 2008; Kumbhakar and Lien, 2010).

5. Data
Our data source is the Norwegian Farm Accountancy Survey collected by the Norwegian Institute of Bioeconomy Research (NIBIO). The survey participants are selected from a list of farmers, randomly drawn from the register of grants kept by the Norwegian Agricultural Agency. The data include production and economic data collected annually by NIBIO from about 1,000 farms in all regions of Norway. The number of participants varies from year to year. Approximately 10 percent of the surveyed farms are replaced per year to incorporate changes in the population of farms in Norway. Participation in the survey is voluntary. There is no limit on the number of years a farm is included in the study. Some of the farmers have participated in more than 20 years, and others have started participating for the first time. To accommodate panel features in estimation, we included only those farms for which at least three consecutive years of data are available.

To assess the efficiency and productivity growth, we needed to be sure that farmers under consideration are comparable. To obtain a homogeneous group, we choose only farms in the two main cropping areas where 98 percent of the cultivated land was located. Figure 1 shows that out of the 286,000 ha of land cultivated for grain and forage production in 2012, 81 percent was located in the Eastern Norway and 17 percent in Central Norway.

The set of data used in this study is a farm-level unbalanced panel data with 3,885 observations from 455 farms specialized in the production of grain and forage crops during the period 1991–2013.

5.1 Variables in the model
The outputs in Equation (2) are grain production in kg, adjusted for quality, i.e., feed units (FU) ($y_1$), forage production in FU ($y_2$) and value of other crop outputs in Norwegian kroner (NOK) ($y_3$). Grain yield is an aggregate of the four main crops: barley, wheat, oats and oilseeds. FU is a measure of the physical output adjusted for differences in the quality of outputs.
1 FRU is defined as 1 kg of grain with 15 percent water content. Thus, the output measure is a quality-adjusted yield of all crops in kilograms per year.

The input prices \( w_j \) in the cost function in Equation (2) are specified as follows: land price is based on market price for land in terms of rents paid for land at the farm level. The price of labor is the wage for hired labor. We computed the implicit prices (opportunity costs) of owned land and family labor based on data for farm-level rents and wages provided by NIBIO. We included a local wage distribution index \( r \) to control for regional variation in wages in our analysis. These data were provided by the Norwegian Tax Administration. The prices of materials and capital costs were constructed as Laspeyres indices based on figures provided by NIBIO. Descriptive statistics of the data are shown in Table I.

6. Estimation results and discussion
6.1 Estimation procedures and hypothesis tests
The cost function was estimated using STATA® version 14. The trend variable was normalized to be zero in the year 2013. We estimated the model for the whole sample.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Label</th>
<th>Unit</th>
<th>Eastern region</th>
<th>Central region</th>
<th>Both regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC</td>
<td>Total cost</td>
<td>NOK</td>
<td>46,016 (30,813)</td>
<td>36,761 (25,183)</td>
<td>43,624 (29,736)</td>
</tr>
<tr>
<td>( y_1 )</td>
<td>Grain output</td>
<td>1,000 FU</td>
<td>68.44 (69.89)</td>
<td>38.50 (38.61)</td>
<td>60.70 (64.64)</td>
</tr>
<tr>
<td>( y_2 )</td>
<td>Forage output</td>
<td>1,000 FU</td>
<td>72.96 (53.08)</td>
<td>87.18 (55.87)</td>
<td>76.63 (54.17)</td>
</tr>
<tr>
<td>( y_3 )</td>
<td>Other outputs</td>
<td>1,000 NOK</td>
<td>9.52 (7.98)</td>
<td>8.73 (6.04)</td>
<td>9.32 (7.54)</td>
</tr>
<tr>
<td>( w_1 )</td>
<td>Rent</td>
<td>NOK/hectare</td>
<td>2,572.60 (159.29)</td>
<td>1,790.60 (122.87)</td>
<td>2,370.50 (154.55)</td>
</tr>
<tr>
<td>( w_2 )</td>
<td>Wage</td>
<td>NOK/hours</td>
<td>141.34 (31.86)</td>
<td>153.56 (30.35)</td>
<td>144.50 (31.93)</td>
</tr>
<tr>
<td>( w_3 )</td>
<td>Materials price</td>
<td>index</td>
<td>66.38 (15.42)</td>
<td>69.51 (15.83)</td>
<td>67.19 (15.59)</td>
</tr>
<tr>
<td>( w_4 )</td>
<td>Capital price</td>
<td>index</td>
<td>80.19 (10.31)</td>
<td>82.33 (10.35)</td>
<td>80.74 (10.36)</td>
</tr>
<tr>
<td>( r )</td>
<td>Region index</td>
<td>index</td>
<td>0.14 (0.01)</td>
<td>0.15 (0.01)</td>
<td>0.14 (0.01)</td>
</tr>
<tr>
<td>( Z_1 )</td>
<td>Crop rotation</td>
<td>ratio</td>
<td>0.22 (0.07)</td>
<td>0.09 (0.02)</td>
<td>0.20 (0.07)</td>
</tr>
<tr>
<td>( Z_2 )</td>
<td>Land tenure</td>
<td>ratio</td>
<td>0.27 (0.25)</td>
<td>0.24 (0.22)</td>
<td>0.26 (0.24)</td>
</tr>
<tr>
<td>( Z_3 )</td>
<td>Off-farm activity</td>
<td>ratio</td>
<td>0.04 (0.09)</td>
<td>0.04 (0.09)</td>
<td>0.04 (0.10)</td>
</tr>
<tr>
<td>( Z_4 )</td>
<td>Gov. support</td>
<td>ratio</td>
<td>0.26 (0.11)</td>
<td>0.32 (0.10)</td>
<td>0.28 (0.11)</td>
</tr>
<tr>
<td>Year</td>
<td>Trend (t)</td>
<td></td>
<td></td>
<td>(1 = year 1993)</td>
<td></td>
</tr>
<tr>
<td>( n )</td>
<td>Sample size</td>
<td></td>
<td>2,881</td>
<td>1,004</td>
<td>3,885</td>
</tr>
</tbody>
</table>

Table I. Descriptive statistics (mean values) for crop production in eastern and central regions

Note: Standard deviations in parentheses
The estimated parameters for the translog frontier models are listed in Table VI. Various specification tests were conducted to obtain the best model and functional form for the data under analysis. A series of hypotheses about the nature of the frontier model and the consistency of the cost function with its properties were tested using Likelihood Ratios. The null hypothesis of an OLS specification was rejected at the 0.01 percent significance level. Before estimating the cost function, the skewness of the data was tested (Schmidt and Lin, 1984). The test of skewness returned a \( p \)-value of less than 0.001 showing that the null hypothesis of no skewness was confidently rejected. Therefore, we found support for a right-skewed error distribution, and hence for the SF specification of the model. We also tested the characteristics of the technology with the result that a Cobb-Douglas technology specification was rejected (see Table II). Thus, we estimated a cost function using a translog function specification.

6.2 Cost efficiency scores

Estimates of CE scores of the farms are presented in Table III. The results show that, at the means, the minimum costs are about 78 and 81 percent of the actual costs for farms in the central and eastern regions, respectively. The implication is that average actual cost per farm could be reduced by 19 to 22 percent. These results are broadly in line with other studies. Kumbhakar et al. (2014) estimated six different models for the Norwegian grain farming and reported that the mean technical efficiency varied from 0.64 to 0.91. Odeck (2007) found a mean technical efficiency for Norwegian grain production of 0.70 for SFA and 0.75 for DEA. Osborne and Trueblood (2006) estimated 70–86 percent inefficiency for Russian crop production. Our estimate is greater than the estimate by Onyenweaku and Okoye (2007) of an average CE of 59 percent for cocoyam farmers in Anabra state, Nigeria.

Table III also shows the distribution of the farms in the sample according to their CE. Thus, 1 percent of the farms are only 60 percent cost efficient while 10 percent of the sample farms are 71 percent cost efficient. We also checked for differences in the efficiency scores between the eastern and central regions using a pairwise comparison of the means test.

<table>
<thead>
<tr>
<th>Restrictions</th>
<th>Parametric restrictions</th>
<th>Wald test statistics</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD technology</td>
<td>( H_0: ) all interaction terms are zero</td>
<td>941.57</td>
<td>0.00</td>
</tr>
<tr>
<td>Normality test/Test return of Skewness</td>
<td>Schmidt and Lin (1984)</td>
<td>11,850.97</td>
<td>0.00</td>
</tr>
<tr>
<td>LR test for random effects</td>
<td></td>
<td>39,235.22</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table II. Properties of grain and forage production technology

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Eastern region</th>
<th>Central region</th>
<th>Both regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.61</td>
<td>0.59</td>
<td>0.60</td>
</tr>
<tr>
<td>5</td>
<td>0.69</td>
<td>0.67</td>
<td>0.68</td>
</tr>
<tr>
<td>10</td>
<td>0.72</td>
<td>0.70</td>
<td>0.71</td>
</tr>
<tr>
<td>25</td>
<td>0.76</td>
<td>0.74</td>
<td>0.76</td>
</tr>
<tr>
<td>Mean</td>
<td>0.81</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>75</td>
<td>0.85</td>
<td>0.82</td>
<td>0.84</td>
</tr>
<tr>
<td>90</td>
<td>0.90</td>
<td>0.87</td>
<td>0.89</td>
</tr>
<tr>
<td>95</td>
<td>0.92</td>
<td>0.90</td>
<td>0.91</td>
</tr>
<tr>
<td>99</td>
<td>0.96</td>
<td>0.93</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table III. Cost efficiency scores for both regions and the pooled data

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Eastern region</th>
<th>Central region</th>
<th>Both regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>0.07</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>Observations</td>
<td>2,881</td>
<td>1,004</td>
<td>3,885</td>
</tr>
</tbody>
</table>
We found the difference to be statically significant (see Table IV). Our results also showed that the average inefficiency level was higher, on average, in the years 2000–2013 compared to the years 1991–2000 (see Table V).

6.3 Elasticities and determinants of cost inefficiency

We checked the robustness of our results by separate estimates for the two regions and different periods. The regression results in Table VII show that all the first-order terms, except materials price ($w_3$) and the regional index ($r$) are significant. The models were found to exhibit positive and highly significant first-order parameters, fulfilling the monotonicity condition as expected for a well-behaved cost function. The elasticities of costs of grain, forage, and other outputs were 0.25, 0.20 and 0.33, respectively. This means, for instance, that if grain output increases by 1 percent, costs increase by an estimated 0.25 percent, ceteris paribus. The elasticity of cost of land was 0.02 and significant at the 5 percent level. If the price of land rose by 1 percent, costs will increase by an estimated 0.02 percent, ceteris paribus. The elasticity of cost of material inputs was 0.15, but was not statistically significant. The coefficient for the price of capital (fixed input) (0.83) is the largest among other partial elasticities and statistically significant ($p < 0.001$). This result implies that crop production in Norway is capital intensive in that the percentage change in the capital price has a larger influence on the costs of crop production than the costs of other inputs. Thus, farm managers who want to improve the crop farming needs to give priority to the wise use of these costs.

The lower part of Table VI presents the coefficients estimated for the determinants included in the inefficiency effects model. The results indicate that agronomic and socioeconomic factors influence CE. Farm management practice, specifically crop rotation, was found to make a positive and significant contribution to reducing the cost of production. This result is in accord with our expectation, suggesting that crop rotation decreases the cost of production, probably via a reduction in the use of variable inputs such as fertilizer and/or by an increase in output for given amounts of input. Thus, our hypothesis is sustained. Crop rotation can improve the fertility of the soil, interrupt the life cycles of insect pests and weeds, and can help control soil-borne diseases (Reckling et al., 2015; Vereijken, 1997; Melander et al., 2005). Deike et al. (2008) claimed that crop rotations increase yields to a higher degree than pesticide intensity and tillage use.

An increase in off-farm activity was found to be associated with a significant reduction in cost inefficiency (increase in CE) among the farm households. While this result might seem to be counterintuitive, possible reasons are that off-farm activities broaden farmers’ experience, leading to improved farm management, or off-farm income may relax cash

<table>
<thead>
<tr>
<th>Group</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central region</td>
<td>1,004</td>
<td>0.783</td>
<td>0.068</td>
<td>-11.86</td>
<td>0.000</td>
</tr>
<tr>
<td>Eastern region</td>
<td>2,881</td>
<td>0.813</td>
<td>0.068</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table IV. Pairwise comparison of means of cost efficiency between eastern and central regions

<table>
<thead>
<tr>
<th>Group</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central region</td>
<td>1,738</td>
<td>0.903</td>
<td>0.040</td>
<td>70.61</td>
<td>0.000</td>
</tr>
<tr>
<td>Eastern region</td>
<td>2,147</td>
<td>0.746</td>
<td>0.085</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table VI
Estimates of parameters in the Translog cost function and inefficiency determinants

<table>
<thead>
<tr>
<th>Variable first order (elasticities(^a))</th>
<th>(y_1)</th>
<th>(y_2)</th>
<th>(y_3)</th>
<th>(w_2)</th>
<th>(w_3)</th>
<th>(w_4)</th>
<th>(r)</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grain (y_1)</td>
<td>0.25*** (0.01)</td>
<td>0.09*** (0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forage (y_2)</td>
<td>0.20*** (0.01)</td>
<td>-0.01 (0.01)</td>
<td>0.07*** (0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (y_3)</td>
<td>0.33*** (0.01)</td>
<td>-0.06*** (0.00)</td>
<td>0.00 (0.01)</td>
<td>0.09** (0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rent (w_2)</td>
<td>0.02* (0.01)</td>
<td>0.03*** (0.01)</td>
<td>0.01 (0.01)</td>
<td>0.01 (0.01)</td>
<td>0.00 (0.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material (w_3)</td>
<td>0.15 (0.18)</td>
<td>0.21** (0.10)</td>
<td>-0.17 (0.11)</td>
<td>-0.13 (0.13)</td>
<td>-0.28 (0.15)</td>
<td>8.40* (4.47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital (w_4)</td>
<td>0.83*** (0.19)</td>
<td>-0.28** (0.11)</td>
<td>0.04 (0.12)</td>
<td>0.10 (0.14)</td>
<td>0.32 (0.17)</td>
<td>-7.02 (4.70)</td>
<td>5.39 (5.01)</td>
<td></td>
</tr>
<tr>
<td>Region index (r)</td>
<td>0.13 (0.11)</td>
<td>0.12** (0.00)</td>
<td>-0.01 (0.00)</td>
<td>-0.10 (0.05)</td>
<td>-0.21** (0.00)</td>
<td>-0.35 (1.24)</td>
<td>0.71 (1.34)</td>
<td>0.97* (0.45)</td>
</tr>
<tr>
<td>Year</td>
<td>0.02 (0.02)</td>
<td>-0.01*** (0.00)</td>
<td>-0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>-0.12* (0.06)</td>
<td>0.10 (0.06)</td>
<td>-0.00 (0.02)</td>
</tr>
</tbody>
</table>

**Exogenous inefficiency determinants\(^b\)**

- Crop rotation: 0.31*** (0.06)
- Land tenure: 0.10*** (0.02)
- Off-farm activity: 0.39*** (0.04)
- Government support: 0.65*** (0.07)
- Log. L. = 122*** (0.00)

\(\gamma = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2) = 0.21\)  
AIC = -139  
\(n = 3,885\)  
BIC = 186

**Notes:**
\(^a\)Elasticities (first order parameters); \(^b\)positive efficiency score parameter estimates show that the variable has a positive effect on cost efficiency. Standard errors in parentheses. *\(p < 0.05\); **\(p < 0.01\); ***\(p < 0.001\)
constraints and allow farmers to spend significantly more on improving farming technologies (Pfeiffer et al., 2009; Stampini and Davis, 2009). However, our finding is not consistent with a study of Norwegian grain farms from 1991–2005 which reported that there was no systematic difference in technical efficiency between part-time and full-time farmers (Lien et al., 2010). Perhaps the different findings might be explained by our use of a longer data series, including more recent years, than that available to Lien et al.

The results show that government support has a positive and significant effect on the efficiency of crop production. Our finding is in line with earlier studies indicating that the subsidies help the technological development of beneficiary farms. Government support may give incentives for technological innovation that increase efficiency. In line with our results, some have claimed that better use of economic resources is achieved (Perjani, 2008; Kumbhakar and Lien, 2010). On the other hand, other studies (e.g. Giannakas et al., 2001) have shown that government payments reduce producer incentives to generate the highest possible income from farming. Neoclassical economists believe that government support distorts the allocation of resources, compared to market equilibrium and leads to higher costs of production. They argue that, in the longer run, subsidies are capitalized into higher land prices, making it harder for potential progressive new entrants to start farming. It seems that further study is needed on the specific influence of various government supports on the efficiency of resource use.

Our results indicate that land tenure also plays an important role in explaining CE differentials among crop producers. In particular, the greater the proportion of the leased land, the higher is the CE. One possible reason might be that productive and efficient farms are spreading the costs of fixed factors such as tractors over larger areas by renting more land. Deininger and Jin (2009) found that land rental markets can help to move toward a more economic distribution of operational farm sizes through transfers of land from producers with low levels of productivity and low comparative advantage in agriculture to more efficient farmers. Following such transfers, agricultural output and incomes will be higher for those renting (Deininger and Byerlee, 2011).

Although, as noted, crop rotation was found to have positively and significantly contributed to reducing the cost of production in both regions, the magnitude of crop rotation contribution in the central region at 0.96 is higher than the 0.29 contribution in the eastern region (Table VII). Moreover, the importance of crop rotation in the latest

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Grain(y1)</td>
<td>0.29*** (0.01)</td>
<td>0.18*** (0.01)</td>
<td>0.30*** (0.01)</td>
<td>0.23*** (0.01)</td>
<td>0.25*** (0.01)</td>
</tr>
<tr>
<td>Forage(y2)</td>
<td>0.19*** (0.01)</td>
<td>0.21*** (0.02)</td>
<td>0.21*** (0.02)</td>
<td>0.15*** (0.01)</td>
<td>0.20*** (0.01)</td>
</tr>
<tr>
<td>Other(y3)</td>
<td>0.31*** (0.01)</td>
<td>0.44*** (0.01)</td>
<td>0.35*** (0.02)</td>
<td>0.36*** (0.02)</td>
<td>0.33*** (0.01)</td>
</tr>
<tr>
<td>Rent(r2)</td>
<td>0.03*** (0.01)</td>
<td>0.03*** (0.02)</td>
<td>0.03*** (0.01)</td>
<td>0.07*** (0.01)</td>
<td>0.02*** (0.01)</td>
</tr>
<tr>
<td>Material(w3)</td>
<td>0.09 (0.22)</td>
<td>0.47 (0.27)</td>
<td>0.01 (0.42)</td>
<td>0.55 (0.35)</td>
<td>0.15 (0.18)</td>
</tr>
<tr>
<td>Capital(w3)</td>
<td>0.85*** (0.23)</td>
<td>0.58** (0.28)</td>
<td>0.97** (0.42)</td>
<td>0.44 (0.35)</td>
<td>0.83*** (0.19)</td>
</tr>
<tr>
<td>Region index (r)</td>
<td>0.01 (0.91)</td>
<td>0.28 (0.52)</td>
<td>−0.02 (0.16)</td>
<td>0.31** (0.16)</td>
<td>0.13 (0.11)</td>
</tr>
</tbody>
</table>

**Exogenous inefficiency determinates**

| Crop rotation | 0.29*** (0.06) | 0.96*** (0.35) | 0.19** (0.07) | 0.97*** (0.14) | 0.31*** (0.06) |
| Land tenure | 0.09*** (0.04) | 0.11*** (0.03) | 0.05 (0.03) | 0.16*** (0.04) | 0.10*** (0.02) |
| Off-farm activity | 0.37*** (0.02) | 0.49*** (0.08) | 0.46*** (0.02) | 0.43*** (0.02) | 0.39*** (0.04) |
| Government support | 0.62*** (0.08) | 0.94*** (0.13) | 0.32** (0.11) | 0.72*** (0.09) | 0.67*** (0.07) |
| Mean cost efficiency | 0.81 | 0.71 | 0.90 | 0.75 | 0.83 |

**Notes:** *The second-order parameters in the TL are dropped, to save space, but are available from the authors on request. Standard errors in parentheses. *p < 0.05; **p < 0.01; ***p < 0.001*
farming period, (2001–2013), was higher than in the previous decade (1991–2000). We found no significant difference between regions, but the contribution of land tenure increased over time. Perhaps scale economies increased over time with the spread of larger and more powerful farm machinery. We found no statistical difference in the positive contribution of off-farm activities between regions or over time. However, the contribution of government support to CE was higher in the central region (0.94) and increased for the years 2001–2013 (0.72). The reasons for these differences are unclear at present.

7. Conclusion and implication of the study
The aim of this study was to investigate whether farm management and socioeconomic factors contributed to improving the performance of crop farms in Norway. We found that farm resources were widely used sub-optimally, i.e., there are farms that produced lower outputs from the inputs they used or used more inputs to produce the same output, compared to the best performing farms.

The findings revealed that agronomic and socioeconomic factors affect the cost-efficient level of crop production. A good agronomic practice – crop rotation – was found to have made a positive and statically significant contribution to reducing cost. Moreover, the magnitude of crop rotation contribution was higher in the central region and increased over time in all regions. Off-farm activity positively and significantly enhanced the performance of the farms, and farmers renting land are more cost-efficient than those not doing so. Our analysis also confirms that government support has an association with improved CE, probably by relaxing farmers’ financial and liquidity constraints, enabling them to purchase new technologies that can enhance farm crop output. We also found that farmers renting land is more cost-efficient than those not doing so.

7.1 Implications of the study
One implication for farmers (and their advisers) is that many of them are less efficient than the estimated benchmark (best performing farms). Thus, those lagging behind the best performing farms need to look at the way they are operating and to seek out ways to save costs or increase crop production. Perhaps there are things for lagging farmers to learn from their more productive farming neighbors. For instance, those farmers not practicing crop rotation might be well advised to try that practice.

More cost-efficient farms likely to be financially sustainable, and also that more cost-efficient farms will contribute to the goal of food self-sufficiency. Thus, policymakers need to help farmers improve their efficiency. This can be done by facilitating the distribution and sharing of information on good farm management. Since it appears that farmers’ management is improved with off-farm experience and/or income, it suggests a need for more focus on the benefits of off-farm work. Policies that promote the development of non-farm businesses and other employment opportunities in rural areas are therefore important to enable more farm people to get local jobs. Since subsidies have a positive effect on efficiency, policymakers should maintain or extend subsidies that help farmers to invest and innovate. Most Norwegian farms are small and several policy measures are in place that holds back structural change, making it hard for operators to reap the scale economies we expect to exist in mechanized crop farming. Policies that limit the transfer of farmland by sale might be reconsidered if improving CE is considered more important than perpetuating a traditional farming culture. Policymakers might also usefully seek and implement measures to facilitate successful farmers sharing their experience with others.

For both taxpayers and consumers, one implication is that the contributions they pay that go to subsidize farmers appear to bring some benefits in terms of more efficient production that, in turn, increase the supply of some foods so possibly making food prices more affordable.
Researchers need to seek out data to explore further the causes of differences in efficiency between farms in order to better identify ways of closing the gaps. Our results show that there is a need to gather more data to be able to identify more farming practices, besides crop rotation, that lead to more efficient production. Similarly, more studies are needed to find out what the farmers in the central region do well so that successful practices might be extended farther in the eastern region.

It appears that the structure of crop farming in Norway, with a preponderance of relatively small farms, is likely to be having an adverse effect on efficiency. That effect may be partially overcome via a healthy rental market. It may well be useful to investigate the operation of this market to see whether any impediments to land renting exist that might be removed and whether there are issues about the security of tenure that need to be addressed.

Farmers are getting different types of support from the government and our study does not account for the different effects of different kinds of subsidy on CE. Different subsidies might have different effects on farm performance. To get more informative and useful results, it would be necessary to repeat the analysis with less aggregated data on subsidy payments.

References


NIBIO (2016), Resultatkontroll for gjennomføringen av landbrukspolitikken: Budgettnemnda for jordbruket, Norwegian Institute of Bioeconomy Research, Oslo (In Norwegian).


Further reading


About the authors

Habtamu Alem (PhD in Economics) is Researcher in the Division for Food Production and Society, Department of Business Economics and Management, Norwegian Institute of Bioeconomy Research (NIBIO), Norway. Habtamu Alem is the corresponding author and can be contacted at: habtamu.alem@nibio.no

Gudbrand Lien is Professor of Economics at Faculty of Economics and Organization Science, Inland Norway University of Applied Sciences, Lillehammer, Norway and he is also working as Senior Researcher in NIBIO, Norway.

J. Brian Hardaker is Emeritus Professor of Agricultural Economics at the University of New England, Armidale, NSW, Australia.

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com
Performance analysis and strategic planning of dairy supply chain in Indonesia

A comparative study

Aries Susanty, Arfan Bakhtiar, Nia Budi Puspitasari and Della Mustika
Department of Industrial Engineering, Universitas Diponegoro, Semarang, Indonesia

Abstract

Purpose – The purpose of this paper is to measure and evaluate the performance of the relationships between farmers, dairy cooperatives and industrial milk processors.

Design/methodology/approach – Data used in this study were primary data collected through personal interviews and closed questionnaires with 1–5 Likert scale. The sample consisted of the representative of the management of 12 dairy cooperatives located in Central Java Province, representative of the management of 12 dairy cooperatives located in West Java Province and some farmers who are members of those dairy cooperatives. This study uses balanced supply chain management scorecard for measuring the performance of dairy supply chain, importance-performance analysis (IPA) for identifying the indicators that are most in need of improvement, and strength, weakness, opportunity and threat (SWOT) analysis for formulating strategic planning.

Findings – The results of balanced supply chain management scorecard combined with IPA analysis showed that the performance relationship between farmers, dairy cooperatives and industrial milk processors in West Java Province is slightly better than that in Central Java Province. It can be seen from the average value of the score of indicator, the category of each indicator and the category of the performance index of each relationship. The major weakness of the relationship between dairy farmers, cooperatives and industrial milk processors in Central Java Province lies in the different perspective (no perspective is dominant), whereas that in West Java Province is dominated by the perspective of the customer. On the other hand, the major strength of the relationship in Central Java Province is dominated by the perspective of the customer, whereas that in West Java Province is dominated by the perspective of learning and growth.

Research limitations/implications – The limitation of this study is related to the number of the dairy cooperatives as the sample and the type of scale used to measure the performance of the relationships between farmers, dairy cooperatives and individual milk processors. So, the future research may replicate this study by surveying all the dairy cooperatives in Central Java and West Java Provinces. It may also enhance the measurement of the performance of the relationships by using a direct measure of each indicator in each perspective, rather than recording the management of dairy cooperative perceptions.

Practical implications – This research provides essential insights for the management of dairy cooperative in the context of strategic planning development. The research reveals that there is a different strategic planning for improving the performance of the relationship between dairy farmers, cooperatives and industrial milk processors in each province. It depends on the major weakness and strength of the relationships, and also, opportunity and threat faced by the dairy industry. One important thing, the management of dairy cooperative in both provinces should have strategic planning related to the use of machine milking by farmers to improve the milk quality.

Social implications – The research revealed that strategic planning could be built after analyzing the internal and external conditions carefully. It may encourage more dairy cooperatives to measure and analyze the internal and external conditions at the bottom of strategic planning of their business.

Originality/value – Although this research only used the balanced supply chain management scorecard and IPA analysis for measuring the performance, and SWOT analysis in formulating the strategic planning for improving the current performance, it will make a difference. First, instead of measuring the performance of dairy cooperatives, this research measured the performance of the relationships between dairy farmers, cooperatives and industrial milk processors. This way, the dairy cooperatives were only sources of data collection. Second, the investigation was quite complicated since the objects of the research were represented by the relationships between farmers, dairy cooperatives and industrial milk processors in Central Java Province and West Java Province.

Keywords Performance, Balanced scorecard, Supply chain management

Paper type Research paper
1. Introduction

Neely et al. (2000) defined a performance measurement system as the set of metrics used to quantify the efficiency and effectiveness of actions. Research on performance measurement mostly focused on the single company. However, in the last few years, the focus has shifted to incorporate a supply chain perspective with several performance measurement systems proposed. In this case, performance measurement is helpful in the improvement of the performance of the supply chain (Chan and Qi, 2002). Supply chain performance measurement can be seen as a set of metrics indicating how well the supply chain system is functioning. Measuring the supply chain performance can facilitate greater understanding of the supply chain and improve its overall performance (Charan et al., 2008). Designing an effective supply chain performance, which can evaluate the performance of the whole supply chain, is important due to a number of reasons namely increase in middleman income, customer value orientation, global competition, stakeholder needs, technology and international rules and regulations (Shashi and Singh, 2015). Besides, supply chain consists of different levels, namely supplier, manufacturer, distributor and consumer, and it is a network of companies influencing each other and affecting one another’s performance (Bigiardi and Bottani, 2010).

Measurement of supply chain performance was introduced in the 1990s, which is based on time and inventories. Levy (1995) introduced performance measures such as average finished goods inventory and demand fulfillment. Christopher (1992) introduced supply chains performance measures such as order cycle time, order completeness and delivery reliability. Meanwhile, delivery performance, lead time, level of defects and responsiveness were Lambert’s and Sharma’s approach to supply chain performance measures (Lambert and Sharma, 1990). Cohen and Lee (1990) introduced material inventory, work in process inventory, finished goods inventory, and fill rates, stock out frequencies and lead time measures. Davis presented inventory levels, inventory investment, order fill rate, line item fill rate and an average number of day’s late measures.

In the 2000s, the measurement of supply chain management performance has used different approaches. Shepherd and Günter (2006) categorized studies on this topic into operational, design and strategic research. Operational research develops mathematical models for improving supply chain performance (Lin et al., 2005; Smith et al., 2005). Design research focuses on optimizing performance through redesigning the supply chain (Shepherd and Günter, 2006). It can be categorized according to the type of research models, such as deterministic analytical models (Chen et al., 2005), stochastic analytical models (Chiang and Monahan, 2005), economic models (Wu, 2005) and simulation models (Hwarng et al., 2005; Reiner, 2005). Meanwhile, strategic research evaluates how to align the supply chain with a firm’s strategic objectives (Balasubramanian and Tewary, 2005). In this category, some authors have been using balanced scorecard (BSC) approach as an appropriate tool for the measurement. Also, a modified version of the BSC has been used by some researchers, such as Brewer and Speh (2001), Kleijnen and Smits (2003) and Park et al. (2005). This modified version is named by Park et al. (2005) as balanced supply chain management scorecard (balance SCM scorecard).

In this research, balanced supply chain management scorecard has been applied in measuring the performance of the relationships in the dairy supply chain. Moreover, related with strategic objective, this research has utilized the results of performance measurement and combined with importance-performance analysis (IPA) as a starting point to formulate some strategic planning for the indicators that are in need of improvement the most. In this case, strength, weakness, opportunity and threat (SWOT) analysis has been used as a method for formulating the strategic plans. The dairy supply chain is chosen as the context for the measurement because its performance has received a great deal of attention in the last decade, due to issues related to food self-sufficiency and dairy supply chain need to formulate the strategy to achieve the target established by the Government of Indonesia.
To achieve food self-sufficiency, the Indonesian Coordinating Ministry of Economy launched the Road Map of Indonesian Dairy 2015–2025 in February 2014. According to the roadmap, the production of milk should achieve 2.75 million tons in 2020 and 5.32 milk tons in 2025. Besides, dairy cattle should achieve 1.3m heads, which will produce an average daily production of 13.11 liters of milk per day in 2020 and dairy cattle should achieve 1.7m heads, which will produce an average daily production of 19.67 liters per day in 2025 (Wright and Meylinah, 2014). It has become apparent that shortly the dairy supply chain in Indonesia will need to formulate a strategy to achieve the target and also to systematically identify the most appropriate metric that has a high impact on the target achievement. Then, the relationship between individual farmers, dairy cooperatives and industrial milk processors has become the focus of the measurement because of the dominance of those three actors in the supply side of the dairy supply chain. The amount of individual dairy farmers in Indonesia has reached 192,160 farmers (Morey, 2011). The majority milk production from dairy farmers is marketed through local dairy cooperatives, and then they sell it to the industrial milk processors (Susanty et al., 2017).

Shortly, to implement balanced supply chain management scorecard in measuring the performance of the relationship between individual farmers, dairy cooperatives and industrial milk processors in the dairy supply chain and also to propose some strategy to achieve the target of the government, this study has several objectives. First, this study aims to measure current performance of the relationship between individual farmers, dairy cooperatives and industrial milk processors with the proposed balance SCM scorecard. Second, this study aims to develop IPA to identify the indicators that are most in need of improvement. The last, this study aims to utilize SWOT analysis to formulate strategic plans. So, the research questions are as follows:

RQ1. How would the current performance of the relationships between individual farmers, dairy cooperatives, and industrial milk processors be if measured by balance SCM scorecard?

RQ2. Which indicators in the relationships between individual farmers, dairy cooperatives, and industrial milk processors are in need of improvement the most?

RQ3. What strategic plan should be formulated to improve the performance of the relationships between individual farmers, dairy cooperatives, and industrial milk processors?

The remainder of the paper is structured as follows. The next section describes the literature review about various metrics used in balanced supply chain management scorecard and is followed by the discussion of the research methodology, such as the objects of the research, instrument and measurement, data collection procedures and data processing techniques. The results are discussed subsequently. Finally, the theoretical and managerial implications and the limitations of the study are presented in conclusion, along with the future research directions.

2. Literature review

BSC was introduced by Kaplan and Norton in the early 1990s. The BSC is considered as the method to operationalize a business’s vision and strategy. The BSC is seen as an answer to the limitations of using the measurements system that is historically based and cannot be used to assist the manager moving forward (e.g. financial ratios). The BSC tries to take objective value to the non-financial measures such as customer satisfaction and operational processes. In detail, there are four broad perspectives associated with BSC, namely, customer, internal business processes, innovation and learning or leaning and growth and financial perspectives (Kaplan and Norton, 1996; Mathiyalagan et al., 2014).

According to several authors, the four perspectives of BSC are suitable for overcoming the problems related to performance measurement in the supply chain. The modified...
version of a BSC for measuring supply chain performance is named by Park et al. (2005) as balanced supply chain management scorecard (balanced SCM scorecard). Besides, studies exploring the application of the BSC as a performance measurement system for the performance management of SCs are cited in several other studies, such as those of Brewer and Speh (2000, 2001), Kleijnen and Smits (2003), among others. Although the application of the BSC as a performance measurement system for SCM can be found in several authors, there is no consensus about the indicators used in each perspective on the balanced SCM scorecard. In detail, several indicators in each perspective which are used by previous researchers can be seen in Table I.

3. Methods of the research

Object of the research
The objects of this research are dairy milk supply chains in Provinces of Central Java and West Java. These provinces are two among three provinces as the largest milk producers in Indonesia. According to Morey (2011), there are 97,589 cows in Central Java, located in Boyolali and Semarang regions producing 77.1 tons of milk per year. In West Java, there are 58,001 cows, located in Bandung, Lembang, Pangalengan, Sumedang, Kuningan and Garut regions producing 126,221 tons of milk per year. The sample of this research is represented by 24 dairy cooperatives consisting of 12 cooperatives out of 24 active dairy cooperatives located in Boyolali and Semarang and 12 cooperatives among 18 active dairy cooperatives located in Lembang, Pangalengan, Sumedang and Garut. Although the sample size is only 24, those cooperatives can represent the condition of the relationship between farmers, cooperatives, and industrial milk processors in every province since 80 percent of the milk produced from Central Java and West Java Provinces comes from those cooperatives. In each of the selected dairy cooperative, one person from the management has represented the cooperative as the source of information. Moreover, one of the dairy farmers among the members of each selected dairy cooperative has also been chosen to represent farmers.

This research has used non-probability purposive sampling for selecting the dairy cooperative and also for selecting the representative of management. Hence, the selection of dairy cooperative and representative of the management of each dairy cooperative as the sample is based upon certain appropriate characteristics. For dairy cooperatives, this research chooses the dairy cooperative belonging to 80 percent of the milk producer in Central Java and West Java Provinces. Concerning the characteristics of the representatives of management, they should be within the cooperative for more than five years, or they should be in the management position for more than three years. The inclusion characteristics of the representatives of management for filling out the questionnaire were as follows: they should have the knowledge about the condition of relationship between dairy farmers, cooperatives and industrial milk processor; they should have the knowledge about the current condition of the cooperative; and, they should have time to participate in the research. From each of cooperative, this research will collect the data needed for measuring the current condition of the relationship between dairy farmers, cooperatives and industrial milk processors which is where the cooperative sells its milk. In this case, the selected dairy cooperatives mentioned previously have been the focal point of data collection because they have intent relationships with individual dairy farmers as members of the cooperatives. The selected dairy cooperatives also have determined relationships with industrial milk processors since they market their milk to the processors. The information from the dairy farmers was used as the other source to validate some information received from the cooperatives.

Instrument and measurement
A total of 28 indicators have been used in this research. They were used to measure the four perspectives, i.e., customer (ten indicators), financial (six indicators), internal business process (four indicators) and learning and growth (eight indicators) perspectives. For the
<table>
<thead>
<tr>
<th>Perspective</th>
<th>Reference</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Financial</strong></td>
<td>Santos et al. (2006)</td>
<td>Material acquisition costs; non-quality costs; warehousing costs; manufacturing unit costs; cost of carrying inventory; logistics cost; transportation costs; cash flow; EBITDA; income; EVA (economic value added); operating ratio; return on investment- ROI; revenue per employee; and return on asset</td>
</tr>
<tr>
<td></td>
<td>Bhagwat, and Sharma (2007)</td>
<td>Net profit vs productivity ratio rate of return on investment; variations against budget; buyer-supplier partnership level; delivery performance; supplier cost-saving initiatives; delivery reliability; cost per operation hour; information carrying cost; and supplier rejection rate</td>
</tr>
<tr>
<td></td>
<td>Bigliardi and Bottani (2010)</td>
<td>Information carrying cost; supplier cost-saving activities; variations against budget; cost per operation hour; and return on investment</td>
</tr>
<tr>
<td><strong>Customer</strong></td>
<td>Santos et al. (2006)</td>
<td>Quality – % non-conformity; forecast accuracy; market share; on-time delivery; number of products/distribution channel; and damaged shipments</td>
</tr>
<tr>
<td></td>
<td>Bhagwat, and Sharma (2007)</td>
<td>Customer query time; level of customer perceived value of product; range of products and services; order lead time; flexibility of service systems to meet particular customer needs; buyer-supplier partnership level; delivery lead time; delivery performance; effectiveness of delivery invoice methods; delivery reliability; responsiveness to urgent deliveries; effectiveness of distribution planning schedule; information carrying cost; quality of delivery documentation; driver reliability for performance; quality of delivered goods; and achievement of defect-free deliveries</td>
</tr>
<tr>
<td></td>
<td>Bigliardi and Bottani (2010)</td>
<td>Customer query time; order lead time; distribution lead time; distribution performance; delivery reliability; effectiveness of distribution planning schedule; quality of delivery goods; customer perceived value of product’ flexibility of service system to meet particular customer needs, and responsive to urgent delivery</td>
</tr>
<tr>
<td><strong>Internal business</strong></td>
<td>Santos et al. (2006)</td>
<td>Supplier on-time delivery; material inventories; material quality; supplier cycle time; % of orders delivered according to plan; schedule changes; BOM accuracy; adherence to schedule; % defect products; number of finished products;SKUs; manufacturing cycle time; setups/ changeovers; plant utilization; finished goods inventory turnover; stock keeping units</td>
</tr>
<tr>
<td></td>
<td>Bhagwat, and Sharma (2007)</td>
<td>Total supply chain cycle time; total cash flow time; flexibility of service systems to meet particular customer needs; supplier lead time against industry norms; level of supplier’s defect-free deliveries; accuracy of forecasting techniques; product development cycle time; purchase order cycle time; planned process cycle time; effectiveness of master production schedule; capacity utilization; total inventory cost; supplier rejection rate; efficiency of purchase order cycle time; and frequency of delivery</td>
</tr>
<tr>
<td></td>
<td>Bigliardi and Bottani (2010)</td>
<td>Accuracy of forecasting technique; planned process cycle time; purchase order cycle time; effectiveness of master production schedule; supplier rejection rate; total inventory cost; and frequency of delivery</td>
</tr>
<tr>
<td><strong>Innovation and learning</strong></td>
<td>Santos et al. (2006)</td>
<td>% new product development; social programs investments; absenteeism; % employee training; employee productivity; motivation; and employee turnover</td>
</tr>
<tr>
<td></td>
<td>Bhagwat and Sharma (2007)</td>
<td>Supplier assistance in solving technical problems; supplier ability to respond to quality problems; supplier cost-saving initiatives; supplier’s booking in procedures; capacity utilization; order entry methods; accuracy of forecasting techniques; product development cycle time; flexibility of service systems to meet particular customer needs; buyer-supplier partnership level; range of products and services; and level of customer perceived value of product</td>
</tr>
<tr>
<td></td>
<td>Bigliardi and Bottani (2010)</td>
<td>Supplier assistance in solving technical problem; supplier ability to respond to quality problem, buyer-supplier collaboration in problem-solving; order entry method; and level of information sharing</td>
</tr>
</tbody>
</table>

Table I: Indicators in each perspective from previous studies
customer perspective, the indicators are developed from Hong and Zhong-Hua (2013), Prakash and Pant (2013), National Standard Indonesia (2011), Wright and Meylinah (2014), Callado and Jack (2015) and Susanty et al. (2017). As for the financial perspective, the indicators are developed from Callado and Jack (2015), Hong and Zhong-Hua (2013) and Prakash and Pant (2013). The indicators for measuring the perspective of the internal business process are developed from Prakash and Pant (2013). Finally, for measuring the perspective of learning and growth, the indicators are developed from Prakash and Pant (2013), Callado and Jack (2015) and Susanty et al. (2017). Some indicators in this research were also developed based on the results of interviews with the representatives of management of the dairy cooperatives.

Based on those 28 indicators, this study has used two types of closed questionnaire:

- The first type is the analytical hierarchy process (AHP) questionnaire. This questionnaire is used to compare the level of importance of each perspective (the perspectives of the customer, financial, internal business process, and learning and growth) and the level of importance of each indicator which belongs to each of the perspectives. The results of this questionnaire have indicated the relative weight of each perspective and indicator that contributes to the relationships between dairy farmers, cooperatives and industrial milk processors. The Saaty’s nine-point scale has been used for the first type of questionnaire, ranging from 1 (equal importance between element i and j) to 9 (absolute dominance of me over j), and reciprocal values, respectively. All values within the range of 1 to 9 and 1/9 to 1 are possible; the respondents were not restricted to the integer data points 1, 2, etc., and their reciprocals (Saaty, 1995).

- The second questionnaire has been used to measure the current condition of each indicator. The five-point Likert scale was used for the purpose. Although the higher the score, the better the condition (1 = the worst condition and 5 = the best condition), the five-point Likert scale used in the study may have different meaning depending on the condition asked on each indicator. As an example, the meaning of value 1 to 5 for the questionnaire “the level of conformity of total plate count (TPC) contained in the milk delivered by the farmers with the limit set by the Indonesian National Standard,” can be described as follows. Value 1 means the level of TPC is between 800,001 CFU/ and 1,000,000 CFU/mL; value 2 means the level of TPC is between 600,001 CFU/ and 8,000,000 CFU/mL; value 3 means the level of TPC is between 400,001 CFU/ and 600,000 CFU/mL; value 4 means the level of TPC is between 200,001 CFU/ and 400,000 CFU/mL; and value 5 means the level of TPC is between 1 CFU/ and 200,000 CFU/mL. In detail, list of indicators and their scale can be seen in Table II. The results of the second questionnaire indicate the relative strength and weakness or the relative condition of the internal factor of the relationship between dairy farmers, cooperatives and industrial milk processors. Moreover, based on those internal factors and combined with threat and opportunity faced by dairy industry or external factors, some strategies will be formulated using SWOT analysis. Specifically, the strategy will focus on the indicators that are most in need of improvement.

Data collection procedure
This study has used both primary and secondary data. The primary sources of data were questionnaire and personal interviews. The 24 copies of the first and second type of questionnaire were administered to the representatives of the management of dairy cooperatives in Semarang, Boyolali, Bandung, Lembang, Pangalengan, Cirebon and Garut. Besides the questionnaire, personal interviews have been conducted with the
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Indicators</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perspective of customers (CP)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer satisfaction (CP1) (Hong and Zhong Hua, 2013)</td>
<td>Level of satisfaction of dairy farmers with price offered by the cooperative (CP11) (Susanty et al., 2017)</td>
<td>1 = strongly disagree to 5 = strongly agree</td>
</tr>
<tr>
<td>(Hong and Zhong Hua, 2013)</td>
<td>Level of satisfaction of cooperatives with the commitment of dairy farmers to produce milk with specific quantity (CP12) (Susanty et al., 2017)</td>
<td>1 = strongly disagree to 5 = strongly agree</td>
</tr>
<tr>
<td>Compliance to food quality and Codex standards (CP2) (Indonesian National Standard, 2011)</td>
<td>The level of conformity of total plate count (TPC) contained in the milk delivered by the farmers with Indonesia National Standard or SNI (CP21) (Indonesian National Standard, 2011)</td>
<td>1 = $800,000 \leq \text{TPC} \leq 1,000,000 \text{ CFU/mL}$ to 5 = 1 \leq \text{TPC} \leq 200,000 \text{ CFU/mL}</td>
</tr>
<tr>
<td>(Indonesian National Standard, 2011)</td>
<td>The level of conformity of total plate count (TPC) contained in the milk delivered by the cooperatives with standard of industrial milk processor (CP22) (SNI 3141.1:2011) (Indonesian National Standard, 2011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The level of conformity of fat contained in the milk delivered by the farmers with SNI (CP23) (SNI 3141.1:2011) (Indonesian National Standard, 2011)</td>
<td>1 = Level of fat $\leq 3,000%$ to 5 = Level of fat $&gt; 4,500%$</td>
</tr>
<tr>
<td></td>
<td>The level of conformity of Solid Nonfat (SNF) contained in the milk delivered by the farmers with SNI (CP25) (Indonesian National Standard, 2011)</td>
<td>1 = Level of SNF $\leq 7,800%$ to 5 = Level of SNF $&gt; 8,200%$</td>
</tr>
<tr>
<td>(Indonesian National Standard, 2011)</td>
<td>The level of conformity of Solid Nonfat (SNF) contained in the milk delivered by the farmers with standard of industrial milk processor (CP26) (Indonesian National Standard, 2011)</td>
<td></td>
</tr>
<tr>
<td>Ease of contact (CP3) (Wright and Meylanah, 2014)</td>
<td>Level of ease of the farmers to get in touch with the cooperatives (CP31) (Susanty et al., 2017)</td>
<td>1 = It is very difficult for farmers to get contact with the cooperatives, even they have critical problem to solve 5 = The farmers can get contact with the cooperatives easily</td>
</tr>
<tr>
<td></td>
<td>Duration of farmers to become a member of cooperatives (from interviews with the cooperatives) (CP41)</td>
<td>1 = less than 5 years to 2 = more than 20 years</td>
</tr>
<tr>
<td><strong>Perspective of financial (PF)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profitability (PF1) (Callado and Jack, 2015; Hong and Zhong-Hua, 2013)</td>
<td>Percentage of profit sharing received by the farmers from selling their milk to cooperatives (From interviews with the cooperatives) (PF11)</td>
<td>1 = less than 1% to 5 = more than 100%</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Indicators</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of profit sharing received by the</td>
<td>Prices and conditions offered by dairy cooperative do not change unexpectedly and it can be seen from the frequency of price changes offered by the cooperatives to the farmers (From interviews with the cooperative) (PF21)</td>
<td>1 = more than four times in a year to 5 = cooperatives never change the price offered</td>
</tr>
<tr>
<td>cooperatives from selling their milk to</td>
<td>Prices and conditions offered by industrial milk processor do not change unexpectedly and it can be seen from the frequency of price changes offered by the industrial processing milk to the cooperatives (From interviews with the cooperatives) (PF22)</td>
<td></td>
</tr>
<tr>
<td>industrial milk processors (from interviews</td>
<td></td>
<td>1 = less than IDR 4,000 to 5 = more than IDR 5,000</td>
</tr>
<tr>
<td>with the cooperatives) (PF12)</td>
<td></td>
<td>1 = less than IDR 5,000 to 5 = more than IDR 6,000</td>
</tr>
<tr>
<td>Prices and conditions offered by dairy</td>
<td>Time span between cooling process in the dairy cooperative and milking process (BP11) (Thai Agricultural Standard, 2008)</td>
<td>1 = more than 120 min to 5 = less than 30 min</td>
</tr>
<tr>
<td>cooperative do not change unexpectedly and</td>
<td>The number of types of quality checking of milk conducted by the cooperative before they sent it to the industrial milk processor (BP21) (from interviews with the cooperatives)</td>
<td>1 = less than three type of quality checking to 5 = more than 13 types of quality checking</td>
</tr>
<tr>
<td>it can be seen from the frequency of price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>changes offered by the cooperatives to the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>farmers (From interviews with the cooperative)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The prices per liter the farmers get from</td>
<td>Implementation of HACCP and other quality control measures by the farmers (BP31) (From interviews with the cooperatives)</td>
<td>1 = no implementation of HACCP and other quality control measures to 2 = have been implementing the HACCP and other quality control measures</td>
</tr>
<tr>
<td>the cooperative according to the quality of</td>
<td>The level of implementation of HACCP and other quality control measures by the cooperatives (BP32) (from interviews with the cooperatives)</td>
<td></td>
</tr>
<tr>
<td>their milk (PF31) (From interviews with the</td>
<td></td>
<td>1 = no information sharing to 5 = more than four times in a years, conduct regular meeting for sharing the information</td>
</tr>
<tr>
<td>Dimension</td>
<td>Indicators</td>
<td>Scale</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td><strong>Level of collaboration in problem-solving between the cooperatives and industrial</strong></td>
<td><strong>5</strong></td>
</tr>
<tr>
<td></td>
<td><strong>processing milk (LG22)</strong> (from interviews with the cooperatives)</td>
<td><strong>4</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Frequency of training for capacity building from cooperatives to farmers in a year (LG31)</strong> (from interviews with the cooperatives)</td>
<td><strong>3</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Frequency of training for capacity building from industrial milk processor to the cooperatives in a year (LG32)</strong> (from interviews with the cooperatives)</td>
<td><strong>2</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Level of sophistication of equipment used by the farmers for milking process (LG41)</strong> (Budiyanto and Usmiati, 2008; Woolford et al., 2004)</td>
<td><strong>1</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Level of sophistication of equipment used by cooperative for cooling the milk (LG42)</strong> (Roberts and Larson, 1941; Vagany and Dunay, 2004)</td>
<td><strong>0</strong></td>
</tr>
</tbody>
</table>

| Table II. Dairy supply chain in Indonesia |
management representatives for further explanation about the values of the scores in the questionnaire. Interviews have also been used to validate some indicators to the dairy farmers such as the indicator about the level of satisfaction of dairy farmers to the cooperatives, the level of information sharing, the level of collaboration, the frequency of training, etc. The secondary data consist of several documents owned by the cooperatives as a complementary of the results of the questionnaire and personal interviews.

Data processing technique
The data were analyzed using AHP (Saaty, 1995) through Expert Choice Software, Snorn de Boer (Trienekens and Hvolsby, 2000) for converting the five-point Likert Scale to the value between 1 and 100. A modified IPA was also used to identify the indicators in need of improvement the most. The application of the IPA was introduced by Martilla and James (1977).

4. Results
Profile of respondents
The profile of 24 dairy cooperatives and their representatives who filled in the questionnaire can be seen in Table III. Half of the representatives of the management were aged between 41 and 55 years old, followed by 61 and 65 years old, 56 and 60 years old, 36 and 40 years old, 60 years old or more and 31 and 35 years old. Concerning the level of education, many of the representatives have a Bachelor's degree, followed by senior high school, diploma, elementary school and only two of them hold a Master's degree. Then, regarding the duration of working with the cooperatives, many representatives have been with the cooperatives for 16 to 20 years, followed by 5 to 10 years, 11 to 15 years and 26 to 30 years, 21 to 25 years and 31 years or more.

Concerning the dairy farmers as the respondents of this research, all of them were male with aged between 45 and 50 years old, followed by 41 and 45 years old, 51 and 55 years old, 56 and 60 years old, 36 and 40 years old, 61 years old or more and 31 and 35 years old. Regarding their duration of working, many respondents have been farmers for 21 to 25 years, followed by 5 to 10 years and 16 to 20 years, 31 to 35 years, 26 to 3 years and 36 to 40 years. Then, most of the respondents have senior high school education, followed by senior high school, bachelor degree and elementary school.

The result of computing the priority weight of each perspective and indicators
After obtaining the individual pairwise judgments from the representatives of the management, the next step was computing the priority weight of each perspective and each indicator using the expert choice software. The results show the approximate priority weight of each perspective and indicator by the individual member of the group of respondents. Then, the final priority weight of each perspective and indicator should be aggregated to arrive at the consensus group. There are several methods to aggregate the opinion of several decision makers. One may choose to aggregate the individual judgments (AIJ) or the resulting priorities (AIP). The choice of methods depends on whether the group is assumed to act together as a unit or as separate individuals. For the first assumption, the geometric mean of aggregate the individual judgments (AIJ) satisfies the reciprocity requirement, implying a synergistic aggregation of individual preferences in such a way that the group becomes a new “individual” and behaves like one. Individual identities are lost with every stage of aggregation, and the Pareto principle is irrelevant. When group members act as individuals (AIP), Dong et al. (2010) gives the following formula to aggregate the priorities weight of several decision makers. Let $\mathbf{w}^{(k)} = (w_{1}^{(k)}, \ldots, w_{n}^{(k)})^{T}$ be the individual priority vector derived from individual judgment matrix $A^{(k)}$ using certain prioritization method. Then, the aggregate of the priorities weight of several decision
<table>
<thead>
<tr>
<th>No.</th>
<th>Name of cooperative (location)</th>
<th>Average milk production per month (kg)</th>
<th>Age (years)</th>
<th>Level of education</th>
<th>Duration working in cooperative (years)</th>
<th>Name of cooperative (location)</th>
<th>Average milk production per month (kg)</th>
<th>Age (years)</th>
<th>Level of education</th>
<th>Duration of working in cooperative (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mardi Mulya (Mojosongo)-Boyolali</td>
<td>998.194</td>
<td>42</td>
<td>Bachelor Degree</td>
<td>12</td>
<td>KPSBU-Lembang</td>
<td>4.378.772</td>
<td>53</td>
<td>Bachelor Degree</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>Andini Luhur-Semarang District</td>
<td>927.000</td>
<td>65</td>
<td>Bachelor Degree</td>
<td>18</td>
<td>Puspa Mekar-Bandung</td>
<td>354.200</td>
<td>40</td>
<td>Senior high school</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>Musul Boyolali</td>
<td>873.234</td>
<td>35</td>
<td>Bachelor Degree</td>
<td>15</td>
<td>Mitra Jaya-Bandung</td>
<td>186.390</td>
<td>48</td>
<td>Masters Degree</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>Wabyu Agung-Semarang District</td>
<td>710.700</td>
<td>45</td>
<td>Bachelor Degree</td>
<td>8</td>
<td>KPBS-Pangalengan</td>
<td>2.439.862</td>
<td>37</td>
<td>Senior high school</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>Gepogo-Boyolali</td>
<td>235.767</td>
<td>65</td>
<td>Bachelor Degree</td>
<td>29</td>
<td>Sinar Jaya-Bandung</td>
<td>86.361</td>
<td>50</td>
<td>Bachelor Degree</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>Getusun-Semarang District</td>
<td>235.767</td>
<td>53</td>
<td>Bachelor Degree</td>
<td>5</td>
<td>Sarwa Multi-Bandung</td>
<td>217.775</td>
<td>47</td>
<td>Bachelor Degree</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>Sumber Karya -Semarang District</td>
<td>221.028</td>
<td>41</td>
<td>Bachelor Degree</td>
<td>7</td>
<td>Tandangsari-Sumedang</td>
<td>670.065</td>
<td>56</td>
<td>Bachelor Degree</td>
<td>28</td>
</tr>
<tr>
<td>8</td>
<td>Banyumanik-Semarang City</td>
<td>117.884</td>
<td>66</td>
<td>Master Degree</td>
<td>6</td>
<td>Cikajang-Garut</td>
<td>735.940</td>
<td>65</td>
<td>Senior high school</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>Ngudi Luhur-Semarang District</td>
<td>108.150</td>
<td>49</td>
<td>Master Degree</td>
<td>18</td>
<td>Bayongbong-Garut</td>
<td>719.946</td>
<td>58</td>
<td>Bachelor Degree</td>
<td>34</td>
</tr>
<tr>
<td>10</td>
<td>Mekar-Uguran, Semarang District</td>
<td>88.405</td>
<td>63</td>
<td>Bachelor Degree</td>
<td>10</td>
<td>Suhuyu-Kuningan</td>
<td>320.103</td>
<td>51</td>
<td>Diploma</td>
<td>30</td>
</tr>
<tr>
<td>11</td>
<td>Kota-Boyolali</td>
<td>58.957</td>
<td>52</td>
<td>Bachelor Degree</td>
<td>25</td>
<td>Karya Ngrahak-Kuningan</td>
<td>877.860</td>
<td>41</td>
<td>Senior high school</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>Subur-Gunung Pati, Semarang City</td>
<td>29.479</td>
<td>58</td>
<td>Bachelor Degree</td>
<td>35</td>
<td>Larassat-Kuningan</td>
<td>150.000</td>
<td>72</td>
<td>Elementary school</td>
<td>19</td>
</tr>
</tbody>
</table>
makers obtained is $\mathbf{w}^{(c)} = (w_1^{(c)}, w_2^{(c)}, \ldots, w_n^{(c)})^T$ (Dong et al., 2010) in the following equation:

$$
w_i^{(c)} = \frac{\prod_{k=1}^{m} (w_i^{(k)})^{\lambda_k}}{\sum_{i=1}^{n} \prod_{k=1}^{m} (w_i^{(k)})^{\lambda_k}}
$$

where $w_i^{(c)}$ = the aggregate of priority weight of the indicator $i$; $k$, the priority weight from the individual decision makers $k$ ($k = 1, 2, 3, \ldots, m$); $m$, the number of decision maker; $i$, indicator $i$ ($i = 1, 2, 3, \ldots, n$); $n$, the number of indicators; $\lambda$, weight vector of decision makers (in this research, the value of $\lambda$ is equal to 1 because all decision makers have the same priority to answer the question).

This study has chosen to use AIP than AJ because the group that consists of dairy cooperatives from several districts in the Provinces of Central Java and West Java is not homogenous and the decision makers are not willing to act like one single individual. In this case, perhaps the decision context regarding the priority weight of each perspective and indicator is attended by a conflict of interest, and each dairy cooperative as the group member is individually acting with its value systems.

The results in the Province of Central Java show that the perspective of the customer has the highest priority weight (0.752), followed by the financial perspective (0.133), the perspective of learning and growth (0.105) and the perspective of the internal business process (0.01). On the other hand, those in the Province of West Java show that the perspective of learning and growth has the highest priority weight (0.9854183) followed by the perspective of the customer (0.0145565), the perspective of the internal business process (0.0000238) and the financial perspective (0.0000015). In detail, the results of the aggregation of the priority weight from the representatives of the management of dairy cooperatives can be seen in Table IV.

In the Province of Central Java, based on the rearranged priority weight in descending order, the top five ranks of the indicators are: the level of satisfaction of cooperatives with the commitment of dairy farmers to produce milk (CP12) (0.35736400600); the duration of farmers to become a member of cooperatives (CP41) (0.33878001465); frequency of training for capacity building from cooperatives to farmers (LG31) (0.10134217904); percentage of profit sharing received by the farmers from selling their milk to cooperatives (PF11) (0.07971709706); and percentage of profit sharing received by the cooperatives from selling their milk to industrial milk processors (PF12) (0.05322452827). In the Province of West Java, the top five ranks are: the level of collaboration in problem-solving between the farmers and cooperative (LG21) (0.6261700018532); the level of sophistication of equipment used by cooperative for cooling the milk (LG42) (0.17814203984952); the frequency of training for capacity building from cooperatives to farmers (LG31) (0.17328713847026); the duration of farmers to become a member of cooperatives (CP41) (0.0707715025267); and the frequency of information sharing between the farmers and cooperative (CP11) (0.0075067456926). It is apparent that the representatives of the management in both provinces agree about the importance of the duration of farmers in becoming members of the cooperatives and the frequency of the training for capacity building from the cooperatives to farmers.

Result of the performance measurement of the relationships between dairy farmers, cooperative and individual milk processor based on balanced supply chain management scorecard

The aggregate value of the performance of the relationship between dairy farmers, cooperatives and individual milk processors in each dairy cooperative is the sum of the performance index of each indicator, which represents multiplication between the score
value of indicators with the weight of the indicator. In this research, before aggregation, the score value of each indicator will be equalized using normalization process (Snorm) of DeBoer (Trienekens and Hvolby, 2000), so the measurement scale from 0 to 100 for each indicator could be obtained. After normalization process, the value of each indicator and the aggregate value of the performance can be grouped using interval value from Trienekens and Hvolby (2000) which consists of poor (score \( \leq 40 \)), marginal (40 < score \( \leq 50 \)), average (50 < score \( \leq 70 \)), good (70 < score \( \leq 90 \)) and excellent (score > 90). This interval value was used because Trienekens and Hvolby (2000) also used this interval value for measuring the performance in supply chain although they did not use balanced supply chain scorecard as a framework. The Snorm equation of DeBoer for normalization process can be seen in the following equations:

If larger is better, then \[ \text{Snorm} = \frac{(S_i-S_{\text{min}})}{(S_{\text{max}}-S_{\text{min}})} \times 100, \] (2)

If lower is better, then \[ \text{Snorm} = \frac{(S_{\text{max}}-S_i)}{(S_{\text{max}}-S_{\text{min}})} \times 100, \] (3)

where \( S_i \) is the actual score of each indicator; \( S_{\text{min}} \), the minimum score; and \( S_{\text{max}} \), the maximum score.

<table>
<thead>
<tr>
<th>Central Java Province Perspective</th>
<th>Indicators</th>
<th>Global priority weight</th>
<th>West Java Province Perspective</th>
<th>Indicators</th>
<th>Global priority weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer (0.752239)</td>
<td>CP11</td>
<td>0.03573640060</td>
<td>Customer (0.0145565)</td>
<td>CP11</td>
<td>0.0042320811240</td>
</tr>
<tr>
<td></td>
<td>CP12</td>
<td>0.35736406000</td>
<td></td>
<td>CP12</td>
<td>0.000000000052714</td>
</tr>
<tr>
<td></td>
<td>CP21</td>
<td>0.00501587722</td>
<td></td>
<td>CP21</td>
<td>0.00008220212712</td>
</tr>
<tr>
<td></td>
<td>CP22</td>
<td>0.00001052400</td>
<td></td>
<td>CP22</td>
<td>0.000000549323664</td>
</tr>
<tr>
<td></td>
<td>CP23</td>
<td>0.00587329651</td>
<td></td>
<td>CP23</td>
<td>0.00000000072848</td>
</tr>
<tr>
<td></td>
<td>CP24</td>
<td>0.00014568638</td>
<td></td>
<td>CP24</td>
<td>0.00000000039282</td>
</tr>
<tr>
<td></td>
<td>CP25</td>
<td>0.00819161620</td>
<td></td>
<td>CP25</td>
<td>0.00000000000008</td>
</tr>
<tr>
<td></td>
<td>CP26</td>
<td>0.00023143686</td>
<td></td>
<td>CP26</td>
<td>0.00000000000001</td>
</tr>
<tr>
<td></td>
<td>CP31</td>
<td>0.00079011918</td>
<td></td>
<td>CP31</td>
<td>0.00115321724745</td>
</tr>
<tr>
<td></td>
<td>CP41</td>
<td>0.323678201465</td>
<td></td>
<td>CP41</td>
<td>0.00097150252679</td>
</tr>
<tr>
<td>Financial (0.133117655)</td>
<td>PF11</td>
<td>0.075917097062</td>
<td>Financial (0.0000015)</td>
<td>PF11</td>
<td>0.0000000002582898</td>
</tr>
<tr>
<td></td>
<td>PF12</td>
<td>0.053224528270</td>
<td></td>
<td>PF12</td>
<td>0.000000000052713</td>
</tr>
<tr>
<td></td>
<td>PF21</td>
<td>0.000000000940</td>
<td></td>
<td>PF21</td>
<td>0.00000000000019</td>
</tr>
<tr>
<td></td>
<td>PF22</td>
<td>0.000018441490</td>
<td></td>
<td>PF22</td>
<td>0.0000000001218888</td>
</tr>
<tr>
<td></td>
<td>PF31</td>
<td>0.000011788870</td>
<td></td>
<td>PF31</td>
<td>0.000000000034158</td>
</tr>
<tr>
<td></td>
<td>PF32</td>
<td>0.0000392986230</td>
<td></td>
<td>PF32</td>
<td>0.00000014802186</td>
</tr>
<tr>
<td>Internal business process (0.00958233)</td>
<td>BP11</td>
<td>0.00069767350</td>
<td>Internal business process (0.0000623)</td>
<td>BP11</td>
<td>0.0000002510941</td>
</tr>
<tr>
<td></td>
<td>BP21</td>
<td>0.00053814120</td>
<td></td>
<td>BP21</td>
<td>0.00000000028318</td>
</tr>
<tr>
<td></td>
<td>BP31</td>
<td>0.000000000560</td>
<td></td>
<td>BP31</td>
<td>0.000000237327518</td>
</tr>
<tr>
<td></td>
<td>BP32</td>
<td>0.000004446480</td>
<td></td>
<td>BP32</td>
<td>0.00000000000815</td>
</tr>
<tr>
<td>Learning and growth (0.104687108)</td>
<td>LG11</td>
<td>0.00333633102</td>
<td>Learning and growth (0.9854183)</td>
<td>LG11</td>
<td>0.00075067459262</td>
</tr>
<tr>
<td></td>
<td>LG12</td>
<td>0.00000372486</td>
<td></td>
<td>LG12</td>
<td>0.00000000002255</td>
</tr>
<tr>
<td></td>
<td>LG21</td>
<td>0.000000000970</td>
<td></td>
<td>LG21</td>
<td>0.62617000185320</td>
</tr>
<tr>
<td></td>
<td>LG22</td>
<td>0.000000000870</td>
<td></td>
<td>LG22</td>
<td>0.000000000519384</td>
</tr>
<tr>
<td></td>
<td>LG31</td>
<td>0.010342179040</td>
<td></td>
<td>LG31</td>
<td>0.17328713847026</td>
</tr>
<tr>
<td></td>
<td>LG32</td>
<td>0.00000000057110</td>
<td></td>
<td>LG32</td>
<td>0.00000001415786</td>
</tr>
<tr>
<td></td>
<td>LG41</td>
<td>0.00000000054747</td>
<td></td>
<td>LG41</td>
<td>0.000031229066579</td>
</tr>
<tr>
<td></td>
<td>LG42</td>
<td>0.000000408577</td>
<td></td>
<td>LG42</td>
<td>0.17814203984952</td>
</tr>
</tbody>
</table>

Table IV. The results of aggregation of the priorities weight from several representatives of management of dairy cooperative in the provinces of Central Java and West Java.
Larger-is-better is used when the higher observed value represents better performance, as in the case of indicator CP11 (level of satisfaction of dairy farmers with the price offered by the cooperative). Below is an example of the calculation to get the measurement scale of 0–100 of the indicator CP11 for the Dairy Cooperative of Mojosongo in Boyolali with the larger-is-better in the following equation:

\[
\text{Larger is better } : \quad S_{\text{norm}} = \frac{S_i - S_{\text{min}}}{S_{\text{max}} - S_{\text{min}}} \times 100, \tag{4}
\]

where \((4 - 1) \times 100 = 75\).

On the other hand, lower-is-better is used when the lower value represents better performance. In this study, lower-is-better is not used to get the measurement scale of 0–100 of the indicators since the best condition of each indicator is achieved when the indicator can get the highest value. In detail, the score of each indicator from each dairy cooperative in Central Java and West Java Provinces after converted to the measurement scale of 0–100 can be seen in Tables V and VI. Then, Tables VII and VIII show the performance index of each indicator (multiplication between the score value of the indicator with its priority weight) and the aggregate value of the performance of the relationship in each surveyed dairy cooperative.

Table V to Table VII indicate that the relationship between dairy farmers, cooperatives and industrial milk processor in West Java Province is better than that in Central Java. It can be seen from the mean score of all indicators, i.e., 55.43 in Central Java Province and 58.75 in West Java Province. Moreover, Central Java Province is also less in the number of indicators categorized as good and excellent compared with West Java Province. In Central Java Province, only 21.43 percent of the indicators belong to good and excellent categories, and the rest (78.57 percent) is in the categories of poor, marginal, and average. Meanwhile, in West Java Province, 39.29 percent of the indicators are categorized as good and excellent, while 60.71 percent of the indicators belong to the categories of poor, marginal and average.

In line with the results shown in Tables V and VI, those of the calculation of an aggregate value of the performance of the relationships between farmers, cooperatives and industrial milk processors shown in Tables VII and VIII also indicate similar finding. The mean aggregate value in Central Java Province is only 53.97 (belong to average category), whereas that in West Java Province achieves 87.02 (belong to good category). Only two dairy cooperatives in West Java Province have the aggregate value less than 70, while 10 dairy cooperatives have the aggregate value more than 70. It means only 2 from 12 relationships are categorized as average, while the other 10 relationships belong to the good and excellent categories. On the contrary, there are 10 dairy cooperatives in Central Java Province which have the aggregate value below 70, and only 2 dairy cooperatives have the aggregate value more than 70. No dairy cooperative in Central Java Province has the aggregate value more than 90. It means that no dairy cooperative is categorized as excellent.

Result of the IPA of the relationships between dairy farmers, cooperative and individual milk processor

IPA, which is an important and applicable tool for mapping the condition of each indicator based on its performance (x-axis) and importance (y-axis), has been used for determining the indicators that are most in need of improvement. The performance of each indicator is expressed by its score, whereas its importance is expressed by its priority weight. Moreover, the median values of the score and priority weight are used as coordinates for plotting individual indicators on a two-dimensional matrix which has four quadrants (concentrate here, keep up with the good work, low priority, and possible overkill) as shown in Figures 1 and 2. The median values as a measure of central tendency of score and priority weight are theoretically preferable to the means because a true interval scale may not exist.
<table>
<thead>
<tr>
<th>No.</th>
<th>Indicators</th>
<th>Mojosongo</th>
<th>Musuk</th>
<th>Cepogo</th>
<th>Kota</th>
<th>Banyumanik</th>
<th>Subur</th>
<th>Mekar</th>
<th>Sumber</th>
<th>Karya</th>
<th>Getasan</th>
<th>Wahyu Agung</th>
<th>Andini Luhur</th>
<th>Ngudi Luhur</th>
<th>Mean</th>
<th>SD</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CP11</td>
<td>50</td>
<td>75</td>
<td>75</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>66.67</td>
<td>18.63</td>
<td>Average</td>
</tr>
<tr>
<td>2</td>
<td>CP12</td>
<td>50</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>25</td>
<td>50</td>
<td>100</td>
<td>75</td>
<td>25</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>64.58</td>
<td>27.87</td>
<td>Average</td>
</tr>
<tr>
<td>3</td>
<td>CP21</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>47.92</td>
<td>27.87</td>
<td>Marginal</td>
</tr>
<tr>
<td>4</td>
<td>CP22</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>50.00</td>
<td>27.87</td>
<td>Marginal</td>
</tr>
<tr>
<td>5</td>
<td>CP23</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>43.75</td>
<td>10.83</td>
<td>Marginal</td>
</tr>
<tr>
<td>6</td>
<td>CP24</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>52.08</td>
<td>16.00</td>
<td>Average</td>
</tr>
<tr>
<td>7</td>
<td>CP25</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>25</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>60.42</td>
<td>36.02</td>
<td>Average</td>
</tr>
<tr>
<td>8</td>
<td>CP26</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>25</td>
<td>50</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>50</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>64.58</td>
<td>37.44</td>
<td>Average</td>
</tr>
<tr>
<td>9</td>
<td>CP31</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>97.92</td>
<td>6.91</td>
<td>Excellent</td>
</tr>
<tr>
<td>10</td>
<td>CP41</td>
<td>50</td>
<td>100</td>
<td>25</td>
<td>75</td>
<td>25</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>50</td>
<td>43.75</td>
<td>25.26</td>
<td>Marginal</td>
</tr>
<tr>
<td>11</td>
<td>PF11</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>41.67</td>
<td>18.63</td>
<td>Marginal</td>
</tr>
<tr>
<td>12</td>
<td>PF12</td>
<td>50</td>
<td>75</td>
<td>75</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>56.25</td>
<td>20.73</td>
<td>Average</td>
</tr>
<tr>
<td>13</td>
<td>PF21</td>
<td>100</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>50</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>50.00</td>
<td>25.00</td>
<td>Marginal</td>
</tr>
<tr>
<td>14</td>
<td>PF22</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>50.00</td>
<td>17.68</td>
<td>Marginal</td>
</tr>
<tr>
<td>16</td>
<td>PF32</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>75</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>37.50</td>
<td>21.65</td>
<td>Poor</td>
</tr>
<tr>
<td>17</td>
<td>BP11</td>
<td>50</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>75</td>
<td>25</td>
<td>0</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>41.67</td>
<td>31.18</td>
<td>Marginal</td>
</tr>
<tr>
<td>18</td>
<td>BP21</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>39.58</td>
<td>25.94</td>
<td>Poor</td>
</tr>
<tr>
<td>19</td>
<td>BP31</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>50</td>
<td>25</td>
<td>50</td>
<td>50</td>
<td>45.83</td>
<td>26.68</td>
<td>Marginal</td>
</tr>
<tr>
<td>20</td>
<td>BP32</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>25</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>70.83</td>
<td>36.56</td>
<td>Good</td>
</tr>
<tr>
<td>21</td>
<td>LG11</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>83.33</td>
<td>21.25</td>
<td>Good</td>
</tr>
<tr>
<td>22</td>
<td>LG12</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>72.92</td>
<td>34.55</td>
<td>Good</td>
</tr>
<tr>
<td>23</td>
<td>LG21</td>
<td>75</td>
<td>25</td>
<td>100</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>72.92</td>
<td>25.94</td>
<td>Good</td>
</tr>
<tr>
<td>24</td>
<td>LG22</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>64.58</td>
<td>31.39</td>
<td>Average</td>
</tr>
<tr>
<td>25</td>
<td>LG31</td>
<td>50</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>62.50</td>
<td>21.65</td>
<td>Average</td>
</tr>
<tr>
<td>26</td>
<td>LG32</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>0</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>18.75</td>
<td>18.04</td>
<td>Poor</td>
</tr>
<tr>
<td>27</td>
<td>LG41</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>75</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>18.75</td>
<td>20.75</td>
<td>Poor</td>
</tr>
<tr>
<td>28</td>
<td>LG42</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100.00</td>
<td>43.75</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

**Mean**: 55.43, **SD**: 23.31, **Average**
## Table VI.
The score of indicators from each surveyed dairy cooperative in West Java Province

<table>
<thead>
<tr>
<th>No.</th>
<th>Indicators</th>
<th>KPSBU</th>
<th>Puspa Mekar Mandiri</th>
<th>Mitra Jaya</th>
<th>Merta</th>
<th>Sarwa Mukti</th>
<th>Tandang Sari Cikajang</th>
<th>Bayongbong Sahyu</th>
<th>Karya Nugraha</th>
<th>Koptan Larasati</th>
<th>Mean SD Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CP11</td>
<td>58</td>
<td>58</td>
<td>33</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>42</td>
<td>25</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>CP12</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>75</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>CP21</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>CP22</td>
<td>75</td>
<td>0</td>
<td>25</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>CP23</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>75</td>
<td>75</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>CP24</td>
<td>25</td>
<td>75</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>75</td>
<td>75</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>CP25</td>
<td>75</td>
<td>75</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>100</td>
<td>25</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>8</td>
<td>CP26</td>
<td>75</td>
<td>75</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>100</td>
<td>25</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>CP31</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>83</td>
<td>50</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>CP41</td>
<td>42</td>
<td>25</td>
<td>17</td>
<td>58</td>
<td>25</td>
<td>42</td>
<td>33</td>
<td>50</td>
<td>75</td>
<td>67</td>
</tr>
<tr>
<td>11</td>
<td>PF11</td>
<td>100</td>
<td>25</td>
<td>25</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>12</td>
<td>PF12</td>
<td>100</td>
<td>25</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>75</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>13</td>
<td>PF21</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>PF22</td>
<td>75</td>
<td>25</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>50</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>PF31</td>
<td>75</td>
<td>75</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>75</td>
</tr>
<tr>
<td>16</td>
<td>PF32</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>17</td>
<td>BP11</td>
<td>83</td>
<td>92</td>
<td>25</td>
<td>92</td>
<td>25</td>
<td>25</td>
<td>58</td>
<td>75</td>
<td>50</td>
<td>83</td>
</tr>
<tr>
<td>18</td>
<td>BP21</td>
<td>75</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>50</td>
</tr>
<tr>
<td>19</td>
<td>BP31</td>
<td>75</td>
<td>83</td>
<td>75</td>
<td>100</td>
<td>75</td>
<td>75</td>
<td>92</td>
<td>100</td>
<td>100</td>
<td>83</td>
</tr>
<tr>
<td>20</td>
<td>BP22</td>
<td>25</td>
<td>75</td>
<td>75</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>75</td>
</tr>
<tr>
<td>21</td>
<td>LG11</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>83</td>
<td>100</td>
</tr>
<tr>
<td>22</td>
<td>LG12</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>23</td>
<td>LG21</td>
<td>92</td>
<td>100</td>
<td>100</td>
<td>83</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>24</td>
<td>LG22</td>
<td>75</td>
<td>75</td>
<td>25</td>
<td>100</td>
<td>25</td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>25</td>
<td>LG31</td>
<td>100</td>
<td>50</td>
<td>25</td>
<td>100</td>
<td>50</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>26</td>
<td>LG32</td>
<td>50</td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>LG41</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>LG42</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Mean 58.75 26.11 Average
2.6802 1.7868
8.9341 35.7364
0.1254 0.5016
0.0003 0.0011
0.2937 0.2937
0.0073 0.0073
0.0000 0.8192
0.0000 0.0331
0.0790 0.0790
8.4695 25.4085
5.9788 1.9929
3.9918 2.6612
0.0000 0.0000
0.0005 0.0014
0.0029 0.0029
0.0010 0.0010
0.0224 0.0000
0.2263 0.2263
0.0000 0.0000
0.0004 0.0004
0.2502 0.3336
0.0003 0.0004
0.0000 0.0000
0.0000 0.0000
7.6007 5.0671
0.0000 0.0000
0.0000 0.0000
0.0004 0.0004

Kota
1.7868
8.9341
0.5016
0.0011
0.1468
0.0036
0.2048
0.0083
0.0790
8.4695
1.9929
1.3306
0.0000
0.0009
0.0000
0.0029
0.0673
0.4527
0.0000
0.0001
0.0834
0.0001
0.0000
0.0000
7.6007
0.0000
0.0000
0.0004

2.6802
26.8023
0.2508
0.0005
0.2937
0.0073
0.6144
0.0249
0.0790
16.9390
3.9859
3.9918
0.0000
0.0009
0.0059
0.0020
0.0673
0.4527
0.0000
0.0004
0.2502
0.0003
0.0000
0.0000
7.6007
0.0000
0.0000
0.0004

2.6802
8.9341
0.3762
0.0008
0.2937
0.0073
0.8192
0.0331
0.0593
16.9390
3.9859
3.9918
0.0000
0.0009
0.0088
0.0020
0.0673
0.4527
0.0000
0.0004
0.3336
0.0003
0.0000
0.0000
7.6007
0.0000
0.0000
0.0004

2.6802
26.8023
0.1254
0.0005
0.2937
0.0073
0.6144
0.0249
0.0790
0.0000
5.9788
3.9918
0.0000
0.0014
0.0059
0.0029
0.0673
0.9054
0.0000
0.0004
0.3336
0.0004
0.0000
0.0000
10.1342
0.0000
0.0000
0.0004

Wahyu
Agung
2.6802
26.8023
0.1254
0.0003
0.2937
0.0073
0.4096
0.0166
0.0790
8.4695
3.9859
3.9918
0.0000
0.0005
0.0059
0.0010
0.0224
0.4527
0.0000
0.0002
0.2502
0.0004
0.0000
0.0000
5.0671
0.0000
0.0000
0.0004

Andini
Luhur
Mean

2.6802 2.3824
26.8023 23.0798
0.1254 0.2403
0.0003 0.0005
0.2937 0.2570
0.0073 0.0076
0.8192 0.4949
0.0331 0.0214
0.0790 0.0774
16.9390 14.8216
3.9859 3.3215
2.6612 2.9939
0.0000 0.0000
0.0009 0.0009
0.0029 0.0039
0.0010 0.0015
0.0673 0.0374
0.4527 0.3584
0.0000 0.0000
0.0002 0.0003
0.3336 0.2780
0.0004 0.0003
0.0000 0.0000
0.0000 0.0000
7.6007 6.3339
0.0000 0.0000
0.0000 0.0000
0.0004 0.0004

Ngudi
Luhur

39.2983 63.1701 64.0506 46.5878 52.0503 52.6623 62.8868 53.97036364
Poor
Average Average Marginal Average Average Average Average

0.8934 3.5736
17.8682 35.7364
0.1254 0.1254
0.0003 0.0003
0.1468 0.1468
0.0073 0.0146
0.4096 0.4096
0.0166 0.0331
0.0790 0.0790
8.4695 16.9390
1.9929 1.9929
1.3306 1.3306
0.0000 0.0000
0.0005 0.0014
0.0000 0.0029
0.0010 0.0000
0.0224 0.0000
0.0000 0.0000
0.0000 0.0000
0.0000 0.0001
0.3336 0.2502
0.0001 0.0000
0.0000 0.0000
0.0000 0.0000
7.6007 2.5336
0.0000 0.0000
0.0000 0.0000
0.0004 0.0004

Performance index of each indicator
Sumber
Banyumanik Subur Mekar
Karya Getasan

81.9269 38.6653 74.9545
31.6678
Good
Poor
Good
Poor

2.6802
35.7364
0.2508
0.0005
0.2937
0.0073
0.0000
0.0000
0.0790
33.8780
1.9929
3.9918
0.0000
0.0009
0.0029
0.0010
0.0000
0.2263
0.0000
0.0004
0.2502
0.0004
0.0000
0.0000
2.5336
0.0000
0.0000
0.0004

Mojosongo Musuk Cepogo

1 CP11
0.03573640060
1.7868
2 CP12
0.35736400600 17.8682
3 CP21
0.00501587722
0.2508
4 CP22
0.00001052400
0.0005
5 CP23
0.00587329515
0.2937
6 CP24
0.00014568638
0.0073
7 CP25
0.00819161620
0.8192
8 CP26
0.00033146386
0.0331
9 CP31
0.00079011918
0.0790
10 CP41
0.33878001465 16.9390
11 PF11
0.07971709706
1.9929
12 PF12
0.05322452827
2.6612
13 PF21
0.00000000094
0.0000
14 PF22
0.00001884419
0.0009
15 PF31
0.00011788870
0.0059
16 PF32
0.00003929623
0.0020
17 BP11
0.00089796735
0.0449
18 BP21
0.00905381412
0.4527
19 BP31
0.00000000506
0.0000
20 BP32
0.00000444648
0.0004
21 LG11
0.00333633102
0.3336
22 LG12
0.00000372486
0.0004
23 LG21
0.00000000097
0.0000
24 LG22
0.00000000087
0.0000
25 LG31
0.10134217904
5.0671
26 LG32
0.00000073110
0.0000
27 LG41
0.00000005474
0.0000
28 LG42
0.00000408577
0.0004
Aggregate value of the
performance of relationship
48.6401
Category
Marginal

No. Indicators

Priority
Weight

Dairy supply
chain in
Indonesia

1451

Table VII.
The performance
index of indicators
and the aggregate
value of the
performance of
relationship in each
surveyed dairy
cooperative in Central
Java Province


Table VIII.
The performance
index of indicators
and the aggregate
value of the
performance of
relationship in each
surveyed dairy
cooperative in West
Java Province

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

CP11
0.00424202811240
CP12
0.00000000052714
CP21
0.00008222012712
CP22
0.00000549322664
CP23
0.00000000072848
CP24
0.00000000003928
CP25
0.00000000000008
CP26
0.00000000000001
CP31
0.00115521724745
CP41
0.00907150252679
PF11
0.00000000258298
PF12
0.00000000005271
PF21
0.00000000000019
PF22
0.00000001201888
PF31
0.00000000034158
PF32
0.00000148012816
BP11
0.00000002510941
BP21
0.00000000028318
BP31
0.00002373827518
BP32
0.00000000000815
LG11
0.00750674256926
LG12
0.00000000002255
LG21
0.62617000185320
LG22
0.00000005109384
LG31
0.17328713847026
LG32
0.00000001413978
LG41
0.00031229066579
LG42
0.17814203984952
Aggregate value of the
performance of relationship
Category

Priority Weight
0.141387
0.000000
0.002056
0.000137
0.000000
0.000000
0.000000
0.000000
0.086641
0.151222
0.000000
0.000000
0.000000
0.000001
0.000000
0.000000
0.000001
0.000000
0.001780
0.000000
0.750674
0.000000
62.617000
0.000001
4.332178
0.000000
0.000000
17.814204

89.874961 85.897284 88.928954
Excellent Good
Good

0.247437 0.247437
0.000000 0.000000
0.002056 0.000000
0.000412 0.000000
0.000000 0.000000
0.000000 0.000000
0.000000 0.000000
0.000000 0.000000
0.115522 0.115522
0.378010 0.226788
0.000000 0.000000
0.000000 0.000000
0.000000 0.000000
0.000001 0.000000
0.000000 0.000000
0.000148 0.000000
0.000002 0.000002
0.000000 0.000000
0.001780 0.001978
0.000000 0.000000
0.750674 0.187669
0.000000 0.000000
57.401004 62.617000
0.000004 0.000004
17.328714 8.664357
0.000001 0.000000
0.007807 0.000000
17.814204 17.814204

KPSBU

Mitra Jaya
Mandiri
0.106051
0.000000
0.000000
0.000000
0.000000
0.000000
0.000000
0.000000
0.115522
0.378010
0.000000
0.000000
0.000000
0.000001
0.000000
0.000000
0.000001
0.000000
0.001780
0.000000
0.750674
0.000000
62.617000
0.000001
12.996535
0.000000
0.007807
17.814204

0.176765
0.000000
0.002056
0.000137
0.000000
0.000000
0.000000
0.000000
0.115522
0.302353
0.000000
0.000000
0.000000
0.000001
0.000000
0.000000
0.000001
0.000000
0.002176
0.000000
0.750674
0.000000
62.617000
0.000005
17.328714
0.000001
0.007807
17.814204

89.874961 94.787588 99.117417 99.195931
Good
Average
Excellent Excellent

0.212101 0.212101
0.000000 0.000000
0.008222 0.004111
0.000549 0.000275
0.000000 0.000000
0.000000 0.000000
0.000000 0.000000
0.000000 0.000000
0.096264 0.057761
0.529141 0.226788
0.000000 0.000000
0.000000 0.000000
0.000000 0.000000
0.000001 0.000001
0.000000 0.000000
0.000148 0.000000
0.000002 0.000000
0.000000 0.000000
0.002374 0.001780
0.000000 0.000000
0.750674 0.500475
0.000000 0.000000
52.178746 41.746754
0.000005 0.000001
17.328714 8.664357
0.000001 0.000000
0.007807 0.000000
17.814204 0.000000

KPBS

Score of each indicator
Sarwa
Sinar Jaya
Mukti
Tandangsari
0.176765
0.000000
0.004111
0.000275
0.000000
0.000000
0.000000
0.000000
0.115522
0.604797
0.000000
0.000000
0.000000
0.000000
0.000000
0.000000
0.000002
0.000000
0.002374
0.000000
0.625537
0.000000
62.617000
0.000005
17.328714
0.000000
0.000000
17.814204

Bayongbong

94.787588 99.2089306
Excellent Excellent

0.106051
0.000000
0.000000
0.000000
0.000000
0.000000
0.000000
0.000000
0.115522
0.453575
0.000000
0.000000
0.000000
0.000001
0.000000
0.000000
0.000002
0.000000
0.002374
0.000000
0.750674
0.000000
62.617000
0.000005
17.328714
0.000001
0.007807
17.814204

Cikajang
0.141387
0.000000
0.002056
0.000137
0.000000
0.000000
0.000000
0.000000
0.115522
0.529141
0.000000
0.000000
0.000000
0.000001
0.000000
0.000074
0.000001
0.000000
0.001780
0.000000
0.750674
0.000000
15.654250
0.000005
17.328714
0.000000
0.007807
17.814204

Karya
Nugraha

98.968319 52.345754
Excellent Average

0.176765
0.000000
0.004111
0.000275
0.000000
0.000000
0.000000
0.000000
0.115522
0.151222
0.000000
0.000000
0.000000
0.000001
0.000000
0.000037
0.000002
0.000000
0.001978
0.000000
0.750674
0.000000
62.617000
0.000005
17.328714
0.000001
0.007807
17.814204

Saluyu

1452

No. Indicators

Puspa
Mekar

90.333913
Excellent

0.141387
0.000000
0.002056
0.000137
0.000000
0.000000
0.000000
0.000000
0.115522
0.226788
0.000000
0.000000
0.000000
0.000001
0.000000
0.000000
0.000002
0.000000
0.001780
0.000000
0.750674
0.000000
62.617000
0.000005
8.664357
0.000000
0.000000
17.814204

Koptan
Larasati

87.016801
Good

0.173803
0.000000
0.002569
0.000195
0.000000
0.000000
0.000000
0.000000
0.106697
0.346486
0.000000
0.000000
0.000000
0.000001
0.000000
0.000034
0.000002
0.000000
0.001995
0.000000
0.672479
0.000000
55.659730
0.000004
13.718565
0.000001
0.004554
16.329687

Mean

IJPPM
67,9


This preference is recommended by Lynch et al. (1996) and also Martilla and James (1977). In detail, the condition of the four quadrants in IPA can be explained as follows. “Concentrate here” is a quadrant that has high importance but low performance, so we should pay more attention to the indicators belonging to this quadrant because they indicate the major weakness of the enterprise. “Keep up the good work” is a quadrant of high importance and high performance. We may give some attention to maintain the indicators in
this quadrant because those indicators have opportunities to improve the enterprise's competitive advantage. The indicators in this quadrant indicate the major strength of the enterprise. “Low priority” is a quadrant for indicators with low importance and low performance. The indicators in this quadrant indicate the minor weakness of the enterprise and the enterprise does not need additional effort to improve those indicators. “Possible overkill” is a quadrant that has low importance but high performance. This quadrant indicates that business resources committed to these indicators would be overkill and should be deployed elsewhere (Martilla and James, 1977).

For Central Java (Figure 1), we could not see the indicators included in the quadrant of “concentrate here” because the position of the indicators is very close to each other. However, by comparing the score of the indicator with its median value (52.08) and the priority weight of the indicator with its median value (0.000331464), it can be concluded that there are six indicators included in quadrant “concentrate here”, namely, the level of conformity of TPC contained in the milk delivered by the farmers with SNI (CP21), the level of conformity of fat contained in the milk delivered by the farmers with SNI (CP23), percentage of profit sharing received by the farmers from selling their milk to cooperatives (PF11), time span between cooling process in the dairy cooperative and milking process (BP11), the number of types of quality checking of milk conducted by the cooperative before they sent it to the industrial milk processor (BP21) and the level of sophistication of equipment used by the farmers for milking process (LG41). These indicators are the major weakness of the relationships between dairy farmers, cooperatives and industrial milk processor in Central Java Province. On the other hand, the major strength consists of eight indicators belonging to quadrant “keep up the good work,” namely, the level of satisfaction of dairy farmers with price offered by the cooperative (CP11), level of satisfaction of cooperatives with the commitment of dairy farmers to produce milk with specific quantity (CP12), the level of conformity of Solid Nonfat (SNF) contained in the milk delivered by the farmers with SNI (CP25), level of conformity of Solid Nonfat (SNF) contained in the milk delivered by the farmers with standard of industrial milk processor (CP26), easiness of the farmers to get in touch with the cooperatives (CP31), percentage of profit sharing received by the cooperatives from selling their milk to industrial milk processors (PF12), the frequency of information sharing between the farmers and cooperatives (LG11) and the frequency of training for capacity building from cooperatives to farmers in a year (LG31).

As for the West Java case (Figure 2), we also could not see the indicators included in the quadrant of “concentrate here” because the position of the indicators is very close to each other. However, using the same method with that on Central Java above, it can be recognized that there are six indicators in the quadrant of “concentrate here”, i.e., level of satisfaction of dairy farmers with price offered by the cooperative (CP11), level of conformity of TPC contained in the milk delivered by the farmers with SNI (CP21), level of conformity of TPC contained in the milk delivered by the cooperatives with the standard of industrial milk processor (CP22), duration of farmers to become a member of cooperatives (CP41), prices per liter that cooperatives get out of industrial milk processor according to the quality of their milk (PF32) and level of sophistication of equipment used by the farmers for milking process (LG41). These indicators are the major weakness of the relationships between dairy farmers, cooperatives, and industrial milk processor in West Java Province. Then, the major strength consists of eight indicators in the quadrant of “keep up the good work”, namely, easiness of the farmers to get in touch with the cooperatives (CP31), time span between cooling process in the dairy cooperative and milking process (BP11), level of implementation of HACCP and other quality control measures by the farmers (BP31), frequency of information sharing between the farmers and cooperatives (LG11), level of collaboration in problem-solving between the farmers and cooperatives (LG11), level of collaboration in problem-solving between the cooperatives and industrial processing milk (LG22), frequency of training for capacity building from cooperatives
Comparing the results of IPA analysis in two different regions, it is found out that two indicators, i.e., the level of conformity of TPC content with SNI in milk delivered by the farmers and the level of sophistication of equipment used by the farmers for milking process, are the major weakness in both provinces. The level of TPC is one of the main problems of milk quality received at the dairy cooperative and industrial milk processor. High TPC levels are a concern for industrial milk processors as it impacts the usage of the milk. To address the milk quality issue, the industrial milk processor has a campaign by its farm advisors with training on hygiene practices; replacing plastic buckets with metal ones; and installing some milk cooling units (Morey, 2011). Moreover, the major weakness in Central Java Province lies in the different perspectives (no perspective is dominant), whereas that in West Java Province is dominated by the perspective of the customer.

The indicator of easiness of the farmers to get in touch with the cooperatives (CP31) and frequency of training for capacity building from dairy cooperatives to farmers become the main strength in both provinces. It is because the cooperative has the important role for the dairy farmers. Dairy cooperatives were introduced by the government to link milk producers with milk processors and to provide farmers with services and inputs. The cooperative’s role is like a buffer between the dairy farmers and the dairy processing industry. The cooperative is an organization with the main purpose to improve the farmers’ welfare (Sebayang, 2013; Susanty et al., 2017). The major strength of the relationships between dairy farmers, cooperatives and industrial milk processors in Central Java is dominated by the perspective of the customer, while that in West Java is dominated by the perspective of learning and growth.

Strategic planning with SWOT analysis
SWOT analysis is one of the most popular tools for strategic planning (Lu, 2010). In this research, SWOT analysis is used to formulate strategic planning in dairy milk supply chain, specifically, to enhance the performance of the relationship between dairy farmers, cooperatives and industrial milk processors. SWOT is an acronym for strengths, weaknesses, opportunities and threats. It has its origins in the 1960s (Learned et al., 1965), and was popularized by Wehrich’s (1982) work. The outcome of the analysis is in terms of suggestions insights regarding the trajectory of the organization categorized in “strengths” that should be sustained (i.e. inner potential), “weaknesses” that must be overcome (i.e. inner barriers), “opportunities” that have to be sought (i.e. environmental prospects), and “threats” that ought to be alleviated (i.e. environmental hindrances) (Hovardas, 2015).

The data for the analysis came from two different sources. The first is with regards to strengths and weaknesses, i.e., the results of assessing the condition of the internal relationship between dairy farmers, cooperatives and milk processors through the IPA. As explained in the previous section, in Central Java Province, there were eight indicators belonging to the major strength and six indicators belonging to the major weakness, while in West Java Province, there were eight and six indicators belong to mayor weakness. Then, the source of data for the second factor of SWOT (opportunities and threats) is the result of the interview with the representatives of management of dairy cooperatives being sampled in this research and also the policy and regulation from the government which is related with the dairy industry. There are three opportunities. First, the growth in demand in dairy products as an impact of “Fresh Milk Campaign.” In this campaign, the government gives subsidies and free milk to primary students (Ditjennak, 2011). The government also cooperates with dairy firms in promoting the health benefits of fresh milk. The government aims to double the consumption of milk by 2024. June 1 was declared as the National Milk Day by the Ministry of Agriculture in 2009 (Ministry of Agriculture Decree No. 2182/KPTS/
PD.420/5/2009) (Vanzetti et al., 2016). Demand in fresh milk also increases in order to fulfill the raw material needed by industrial milk processors. Currently, more than 80 percent of raw material needed by industrial milk processors should be imported from the other countries. Second, credit facility for funding is being endorsed by the Government. By Presidential Instruction Number 6/2007, the government introduced a credit scheme for micro and small enterprises, known as microcredit loans (KUR-Kredit Usaha Rakyat). These are government-guaranteed loans directed to micro, small and medium enterprises as well as cooperatives, which are productive and feasible businesses, but still un-bankable. The KUR scheme initially requires a project or business activity as the principal collateral for the loan. However, since this collateral does not meet with the banks' own requirements, the government initiated a guarantee program for micro, SMEs and cooperatives so they can access loans from banks. KUR is intended to provide working capital and investment credit of up to Rp500m. The credit providers are commercial banks assigned by the government (Machmud and Huda, 2011). Third, Government Regulation No. 6/2013 concerning the empowerment of the farmers. Empowerment of farmers is all efforts made by the government, provincial government, district/city government and stakeholders in the field of animal husbandry and health to enhance independence, facilitate and improve the business, competitiveness and welfare of farmers. Then, the main threat faced by dairy industry is related to the free market of dairy commodity and import of raw milk. One aspect of pillar number one in the ASEAN economic community (AEC) which was effectively implemented by the end of 2015 is free flow of goods. Within this pillar, the tariff will be reduced or eliminated to increase the value of inter and extra-ASEAN trade in the agricultural, including dairy sector (Priyanti and Soedjana, 2015). Then, the regulation of the Minister of Finance No. 145/PMK.011/2008 concerning fiscal incentives in the form of government-borne duties on the import of goods and materials used in the dairy processing industry has caused industrial milk processor free to import milk. The regulation of the Minister of Finance No. 19/PMK.011/2009 regarding the determination of import duty rated on specific milk products from five percent to zero percent has made the industrial milk processors more powerful to determine the price of milk.

So, based on the major strength and weakness in the current relationships between dairy farmers, cooperatives, and industrial milk processors, and also the opportunity and threat faced by the dairy milk industry, the strategic planning for each province can be seen in Tables IX and X.

5. Discussion
Using the balanced supply chain management scorecard, IPA analysis and SWOT analysis, this research has three purposes. First, this study aims to measure the current performance of the relationships between farmers, dairy cooperatives and industrial milk processors. The second one is to identify the indicators that are most in need of improvement, and the third is to formulate some strategic plans. In the case of the two provinces, the results of the balanced supply chain management scorecard combined with the IPA analysis show that the performance of the relationships between farmers, dairy cooperatives and industrial milk processors in West Java Province is slightly better than that in Central Java. It can be seen from the average value of the score of indicator, the category of each indicator and the category of the performance index of each relationship. The average value of the score of the indicator in West Java Province is slightly larger than that in Central Java. Also, the percentage of indicators belonging to the category of average, good and excellent in West Java is slightly larger than that in Central Java. Almost all of the performance indices of the relationships between dairy farmers, cooperatives and industrial milk processors in West Java are included in the category of average and good, whereas, in Central Java, only half of the performance indices are in the category of average and good.
Strength | Weakness
---|---
Most of dairy farmers have been satisfied with price offered by the cooperative | Low level of conformity of TPC content with SNI in milk delivered by the farmers
Most of dairy cooperatives have been satisfied with the commitment of dairy farmers to produce milk with specific quantity | Low level of conformity of fat content with SNI in milk delivered by the farmers
High level of conformity of SNF content with SNI in milk delivered by the farmers | Low percentage of profit sharing received by the farmers from selling their milk to cooperatives
Most of the farmer feel easy to get in touch with the cooperatives | Time span between cooling process in the dairy cooperative and milking process is too long
High percentage of profit sharing received by the cooperatives from selling their milk to industrial milk processors | Types of quality checking of milk conducted by the cooperative is very limited
High frequency of information sharing between the farmers and cooperatives | Most of dairy farmer still use their hand for milking process
Training is often carried out by dairy cooperative to increase the capacity of the farmers | Encourage the cooperatives to conduct comprehensive analysis in determining the most optimal route for milk collection from the farmers, so time span between milking process and storage in cooling units in the cooperatives can be shortened
Opportunity:
Growth in demand in dairy product, including fresh milk as the impact of “Fresh Milk Campaign” | Encourage the cooperative to utilize the credit facility from government to purchase more vehicles for milk collection and also equipment needed for milk quality testing
The availability of credit from government for micro, SMEs, and cooperatives, so they can access loans from banks to provide working capital and investment credit | Encourage the farmer to utilize the credit facility from government to purchase high quality of fodder and milking machine. Using milking machine will reduce the contact between milk with the farmer, resulting in more cleaner and hygienist milk
Government Regulation No. 652013 concerning empowerment of farmer | Encourage the dairy farmers and cooperative to utilize the aid from government to enhance their capacity in produce high-quality milk
Threat:
Free flow of dairy product from ASEAN countries as impact of AEC | Utilizing training as a means for educating the farmer in producing high quantity and quality of milk so the milk produced by the dairy farmer not only can meet the demands of industrial milk processor but also can be exported to the ASEAN countries
The regulation of the Minister of Finance No. 145/PMK.011/2008 has caused industrial milk processor free to import milk | Regulation of the Minister of Finance No. 19/PMK.011/2009 has caused the industrial milk processors power to determine the price of milk
Regulation of the Minister of Finance No. 19/PMK.011/2009 has caused the industrial milk processors power to determine the price of milk |
<table>
<thead>
<tr>
<th>Strength</th>
<th>Weakness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most of the farmers feel easy to get in touch with the cooperatives</td>
<td>Low level of satisfaction of dairy farmers with price offered by the cooperative</td>
</tr>
<tr>
<td>Short duration time between milking process and stored in cooling unit in cooperative</td>
<td>Low level of conformity of TPC content with SNI in milk delivered by the farmers</td>
</tr>
<tr>
<td>High level of implementation of HACCP and other quality control measures by the farmers</td>
<td>Low level of conformity of TPC content with SNI in milk delivered by the cooperatives to industrial milk processor</td>
</tr>
<tr>
<td>High frequency of information sharing between the farmers and cooperatives</td>
<td>Loyalty of the farmers to dairy cooperatives still low; it can be seen from the duration of farmers become a member of cooperatives</td>
</tr>
<tr>
<td>High level of collaboration in problem-solving between the farmers and cooperatives</td>
<td>Low prices per liter that cooperatives get from of industrial milk processor</td>
</tr>
<tr>
<td>High level of collaboration in problem-solving between the cooperatives and industrial processing milk</td>
<td>According to the quality of their milk</td>
</tr>
<tr>
<td>Training is often carried out by dairy cooperative to increase the capacity of the farmers</td>
<td>Most of dairy farmer still use their hand for milking process</td>
</tr>
<tr>
<td>High level of sophistication of equipment used by cooperative for cooling the milk (LG42)</td>
<td></td>
</tr>
</tbody>
</table>

Opportunity:  
- Growth in demand in dairy product, including fresh milk as the impact of “Fresh Milk Campaign”  
- The availability credit for micro, SMEs and cooperatives, so they can access loans from banks to provide working capital and investment credit  
- Government Regulation No. 6/2013 concerning empowerment of farmers  

Utilize information sharing and the collaboration between dairy farmers, cooperative and industrial processing to discuss and solve the problems of production techniques faced by farmers, so they can produce more quality milk  
Improve the implementation of HACCP and other quality control by utilizing the credit facility and government aid  
Improve the level of satisfaction of equipment used by cooperative for cooling the milk by utilizing credit facility from government  
Utilizing training as a means for educating the farmers in producing high quantity and quality of milk so the milk produced by the dairy farmer not only can meet the demands of industrial milk processor but also can be exported to the ASEAN countries  
Improve the implementation of HACCP and other quality control measures by the farmer and also the level of satisfaction of equipment used by cooperative for cooling the milk to produce the quality milk that meet the standard of industrial processing milk and also standard of ASEAN countries  
Encourage the cooperatives to conduct comprehensive analysis in determining the relationship between price and the quality of milk offered by the farmers  
Utilize the promise from government to empower the dairy farmers as a tool to make an agreement with the industrial processing related with milk price  
Encourage the farmer to utilize the credit facility from government to purchase milking machine so they can produce more hygienic milk  

Threat:  
- Free flow of dairy product from ASEAN countries as impact of AEC  
- The regulation of the Minister of Finance No. 145/PMK.011/2008 has caused industrial milk processor free to import milk  
- The regulation of the Minister of Finance No. 19/PMK.011/2009 has caused the industrial milk processors power to determine the price of milk  

Utilizing training as a means for educating the farmer in producing high quantity and quality of milk so the milk produced by the dairy farmer not only can meet the demands of industrial milk processor but also can be exported to the ASEAN countries  
Improve the implementation of HACCP and other quality control measures by the farmer and also the level of satisfaction of equipment used by cooperative for cooling the milk to produce the quality milk that meet the standard of industrial processing milk and also standard of ASEAN countries  
Encourage the cooperatives to conduct comprehensive analysis in determining the relationship between price and the quality of milk offered by the farmers  
Utilize the promise from government to empower the dairy farmers as a tool to make an agreement with the industrial processing related with milk price  
Encourage the farmer to utilize the credit facility from government to purchase milking machine so they can produce more hygienic milk
There have been several conditions that make the performance in West Java better than that in Central Java. A major milk processor in West Java has worked with its dairy cooperative to install milk cooling units to improve the quality of milk received. This good action attracted all the dairy cooperatives in West Java to have milk cooling units. There is penalty received by the dairy cooperative if milk received by the industrial processor has a temperature of more than eight Celsius degrees. Besides, The Provincial Government of West Java has received a grant from the Central Government to install milk cooling units at the cooperatives. They installed three units in 2010 and four units in 2011. This measure assisted to reduce the TPC by cooling milk more quickly when received from the farmers.

Then, related to improving the capacity of cooperatives and dairy farmers, West Java Province has Cikole Dairy Training Center. This center is funded by the Japan International Cooperation Agency and aimed at providing technology transfer to improve dairy farming and milk production (Morey, 2011). To get long-term profitability, some cooperatives in West Java (such as KPBSU Lembang) have engaged in dairy processing activities as they are aware that much of the value added is created in the chain function (Susanty et al., 2017).

This research has some implications for the dairy cooperative. It provides some insights for the management of the dairy cooperative in the context of strategic planning. The research reveals that there is a difference in strategic planning for improving the performance of the relationships between dairy farmers, cooperatives, and industrial milk processors between the two provinces. The management in Central Java should pay more attention to the transport of milk from dairy farmers to cooperatives through comprehensive analysis in determining the most optimal route for milk collection and also add more vehicles for the purpose and encourage dairy farmers to utilize the credit facility from the government to purchase high-quality fodder. Besides, it is important for the management of dairy cooperative to maintain the satisfaction of dairy farmers to the cooperative and the ease of contact already established as a means to discuss and make a proposal for getting the credit or assistance from the government. So, the assistance will meet the needs of the farmers.

Meanwhile, the management of the dairy cooperative in West Java should pay attention to the relationship between price and the quality of milk offered by the farmers and make a fair agreement with the industrial milk processor in determining the price. It is also important for management to give real support to improve the implementation of HACCP and other quality control conducted by the farmers and the level of satisfaction of equipment for cooling the milk through utilizing the credit facility from the government. Both of the management in Central Java and West Java Provinces should pay more attention to encourage the farmers to use milking machine in the process so they can deliver high-quality milk, which, in turn, can make the cooperative deliver the high-quality milk to the industrial milk processor. To win the free flow of dairy product in ASEAN countries, it is important for the dairy cooperative in both provinces to utilize training as a means for educating the farmers to produce high-quality milk. This way, not only can the milk meet the SNI but can be exported to other ASEAN countries as well because the milk meets the ASEAN standard.

6. Limitations and future research directions
This study has several limitations. First, not all the dairy cooperatives in Central Java and West Java have become the sample; instead, it was only 24 of them. Although the sample represents 80 percent of the milk produced from Central Java and West Java, the limited sample of this research can make the results still bias due to the condition of the surveyed dairy cooperatives. Second, this study used the Likert scale as an approach for measuring the performance of each indicator in the context of the focus of study, which can be the source of bias as well, especially in expressing the level of current performance of the relationships.
In response to this limitation, suggested future research may lie in trying to add the sample size with the remaining dairy cooperatives in both provinces, and also to replicate this study by surveying the dairy cooperatives in other provinces such as East Java (the other province that produces the largest milk in Indonesia). Future research may also enhance the measurement of the performance of the relationships between farmers, dairy cooperatives and industrial milk processors by using a direct measurement of each indicator in each perspective, rather than relying on the cooperative management’s perceptions.

References


Further reading


Corresponding author

Aries Susanty can be contacted at: ariessusanty@gmail.com

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Measurement invariance of employee engagement across race groups

Wesley Herscelle Gallant and Nico Martins
Department of Industrial and Organisational Psychology, University of South Africa, Pretoria, South Africa

Abstract

Purpose – The testing of measurement invariance is important in cross-cultural research to establish whether the psychometric properties of an instrument remain valid and reliable across different sample groups as these assumptions are rarely tested statistically. The purpose of this paper is to determine the factorial invariance of the employee engagement questionnaire across the various race groups by means of structural equation modelling.

Design/methodology/approach – Cross-sectional and descriptive research designs were followed in this study in the form of non-probability, convenience sampling to attract a sample of 1,175 employees in financial institutions. The employee engagement instrument (EEI) was electronically administered to 285,000 people who form part of a research database.

Findings – The results confirmed the reliability and validity of the instrument as determined by the exploratory and confirmatory factor analysis. Lastly, the results indicated that invariance can be assumed across race groups for financial institutions.

Practical implications – It is important for organisations to take cognisance of how specific socio-demographic variables influence the measurement of employee engagement, in this case race. The conclusion reached was that the EEI can be used with confidence in the financial sector for future employee engagement assessments.

Originality/value – These findings add to the current body of literature that exists on employee engagement and race in the South African work context and addresses one of the complexities assessment practitioners might have to comply with regarding questionnaire validity across race groups.

Keywords Employee engagement, Confirmatory factor analysis, Exploratory factor analysis, Multi-group invariance, Race groups

1. Introduction

The world of work has changed dramatically over the last 20 years due to globalisation, technological advances, changes in the social, cultural and economic environments and new ways of doing business that require businesses to keep ahead of their competitors by redefining and re-aligning their business strategies to overcome the challenges of the future. Masibigiri and Nienaber (2011) posit that these constant changes and turbulences have a direct impact on the organisational culture, retention, motivation, satisfaction and engagement of employees, which, in turn, influences organisational effectiveness and performance.

This has also brought about the “war for talent”, which has placed the attraction and retention of talent at the forefront of business strategies to ensure the effective and positive functioning of organisations (Ulbrich, 2015). Organisations are, therefore, constantly looking for new and innovative ways to manage their current workforce by retaining talented and skilled employees to provide them with a competitive advantage and allow them to respond to and overcome the changing market needs successfully. However, retaining the right talent in South Africa, as well as internationally, has proven to be a challenging task due to skills shortages, increased national and international mobility and the retirement of the baby boomers. Furthermore, employee turnover and intention to quit also places tremendous strain on organisations (Brundage and Koziel, 2010; Martins and Martins, 2013).
Employee engagement has been identified as one of the most effective tools organisations can utilise in order to reduce turnover intentions, improve productivity, increase profitability and, most importantly, achieve a competitive advantage (Sardar et al., 2011). Employee engagement has also been found to be positively related to employee satisfaction, individual performance, organisational effectiveness, organisational commitment, employee motivation and productivity, career adaptability, employee retention, intention to quit, customer service, customer loyalty, psychological capital and organisational citizenship behaviour (Coetzee et al., 2014; Hayday et al., 2004).

Against this backdrop, it is important for organisations to attract, retain, motivate and engage the right employees, in other words, employees who demonstrate the right behaviours, especially in the current economic climate. A vast body of literature evidence has been published in support of the notion that employee engagement is a critical ingredient in the short-term survival and long-term business performance and productivity of organisations (Anitha, 2014; Barnes and Collier, 2013; Fearon et al., 2013; Mehrzi and Singh, 2016; Werner, 2014).

In the next section, the research purpose and objectives are discussed and a hypothesis is proposed followed by a literature review. This is followed by the research design discussing the research approach, the method, the research participants and the instrument used in the study. Towards the end the focus is on the results of the study and the decisions regarding the set research hypothesis. Lastly the authors present a brief conclusion and the limitations of the study.

2. Research purpose and objectives
According to Imandin et al. (2014), research on employee engagement has primarily focused on how organisations can enhance employee engagement, what they should do to drive and facilitate employee engagement, the antecedents of employee engagement and the measurement of employee engagement in general. Very few studies have been conducted on the accurate measurement of employee engagement that organisations can use to measure the levels of employee engagement within their organisations. Imandin et al. (2014) further state that the measurement of employee engagement within organisations and for research purposes requires a newly developed, validated and reliable measure that is based on literature from reputable sources. Additionally, there appears to be a paucity of research on testing for measurement invariance of psychometric tools for different race groups. Byrne and Stewart (2006) stated that measurement invariance is of critical importance when conducting multi-group comparisons. Measurement invariance is, therefore, of critical importance, especially with regard to cross-cultural research, in order to establish whether an instrument performs exactly the same way across different sample groups as it has been found that these assumptions are rarely tested statistically.

There is increasing pressure on South African organisations to improve their financial performance and sustain their competitiveness requires engaged employees, especially in the financial industry (João and Coetzee, 2011). Research by E-Trinity (2014) and Deloitte (2014) suggests that employee engagement is on the decline in financial institutions. The financial industry is known for its quest to attract, retain, motivate and develop talent from diverse groups of people due to the national skills shortage in this industry (João and Coetzee, 2011), thus emphasising the importance of employee engagement in this industry.

Research objective of this study
Based on the analysis of existing literature, there seems to be a scarcity of research, both locally and internationally, that focuses on establishing multi-group invariance of psychometric tools, especially with regard to race. Furthermore, there is currently also a lack of research that addresses the impact that race has on employee engagement.
In light of the above, the objectives of this research study were to determine the factorial invariance of employee engagement across the various race groups by means of structural equation modelling (SEM) in financial institutions.

**Hypothesis**

Based on the literature, purpose and objectives of the research, the following hypothesis was formulated:

\( H1. \) Race groups in financial institutions display invariance with regard to the EEI.

### 3. Literature review

#### Employee engagement defined

Employee engagement forms part of the positive psychology movement, which focuses on enhancing the optimum functioning, well-being, passion and health of employees in the business environment (Schaufeli and Bakker, 2004).

Various definitions of employee engagement exist in literature. Engagement was first conceptualised by Kahn (1990), who described employee engagement as the harnessing of individual employees’ selves to their work roles, whereby engaged employees express themselves physically, emotionally and cognitively during the execution of their work. Richman (2006) refers to employee engagement as the emotional and intellectual commitment that employees display towards their work and the organisation.

Macey and Schneider (2008) proposed a theoretical taxonomy based on Kahn’s conceptual framework and states that employee engagement should be viewed as a multidimensional construct which comprises three different forms of engagement, namely, trait, state and behavioural engagement.

Trait engagement is defined as the inclination or orientation to experience the world from a particular vantage point. Psychological state engagement is defined as an antecedent to behavioural engagement (encompassing the constructs of satisfaction, involvement, commitment and empowerment) and behavioural engagement was defined in terms of discretionary effort (Macey and Schneider, 2008, p. 6).

Shuck and Wollard (2010, p. 103) describe employee engagement as “an individual employee’s cognitive, emotional and behavioural state directed toward desired organizational outcomes”. Shuck and Wollard’s definition is partially derived from definitions of employee engagement as conceptualised by Kahn (1990) and Macey and Schneider (2008). Anitha (2014, p. 310) summarised her discussion of significant definitions of employee engagement as follows: “The [above] determinants of employee engagement descriptions of the concept depict that an engaged employee is intellectually and emotionally bound with the organisation, feels passionately about its goals and is committed to live by its values”.

#### Employee engagement some empirical findings

Endres and Manchano-Smoak (2008) and Gruman and Saks (2011) posit that over the last two decades, research on employee engagement has increased exponentially and has even been branded as “a human resource craze”. Rothmann and Rothmann (2010) posit that the need for businesses to minimise their inputs and maximise their outputs has contributed significantly to the rise in employee engagement research and practices. Whilst research studies conducted by other academics and researchers have reported varying and contradicting results, what most of these studies have had in common is that they have found that employee engagement is a critical element to sustaining a competitive advantage.

Despite the increase in research with regard to the importance of employee engagement in an organisational context, various research studies have shown that the levels of
employee engagement have been on a decrease over the last decade (Gallup, 2013; Saks and Gruman, 2014). A recent survey conducted by Gallup (2013), in 142 countries, found that only 13 per cent of the sample was engaged by their organisations and work roles. The results also showed that actively disengaged workers outnumbered the number of engaged employees at a ratio of 2 to 1. Specifically, results obtained from Australia and New Zealand found that 60 per cent of employees are not engaged in their work and that 16 per cent of employees are actively disengaged from their work roles. In sub-Saharan Africa, results indicate that during the course of 2011 and 2012, only 9 per cent of South African employees were engaged and 91 per cent of employees were not engaged. South Africa also reported one of the highest percentages of disengaged employees in the world (Gallup, 2013).

The Gallup Group found that employee engagement is statistically significantly related to positive organisational behaviours and outcomes, such as productivity, profitability, employee retention, employee motivation, job satisfaction, customer satisfaction and psychological capital (Endres and Manchano-Smoak, 2008; Xanthopoulou et al., 2009).

Rothmann and Rothmann (2010) assert that it is important to understand the antecedents (e.g. work-role fit, job enrichment, the availability of personal resources and co-worker support) of employee engagement. Research conducted by these authors found that the antecedents of employee engagement explained only 20 per cent of the variance in employee engagement, compared to 36 per cent of the variance explained by the work activities model (Bakker et al., 2008).

Gruman and Saks (2011) found that work relationships with immediate managers/supervisors that are supportive, trusting and non-autocratic promote employee engagement. Research works by Bakker and Demerouti (2008) and Coetzer and Rothmann (2007) have consistently found that factors such as social support from peers, support from immediate management, open communication channels, performance feedback, autonomy, job enrichment and training and development opportunities positively correlate to employee engagement, whereas job demands displayed a negative relationship with employee engagement. Research conducted by Hassan and Ahmed (2011) on authentic leadership, trust and employee engagement found that authentic leadership enhanced the level of trust employees have in their leadership and that interpersonal trust is a strong predictor of employee engagement.

Xanthopoulou et al. (2007) found a positive relationship between employee engagement and self-efficacy, optimism, hope, resilience and organisational self-esteem. Research works by Patrick and Bhat (2014), on the relationship between employee engagement, critical psychological states and personal resources, found that employee engagement positively correlates with personal resources (optimism, hope and efficacy), indicating that work engagement enhances personal resources and that this can be attributed to the type of work employees do.

A meta-analysis conducted by Harter et al. (2009) reported that engagement was related to nine performance measures used in organisations and that organisations whose employees display high levels of engagement have a significantly better chance at achieving productivity and success as compared to organisations whose employees display low levels of employee engagement or who are actively disengaged. This was also supported by Anitha (2014) who reported that it also shows that there is a strong significant relationship between employee engagement and employee performance. A meta-analysis conducted by Halbesleben (2010) indicates that engaged employees displayed higher commitment to their jobs as well as the organisations, improved health, higher levels of job performance and lower turnover intentions. Furthermore, a meta-analysis by Christian et al. (2011) showed that employee engagement had a positive relationship with task performance and contextual performance. A recent study by Mehrzi and Singh (2016) presents a conceptual framework of employee engagement, demonstrating the relationships between organisational culture, leadership, teamwork and perceived organisational support,
as mediated by the employee motivation and concluded that employee engagement is vital for organisations’ success and sustainability.

From a race or multi-cultural perspective, research by Igbaria (1992) found that black employees showed lower levels of career satisfaction than white employees. Lower career satisfaction was attributed to the perception that black employees received more career development support from their supervisors and managers than the white employees. A study conducted by Somers (2001) found that black employees were more involved and committed to their work than white employees, which means that black employees showed higher levels of job involvement. Research by Jones and Harter (2005) indicated that employees from different racial groups, who displayed low levels of employee engagement, often reported increased intentions to quit than employees from the same racial group. Findings by Bell and Barkhuizen (2011) indicated that white employees displayed higher levels of employee engagement than their coloured and black counterparts. Lastly, studies conducted by Jones et al. (2009) and Bakken et al. (2000) found no statistically significant differences with regard to how employee engagement is experienced by different race/ethnicity groups. For this reason, the researchers find it necessary to determine the factorial invariance of the employee engagement instrument (EEI) across different race groups and to investigate the relationship between race and employee engagement with regard to the current sample.

In the next section of this paper, the research design followed to explore the plausibility of the hypothesis is presented.

4. Research design

Research approach

This research falls within the paradigm of positive psychology (Seligman and Csikszentmihalyi, 2000), focusing on positive organisational outcomes in the South African work environment.

Research method

A scientific quantitative, cross-sectional survey design was used to achieve the research objective. A cross-sectional research design focuses on drawing a sample from the target population at a specific point in time (Babbie, 2013). As the aim of this study was to explore the factorial invariance of the employee engagement questionnaire across the various race groups, the research embarked on was validation research.

Research participants

The database of a research company, made up of approximately 285,000 business people from various cultural and educational backgrounds, industries, sizes of business, job levels and job roles – reflecting the profile of the South African working population – was utilised in this study. The database is known as a permission database, which means that every person whose information is stored in the database gave their permission and indicated that they are willing to participate in research initiatives should they be approached to complete online surveys. Hence, non-probability convenience sampling was used to attract a sample of 1,175 (n = 1,175) respondents from financial institutions. Convenience sampling can be defined as a sample in which only convenient or accessible members of the population are selected (Burt et al., 2009). The demographic profile of the sample is displayed in Table I.

A total of 1,175 completed questionnaires were received. Based on the table above, it becomes apparent that the majority of the sample in terms of race was made up of white employees (62.2 per cent). Furthermore, the majority of the sample consisted of senior management (27.2 per cent), managers (26.7 per cent) and employees (21.3 per cent).
Combined, these three job level categories formed 75.2 per cent of the sample. Additionally, the vast majority of the sample was in possession of postgraduate qualifications (38.6 per cent) and diplomas (19.0 per cent). Finally, male participants made up 51.4 per cent, whilst female participants made up 48.6 per cent of the sample.

**Measuring instrument**

The five-point Likert EEI, developed by Nienaber and Martins (2015), was used to measure employee engagement at both an individual level (individual growth and development) as well as at the organisational and team level (performance quality) (Schaufeli and Bakker, 2004).

The instrument focuses on engagement at an individual level, team level and organisational level and focuses on the individual employee’s work role and the employee’s role as an organisational member. The survey is divided into six dimensions. The first dimension, “Team”, consists of 12 items and focuses on how teamwork enhances employee engagement. The second dimension, “Organisational satisfaction”, which comprises of nine items measuring organisational and job satisfaction among employees. The third dimension, “Customer service”, has six items and measures employees’ perceptions of their customer service and the organisation’s customer service strategy. “Organisational commitment” comprises six items and measures the employee’s commitment to the organisation, their jobs and the organisational strategy. The fifth dimension, “Immediate manager”, is made up of six items and measures how an employee’s immediate manager or supervisor engages the employee through their actions and support. The final dimension, “Strategy and implementation”, comprises 11 items and measures the employee’s perception

<table>
<thead>
<tr>
<th>Variables</th>
<th>Sample</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>197</td>
<td>16.8</td>
</tr>
<tr>
<td>Coloured</td>
<td>100</td>
<td>8.5</td>
</tr>
<tr>
<td>Indian</td>
<td>116</td>
<td>9.9</td>
</tr>
<tr>
<td>White</td>
<td>732</td>
<td>62.2</td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
<td>0.7</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>22</td>
<td>1.9</td>
</tr>
<tr>
<td><strong>Job level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top management</td>
<td>204</td>
<td>17.4</td>
</tr>
<tr>
<td>Senior management</td>
<td>320</td>
<td>27.2</td>
</tr>
<tr>
<td>Manager</td>
<td>314</td>
<td>26.7</td>
</tr>
<tr>
<td>Supervisor</td>
<td>87</td>
<td>7.4</td>
</tr>
<tr>
<td>Employee</td>
<td>250</td>
<td>21.3</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 and below</td>
<td>3</td>
<td>0.3</td>
</tr>
<tr>
<td>Grades 9–10</td>
<td>4</td>
<td>0.3</td>
</tr>
<tr>
<td>Grades 11–12</td>
<td>144</td>
<td>12.3</td>
</tr>
<tr>
<td>Certificate</td>
<td>147</td>
<td>12.5</td>
</tr>
<tr>
<td>Diploma</td>
<td>223</td>
<td>19.0</td>
</tr>
<tr>
<td>First degree</td>
<td>200</td>
<td>17.0</td>
</tr>
<tr>
<td>Postgraduate qualification</td>
<td>453</td>
<td>38.6</td>
</tr>
</tbody>
</table>

**Table I. Demographic representation of the sample**

<table>
<thead>
<tr>
<th>Gender</th>
<th>Sample</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>604</td>
<td>51.4</td>
</tr>
<tr>
<td>Female</td>
<td>571</td>
<td>48.6</td>
</tr>
</tbody>
</table>
of the organisational strategy and the employee’s involvement in strategy development and implementation (Nienaber and Martins, 2014).

A two-fold study by Nienaber and Martins (2015) used exploratory factor and confirmatory analysis to determine the factorial structure of the instrument and subscales. Subsequent to this, Nienaber and Martins (2015) conducted a second study to confirm the construct validity of the measuring instrument by means of confirmatory factor analysis (CFA). Results for the six-factor structure were as follows: root mean square error approximation (RMSEA) = 0.020, goodness-of-fit index (GFI) = 0.828, IFI = 0.923, TLI = 0.917, CFI = 0.923 and PFI = 0.817 (Martins, 2015). The overall reliability of Cronbach’s α coefficient of the employee engagement scale measured at 0.918 and the internal consistency of the subscales ranged from 0.895 and 0.951 (Nienaber and Martins, 2015). According to Nunnally and Bernstein (1994), 0.70 is the recommended minimum cut-off.

Research procedure and ethical considerations
Ethical approval for conducting the study was given by the Research Ethics Review Committee of the College of Economic and Management Sciences of the University of South Africa. Data collection was done by means of an electronic survey that was sent to respondents by means of mass e-mail to invite them to participate in the study. Each respondent was sent a personalised e-mail containing a link to an online survey, informing them of the purpose of the research and inviting them to participate in the survey on an anonymous, voluntary and confidential.

5. Statistical analysis
Exploratory and CFA was used to confirm the factor structure of the instrument. Cronbach’s α coefficients were calculated for each subscale to determine the internal consistency between the items measuring each construct and to evaluate the reliability of the measuring instrument. Pearson’s product-moment correlation coefficients were calculated to determine the relationships between all the subscales of the measuring instrument.

Statistical Programme for Social Sciences (SPSS) (version 22) was used to conduct an exploratory factor analysis (EFA), whilst Analysis of Moment Structures (AMOS) (version 12) was used to conduct SEM in order to determine the factorial invariance of the employee engagement questionnaire across the various race groups.

6. Results
Exploratory factor analysis
To confirm the validity of the questionnaire for financial institutions, an EFA was conducted using principal component analysis (PCA) and was used to explore the factor structure of the EEI (Tinsely and Tinsely, 1987). EFA is used to define and determine the number of continuous latent variables which are used to explain the correlations among a set of observed variables. This is achieved by condensing a large set of variables to obtain a small number of factors. The variables that highly correlate to each other are then grouped together (Pallant, 2001); however, variables that are not very clear or that cross load were eliminated from the analysis. The orthogonal – varimax rotation was performed on the pooled solution (i.e. all the participants were included in the same analysis). SPSS was used to determine whether the factorial structure of the instrument would remain the same, to examine correlations among the items as well as measure the employees’ perceptions of employee engagement in their organisations. The results of the EFA are reflected in Table II.

Subsequently, the factorability of the correlation matrix was examined using Pearson’s product-moment correlation coefficients. An analysis of the distributions indicated that the
<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q61</td>
<td>0.816</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q50</td>
<td>0.797</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q46</td>
<td>0.752</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q45</td>
<td>0.746</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q49</td>
<td>0.737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q52</td>
<td>0.728</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q47</td>
<td>0.724</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q48</td>
<td>0.703</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q61</td>
<td>0.647</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q59</td>
<td>0.637</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.409</td>
</tr>
<tr>
<td>Q59</td>
<td>0.637</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q63</td>
<td>0.632</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q56</td>
<td>0.821</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q55</td>
<td>0.808</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q57</td>
<td>0.807</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q52</td>
<td>0.714</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q57</td>
<td>0.710</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q58</td>
<td>0.636</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q68</td>
<td>0.591</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q36</td>
<td>0.659</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q39</td>
<td>0.655</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q38</td>
<td>0.410</td>
<td>0.653</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q37</td>
<td>0.574</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q35</td>
<td>0.515</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q34</td>
<td>0.509</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q41</td>
<td>0.416</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q54</td>
<td>0.539</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q43</td>
<td>0.498</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q53</td>
<td>0.471</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q42</td>
<td>0.464</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q42</td>
<td>0.427</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q42</td>
<td>0.404</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table II.
Rotated component matrix
notion of normality, linearity and homoscedasticity were not dishonoured. A review of the correlation matrix showed coefficients of 0.3 and above for the majority of the constructs. The Kaiser–Meyer–Olkin (KMO) value calculated was 0.976, which is well above the minimum value of 0.50. KMO values between 0.7 and 0.8 are good, values between 0.8 and 0.9 are great and values above 0.9 are excellent (Hutcheson and Sofroniou, 1999). Bartlett’s (1954) test of sphericity reached high statistical significance, \( p < 0.001 \), showing that the correlations within the \( R \)-matrix were adequately different from zero to warrant factor analysis and, therefore, supports the factorability of the correlation matrix (Field, 2005). Results in Table II show that 50 items were initially subjected to the PCA; however, four of the items (Q24, Q25, Q38 and Q60) were removed as it cross-loaded on two factors, as per Table II. Following the EFA process, five additional items (Q44, Q53, Q55, Q63 and Q65) did not fit the with the new factor structure as suggested by the EFA results, and as a consequence they were removed. The initial 50 items resulted in a six-factor structure and explained 66.734 per cent (Table III) of the variance in the data.

Subsequent to the EFA, the researchers investigated the new factor structures. Upon investigation, the researcher found that the constructs of “Strategy implementation” and “Organisational satisfaction” in the original instrument were no longer applicable and, therefore, renamed these to “Nature of my job” and “Job satisfaction”, respectively.

**Descriptive statistics and Cronbach’s \( \alpha \) coefficients**

Table IV reflects the descriptive statistics, \( \alpha \) coefficients and Pearson’s product-moment correlations of the EEI. According to Odendaal and Roodt (1998), research suggests that mean scores above an average of 3.20 can be considered as a reasonable cut-off score to differentiate between positive and negative perceptions. Scores above 3.20 were accepted as positive, whereas scores below 3.20 were accepted as negative scores.

Overall, the mean scores obtained for the various dimensions were reasonably high, with only two factors reporting average mean scores. The Team dimension reported the highest mean \( (m) \) score of 4.10 and the lowest standard deviation (SD) of 0.65; whereas Nature of my job reported the lowest mean score of 3.42 and a standard deviation of 0.88. The highest standard deviation of 0.97 was obtained for Immediate manager, which means that the vast majority of the responses were scattered/deviated from the mean.

The \( \alpha \) coefficients for all the dimensions were greater than the cut-off of 0.70 (Nunnally and Bernstein, 1994). The overall Cronbach coefficient value for the employee engagement scale was 0.970, indicating internal consistency. Job satisfaction produced the highest

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial eigenvalues</th>
<th>Extraction sums of squared loadings</th>
<th>Rotation sums of squared loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of variance</td>
<td>Cumulative %</td>
<td>% of variance</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>1</td>
<td>22.882</td>
<td>45.765</td>
<td>45.765</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3.911</td>
<td>7.821</td>
<td>53.586</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.301</td>
<td>4.601</td>
<td>58.187</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.133</td>
<td>4.267</td>
<td>62.453</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.109</td>
<td>2.218</td>
<td>64.671</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.032</td>
<td>2.063</td>
<td>66.734</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.900</td>
<td>1.921</td>
<td>68.655</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.852</td>
<td>1.703</td>
<td>70.358</td>
</tr>
<tr>
<td>9</td>
<td>0.786</td>
<td>1.572</td>
<td>71.930</td>
</tr>
<tr>
<td>10</td>
<td>0.711</td>
<td>1.423</td>
<td>73.333</td>
</tr>
</tbody>
</table>

**Note:** Only the top section of the table is displayed
Cronbach’s α value of 0.951, whereas Customer service produced the lowest Cronbach’s α value of 0.763. Based on these results, it becomes evident that the EEI demonstrates acceptable reliability.

Results in Table IV further indicate that all factors are statistically and practically significantly related to each other with Organisational commitment and Nature of my job exhibiting the strongest correlation ($r = 0.772$) and Organisational commitment and Team exhibiting the weakest correlation ($r = 0.523$).

An overall race structural equation model
CFA, using the SEM, was conducted to develop, specify and determine the resultant measurement model on the first-order latent construct level. CFA is typically performed using sample covariances rather than the correlations used in EFA. AMOS was used to conduct the CFA using the six factors identified during the EFA process.

According to Byrne (2010), SEM is a statistical tool of multivariate analysis techniques, which shows the relationships between different variables through the use of measurement equations and structural equations. Measurement equations refer to the process of testing the accuracy of the proposed model by evaluating the relationship between latent variables and their indicators. Structural equations allow for the testing of statistical hypotheses by evaluating the hypothesised relationships between the latent variables (Byrne, 2010). SEM, therefore, facilitates the determination and confirmation of relationships among multiple variables (Hair et al., 2014).

As a requirement for invariance testing, it is necessary to consider the baseline model, which is then estimated separately for each of the different groups. This process involves collecting data to test whether the items of the scale comprise good indicators of a given latent construct. In order to assess factorial invariance, multi-group invariance analyses are conducted (Byrne, 2004). When determining factorial invariance, the baseline model is compared with the observed structure of two or more variables (dependent or independent). Joreskog’s strategy for measuring compatibilities of structures is often followed to measure for invariance (Milfont and Fischer, 2010).

The baseline measurement model, which will be used for comparison purposes, is depicted in Figure 1.

To determine the validity of the model, the researchers analysed the various fit indices that demonstrate how well the a priori model fits the sampled data and shows how the suggested model has the most superior fit. These fit indices provide an indication of how well the suggested theory fits the sampled data, whereas the incremental fit indices calculations demonstrate how well the model fits in comparison to the null model.
(Jöreskog and Sörbom, 2003). The absolute category includes the $\chi^2$ test, GFI, adjusted GFI (AGFI) and RMSEA.

The results of the CFA using SEM are presented in Table V. The EEI was also subjected to a CFA using SEM, of which the results are reflected in Table VI. Absolute fit indices obtained for the baseline model showed a GFI statistic of 0.893 and an RMSEA statistic of 0.048. In terms of the incremental fit indices, results yielded an IFI of 0.947, TLI of 0.942 and a CFI of 0.947. Parsimony adjustment statistics showed a PNFI of 0.844 and RCFI of 0.860. Based on the criteria reflected in Table V, it becomes evident that the adapted EEI demonstrates acceptable validity in financial institutions.

**Multi-group invariance**

Testing for factorial invariance includes a sequence of ordered steps, starting with the creation of a baseline model for each group, followed by tests for metric invariance across groups at each of several progressively more stringent levels (Byrne and Stewart, 2006). Many approaches have been established for testing for factorial invariance. The most commonly used process is the multi-stage, multi-group CFA (Byrne, 2004). This process is referred to as forward or sequential constraint imposition approach and focuses on testing for factorial invariance across groups by investigating the $\chi^2$ difference test ($\Delta \chi^2$) between two nested models; one unconstrained model (invariance not assumed) and one constrained...
model (invariance is assumed), based on specific measurement weights, structural weights, structural covariances, structural residuals and measurement residuals being equal across groups (Dimitrov, 2010).

According to Meade and Lautenschlager (2004), measurement invariance (equivalence) focuses on establishing whether a construct has the same meaning under different conditions (e.g. consistency over different sample groups). Hypotheses with regard to the measurement invariance across different groups are fitted to multiple samples using CFA. This is conducted by simultaneously fitting the covariance matrices from at least two independent samples. The model is then fitted by specifying the same measurement model across the different groups. When running this model, both the factors and the factor indicators are the same; however, all the parameters are freely estimated for each of the sample groups. To determine the extent of the invariance between the different race groups the baseline model is fitted separately for each group. The results from this statistical analysis are presented in Table VI.

### Table V. Goodness-of-fit baseline model

<table>
<thead>
<tr>
<th></th>
<th>African</th>
<th>Coloured</th>
<th>Indian</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$</td>
<td>1,364.949</td>
<td>1,340.280</td>
<td>1,568.417</td>
<td>2,307.584</td>
</tr>
<tr>
<td>$\chi^2$ degrees of freedom</td>
<td>745</td>
<td>745</td>
<td>745</td>
<td>745</td>
</tr>
<tr>
<td>CMIN/DF</td>
<td>1.832</td>
<td>1.799</td>
<td>2.105</td>
<td>3.097</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>GFI index</td>
<td>0.726</td>
<td>0.637</td>
<td>0.627</td>
<td>0.863</td>
</tr>
<tr>
<td>RMSEA</td>
<td>0.065</td>
<td>0.090</td>
<td>0.098</td>
<td>0.054</td>
</tr>
</tbody>
</table>

### Table VI. Parsimony adjustment measures

<table>
<thead>
<tr>
<th></th>
<th>African</th>
<th>Coloured</th>
<th>Indian</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNFI</td>
<td>0.737</td>
<td>0.637</td>
<td>0.645</td>
<td>0.826</td>
</tr>
<tr>
<td>RCFI</td>
<td>0.821</td>
<td>0.761</td>
<td>0.746</td>
<td>0.851</td>
</tr>
<tr>
<td>Participants</td>
<td>197</td>
<td>100</td>
<td>116</td>
<td>732</td>
</tr>
</tbody>
</table>
The goodness-of-fit indices are displayed in Table VI. The four different race groups indicate the following:

- The GFI s for all four race groups are below the suggested cut-off of 0.90 with the white race showing the highest GFI of 0.863, with the Indian race indicating the lowest GFI of 0.627. This means that for all race groups there is mediocre fit in terms of the GFI.

- The incremental fit indices for the white race group are all above the recommended cut-off of 0.90, whereas IFIs for coloured and Indian race groups are all below the suggested cut-off of 0.90. The African race group shows acceptable IFI (0.904) and CFI (0.903), with an IFI (0.894), which is below the cut-off of 0.90.

- The PNFI and RCFI for all race groups are below the recommended 0.90, with the higher values indicating better fit.

- The African race (0.065) and the white race (0.054) were the only two of the four race groups that show a RMSEA below the recommended cut-off of 0.08.

It is important to note that even if the model fits well for each group (which is not the case in this study), it is still necessary to conduct the multi-group CFA as it provides a comparison standard for subsequent tests (Byrne, 2004; Milfont and Fischer, 2010). Invariance between the race groups is then tested by constraining the factorial structure to be exactly the same across the four groups (Byrne, 2004). If the specified model does not fit the data, measurement/configural invariance has not been established (Byrne, 2004; Kline, 2011; Milfont and Fischer, 2010).

Structural invariance measures whether the different sample groups understand and respond to items in the same manner; in other words, it measures whether the strengths of the relations between the items and their specific fundamental construct are the same throughout the various groups. If structural invariance is obtained, the calculated ratings can be compared across groups and the observed item differences will specify group differences in the underlying latent construct. According to Vandenberg and Lance (2000), the establishment of partial structural invariance should be established before continuing with other invariance testing methods (e.g. error variance invariance, scalar invariance). This model is tested by constraining all factor loadings to be the same across groups (Milfont and Fischer, 2010). The results of the goodness-of-fit indices are portrayed below.

Table VII depicts the absolute fit, incremental fit and parsimony adjustment measures indices for the baseline model, the unconstrained model (across all race groups) and the constrained model (across all race groups). The CMIN fit statistic improved from 3.275 to 2.213 in the unconstrained model and 2.174 in the constrained model. The GFI index decreased from 0.893 to 0.790 in the unconstrained model and 0.788 in the constrained model. The AGFI index decreased from 0.876 to 0.758 in the unconstrained model and 0.764 in the constrained model. The RMSEA fit index improved from 0.048 to 0.033 in the unconstrained model and 0.032 in the constrained model. The CFI fit statistic deteriorated from 0.947 to 0.908 in both the constrained and unconstrained models. TLI deteriorated in from 0.942 to 0.899 in the unconstrained model and 0.902 in the constrained model.

Following the above, the researchers saw it fit to determine whether the measurement weights $\chi^2$ ($\chi^2$ differences) model tested significance when comparing it to the baseline and unconstrained model. The results obtained from this analysis are displayed in Table VIII.

Table VIII shows that the $\chi^2$ change from the default model across all four race groups to the constrained model is insignificant; $\chi^2 = 105.617, p = 0.335$. The null hypothesis of factorial invariance across the four race groups can, thus, be rejected as multi-group invariance can be assumed.
The above results, across the four different race groups, do not indicate significant differences with regard to the measurement weights of the latent constructs as these relate to the items. It can, thus, be assumed that for all four race groups the constructs were formed in the same way.

Decisions regarding the research hypotheses

Conclusions with regard to the hypotheses of the study are based on the results as discussed above.

The results indicate that the null hypothesis ($H_0$) of race groups display variance with regard to the EEI is rejected as multi-group invariance (equivalence) can be assumed across the four different race groups.

### 7. Discussion

#### Outline of the results

In order for the researchers to determine the factorial invariance of the instrument, the researchers first had to establish the reliability and validity of the instrument. To achieve this, an EFA was conducted using PCA to explore and determine the factor structure of the instrument. Based on the results obtained from the EFA, nine items were deleted from the original instrument, which resulted in a 41-item instrument. Following the deletion of these items, the researchers reviewed the remainder of the items and the corresponding factors and found that Strategy and implementation and Organisational satisfaction, as depicted in the original EEI, were no longer relevant and were, thus, renamed to Nature of my job and Job satisfaction, respectively.

The deletion of the items as indicated above resulted in a 41-item instrument and 6 dimensions. These six dimensions were then tested for reliability using Cronbach’s $\alpha$ coefficients. The EEI demonstrated acceptable reliability overall, as did the individual subscales.

#### Table VII.
Goodness-of-fit indices

<table>
<thead>
<tr>
<th>Measurement weights</th>
<th>Baseline model</th>
<th>Unconstrained model</th>
<th>Constrained model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$</td>
<td>2,775.451</td>
<td>6,965.406</td>
<td>6,706.023</td>
</tr>
<tr>
<td>$\chi^2$ (CMIN)</td>
<td>745</td>
<td>2,980</td>
<td>3,085</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GFI index</td>
<td>0.893</td>
<td>0.790</td>
<td>0.788</td>
</tr>
<tr>
<td>AGFI</td>
<td>0.876</td>
<td>0.758</td>
<td>0.764</td>
</tr>
<tr>
<td>RMSEA</td>
<td>0.048</td>
<td>0.033</td>
<td>0.032</td>
</tr>
</tbody>
</table>

#### Table VIII.
Model comparison for four races

<table>
<thead>
<tr>
<th>Model</th>
<th>DF</th>
<th>CMIN</th>
<th>$P$</th>
<th>NFI $\delta_1$</th>
<th>IFI $\delta_2$</th>
<th>RFI $\rho_1$</th>
<th>TLI $\rho_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement weights</td>
<td>105</td>
<td>110.617</td>
<td>0.335</td>
<td>0.003</td>
<td>0.003</td>
<td>$-0.003$</td>
<td>$-0.003$</td>
</tr>
</tbody>
</table>
Subsequent to determining the reliability of the instrument, the instrument was subjected to a CFA using SEM. All indices displayed acceptable fit, except the GFI statistic, which proves to be worrisome as it is below the recommended cut-off of 0.90; however, the covariance matrix predicted by the model still explained about 89.3 per cent of the total variability in the sample covariance matrix and the relative fit of the model shows about 94.7 per cent improvement over the independence model fit. Based on these results, it is evident that the instrument demonstrates acceptable construct validity. The data were, thus, used to proceed with invariance testing among the four race groups.

With regard to the goodness-of-fit indices for the different race groups, all four race groups reported a GFI of less than 0.90, with the white race group showing the best fit in terms of GFI, and the Indian race group showing the least favourable fit. In terms of the incremental fit indices, the white race group yielded an IFI, TLI and CFI above the recommended cut-off of 0.90. The African race group yielded an acceptable CFI and TLI, but an IFI below the suggested cut-off of 0.90. The coloured and Indian race groups both yielded incremental fit indices (IFI, TLI and CFI) below the recommended cut-offs. The African and white race groups were the only two of the four race groups which yielded acceptable RMSEAs that are below the recommended cut-off of 0.08. Poor fit indices of the coloured and Indian race group could be as a result of the small sample sizes, as some indices (i.e. $\chi^2$ tests, GFI and RMSEA statistics) are particularly sensitive to small sample sizes. These results, therefore, suggest that for coloured and Indian race groups, the instrument displays poor model/measurement fit, for the African race group the instrument indicates mediocre model/measurement fit and for the white race group the instrument displays good model/measurement fit.

To determine whether the constrained model tested significance when comparing it to the default and unconstrained model, a model comparison was conducted. Results showed that the $\chi^2$ change from the default model across all four race groups to the measurement weights is insignificant. These results, therefore, suggest that the EEI demonstrates multi-group invariance across the different race groups as the instrument does not indicate significant differences with regard to the measurement weights, therefore, it can be assumed that for all four race groups the constructs were formed in the same way. The implication of these results is that the EEI can be used with confidence across race groups in the financial sector in South Africa. As discussed in the introduction, there appears to be a paucity of research on testing for measurement invariance of psychometric tools for different race groups, which is now been addressed in this research in terms of the EEI. Human resources practitioners and organisational development practitioners can now use the EEI with confidence in their financial institutions.

**Limitations of present study**

It is important to note certain limitations of this study. The first limitation is that a cross-sectional research design was utilised in this study, which does not make provision for the measurement of variables over a period of time and does not allow for generalisation of the results.

A second limitation is with regard to the unequal distribution of the race groups. The over-representation of the white race could have resulted in the results being skewed, affecting the reliability and validity of the data.

A third limitation can be attributed to the non-probability, convenience sampling method used to collect data. This sampling method prohibits the generalisation of results to the larger population.

**Recommendations for future research**

Race groups can further be split into various different ethnic, cultural and language groups and future research projects can, therefore, focus on determining the reliability and validity
of the EEI for different ethnic, cultural and language groups. A final recommendation for future research is the use of a longitudinal research design to evaluate and determine the effect that time and changing business environments have on employee engagement.

**Practical implications**
It is important for organisations to take cognisance of how particular socio-demographic variables influence employee engagement and the subsequent organisational commitment, job performance and motivation. By understanding how different employees are engaged it enables organisations to customise their engagement programmes to meet the needs of the various types of employees within the organisation instead of applying a “one size fits all” approach to engagement programmes.

Based on the results of the present study and the literature it becomes apparent that the constructs measured by the EEI play an important role in the effective functioning of both individuals and organisations.

8. Conclusions
In summary, it can be concluded that the adapted EEI demonstrates acceptable reliability and validity. This means that in the financial institutions, interpretation of the results can be done with confidence. Furthermore, the results also indicate that the EEI can be used with confidence to measure employee engagement across the four different (African, coloured, Indian and white) race groups in the financial sector.

References


About the Authors
Wesley Herscelle Gallant was Masters’ Student at the Department of Industrial and Organisational Psychology at the time of the research. He was also busy with his internship at the Military Psychological Institute (MPI) in Pretoria. The MPI is tasked to conduct all military psychological research and development within the Department of Defence in South Africa.

Nico Martins holds a PhD in Industrial Psychology and is with the Department of Industrial and Organisational Psychology at the University of South Africa. Nico specializes in the field of Organisational Psychology and is currently employed as Research Professor. His research focuses on organisational culture/climate, organisational trust, employee engagement and organisational diagnoses. Nico Martins is the corresponding author and can be contacted at: martinmsun@mweb.co.za
Best supply chain management practices and high-performance firms

The case of Gulf manufacturing firms

Moh’d Anwer AL-Shboul
King Talal School of Business Technology, Business Administration Department, Princess Sumaya University for Technology (PSUT), Amman, Jordan

Jose Arturo Garza-Reyes
Centre for Supply Chain Improvement, The University of Derby, Derby, UK, and

Vikas Kumar
Bristol Business School, University of the West of England, Bristol, UK

Abstract

Purpose – The purpose of this paper is to investigate the best supply chain management (SCM) practices that are implemented in medium- and large-sized Gulf manufacturing firms (MFs).

Design/methodology/approach – This study has explored seven SCM practices, i.e. supplier collaboration, flexibility with partners, usage of internet, customer focus (CF), lean production (LP), internal integration (II) and quality management (QM). It assumes that the best-performing firms must be the ones implementing the best practices. t-test and multiple linear regression analyses were used to establish the best practices, implemented by medium and large-sized Gulf MFs.

Findings – The results showed that QM, CF and supplier collaboration are considered as best SCM practices in Gulf MFs. Usage of internet may have been the best practice previously, but not anymore. LP cannot yet be qualified as, but may develop into the best SCM practice.

Practical implications – The study provides a useful contribution to the field of best SCM practices as it provides better decision-making insights and a benchmarking base to top managers, policy makers and academics. It is likely to result in increased overall performance of their firms.

Originality/value – The study provides an understanding of the distinctive characteristics of the best SCM practices, implemented by Gulf MFs. It has broader implications for all MFs, particularly in developing economies where the growth of manufacturing and effective management of their supply chains is a key element for the economic development.

Keywords Best practices, Manufacturing, Supply chain management, Gulf countries,

Medium to large-sized firms

Paper type Research paper

1. Introduction

Globalization, stringent quality requirements and intense competition have forced most manufacturing firms (MFs) to improve their performance by re-examining how they get products into their customer’s hands, and how they can quickly respond to customer’s needs in a constantly changing environment. Therefore, a prerequisite for manufacturers enhances profitability and remains competitive in the current global dynamic market to understand and practice supply chain management (SCM) (Cook et al., 2011). It has contributed for SCM and recognized it as an important field that has generated extensive interests among industrialists and scholars, literally invading world’s businesses (Ou et al., 2010; Li et al., 2006).

SCM is considered as a multidisciplinary field that has been explored from many different perspectives (Papakiriakopoulos and Pramatari, 2010). The practices of SCM are proposed to be a multi-dimensional concept, including downstream and upstream sides of
the supply chain. The concept of SCM has been considered from two alternative perspectives that include: purchasing and supply management. These perspectives emphasize purchasing and materials management as a basic strategic business processes, rather than a narrow-specialized supporting function (Narasimhan et al., 2004), transportation and logistics management. It mainly focuses on integrated logistics systems (e.g. inventory management, vendor relationships, transportation, distribution, warehousing and delivery services) that lead to inventory reduction both within and across firms in the supply chain (Cook et al., 2011).

The best supply chain management practices (SCMPs) have become an essential factor for low-performing firms to remain competitive in the global race (Okongwu et al., 2015). The nature of SCMPs will be able to explain the dual purpose of SCM as it improves the performance of an individual firm as well as the performance of the whole supply chain. This could be achieved through the effective adoption and construction of the best SCM practices (Kim, 2006). Apart from some of the firms, many of them still do not exactly know what best practices to implement, due to a lack of understanding of what constitutes a comprehensive set of SCM practices (Li et al., 2005, 2006). The best SCM practices can positively impact on performance (Tan, 2002). The degree of attention paid to SCM has increased in developing countries since the last two decades; however, the Gulf countries in particular are still behind. Many national entrepreneurs and managers of industrial conglomerates in Gulf countries generally ignore the concept of SCM (Abu-Alrejal, 2007). This phenomenon has halted the manufacturing industry of these countries from adopting and developing practices that enable the effective management of their supply chains.

SCMPs are implemented to achieve and enhance performance by enabling an internal cross-functional integration within the firm, and external integration with suppliers and customers (Kannan and Tan, 2010; Kim, 2006). The seven SCMPs considered in this study (i.e. supplier collaboration, flexibility with partners, usage of internet, customer focus (CF), lean production (LP), internal integration (II) and quality management (QM)) were developed, tested and validated in the literature by researchers, such as Li et al. (2006), Green Jr et al. (2008), Tan (2002) and Cook et al. (2011). These practices are considered crucial, and they cover both upstream and downstream sides of the SC. The study has addressed an importance theoretical gap addressing the lack of empirical studies investigating and/or examining the best SCMPs of high-performing Gulf MFs’ and its impact on performance.

There is a limited knowledge available on the impact of which best practices of high-performing firms affect their performance, although current literature gives a clear understanding of the link between SCM practices and performance. There is a lack of empirical evidence on the relationship between high-performers and best SCMPs, and which SCMPs enhance firms’ performance. Managers and practitioners in the Gulf region are still seeking to identify the best SCMPs, in which they should focus more to enhance firm performance. The insights would provide an opportunity for Gulf managers to effectively utilize the other practices (current non-best practices) to enhance MF’s performance. This study has attempted to focus on the best SCMPs of high-performing Gulf MFs. Thus, the study aims to investigate the current SCMPs that are commonly implemented by high-performing medium and large-sized Gulf Manufacturing Firms (GMFs), and determine the best practices that have the most significant effects on the performance of these firms. Following research question has been developed on the basis of aim:

RQ1. Which best SCM practices are implemented by high-performing GMFs (both medium and large-sized firms)?

The rest of the paper is organized as follows: Section 2 presents the literature review. Theoretical research framework, including the definitions and theory underlying each dimension of the SCMPs, best SCM practices and MFs’ performance constructs are
discussed in Section 3. Section 4 presents the research methodology. The results of the study are then presented in Section 5. Section 6 includes the discussion of measurements model. Finally, Section 7 provides the main conclusions of the research, identifies the main limitations and outlines the future research directions derived from this work.

2. Literature review
MFs have been the backbone of economic growth of many nations by driving industrial development. They play an important role in national economies by providing job opportunities and supporting larger industries (Anuar and Yusuff, 2011). To sustain these contributions, MFs must not only become increasingly advanced in their technologies and manufacturing processes, but also, they should adopt world-class SCMPs. Sandhu et al. (2013) stated that SCM practices are regarded as operational functions and main activities in the firm, which determines the effectiveness and efficiency of its supply chain. The main goal of SCM concept is to enhance the long-term competitive firms’ performance and their supply chains by integrating their functions, processes, and operations internally and externally with other partners. These partners mainly include the suppliers, manufacturers, distributors and customers (Kim, 2006). SCM encompasses various activities such as planning and management, procurement, coordination, collaboration, outsourcing and all other logistics management activities with other channel partners (Soosay et al., 2008). Majority of the studies have emphasized that the ultimate goal of SCM is to enhance and improve the performance of firms (Li et al., 2005, 2006; Chen and Paulraj, 2004; Min and Mentzer, 2004).

Donlon (1996) identified several components and elements of SCM practices, which include supplier partnering, process flow, outsourcing and information sharing. These main elements are considered as the evolution of SCM practices in the MFs in the last decade. The empirical work of Sundram et al. (2016) classified seven important SCM practices, namely, supplier strategic partnership, customer relationship, information sharing, information quality, postponement, agreed vision and goals and risk and rewards sharing. Developing strategic partnerships in the supply chain was also emphasized by Christopher and Jüttner (2000). Whereas, Alvarado and Kotzab (2001) selected customization and information technology through postponement activities as an important factor of SCM practices. Tan (2001) ensured that information sharing among trading partners in supply chain such as customization and postponement are crucial SCM practices that emphasize a well-integrated supply chain. Tan et al. (2002) included six dimensions in their study, namely, supply chain integration, just-in-time capability, customer service management, geographical proximity and information sharing. Whereas, the study of Chen and Paulraj (2004) included several other dimensions, such as communication, supplier involvement, supplier base reduction, cross-functional teams and long-term relationships to measure SCM practices.

The current commercial and competitive situation of GMFs and the lack of specific studies on best SCM practices in this region justify the opportunity of studying this phenomenon in its own right. Seven SCMPs for this study were selected, developed, tested and validated by many researchers in the SCM literature, such as Cook et al. (2011), Green Jr et al. (2008), Li et al. (2006) and Tan (2002). These selected practices by authors are considered crucial and covers both upstream and downstream through the SC. SCM practices are considered as a perfect recipe for the success of several firms from various industries (Gorane and Kant, 2015). The medium- and large-sized MFs are now under increasing pressure due to various international trading and commercial agreements, which have forced them to improve their competitiveness. Majority of the Gulf Manufacturing Firms (GMFs) that used to compete based on supply chain-oriented factors, i.e. flexibility, serviceability and responsiveness (Gunasekaran, 2003).
The Gulf MFs will be able to improve their performance, expand their assets, provide work opportunities and contribute to the economic growth of the Gulf region by identifying and developing best SCM practices. Intensified competition, fast technological development, shortening product life cycle, increasingly customized products and volatility in input prices have created a dynamic environment, where manufacturers should be more flexible, adaptive and responsive to fulfill their customer orders (Arlbjørn and Mikkelsen, 2014; Anuar and Yusuff, 2011). The identification, adoption and continuous development of “best practices” are likely to result in a superior business capability, which will ultimately lead to increased competitiveness (Laugen et al., 2005). Table I has summarized certain studies that have focused on studying different types of best practices and their effect on various dimensions of performance.

3. Theoretical research framework
Gulf MFs have been classified into two categories on the basis of their performance: low- and high-performing firms. This study is based on the assumption that best-performing Gulf MFs are those that possess the best practices. Figure 1 has illustrated the theoretical research framework developed from the review of literature to understand the antecedents and consequences brought by the casual relationship between SCMPs and the performance

<table>
<thead>
<tr>
<th>Best Practice (BPs)</th>
<th>Effect on</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous improvement (CI), just-in-time (JIT) and total quality management (TQM)</td>
<td>World-class manufacturing (WCM) and competitive performance</td>
<td>Flynn et al. (1999)</td>
</tr>
<tr>
<td>Cross-functional and cooperation</td>
<td>Better performing manufacturing managers such as team building and support</td>
<td>Kathuria and Partovi (1999)</td>
</tr>
<tr>
<td>Customer focus, employee focus, community focus and productivity focus</td>
<td>Customer retention and time to market</td>
<td>Fazli (2011)</td>
</tr>
<tr>
<td>Supplier involvement, facility control and vendor and material management</td>
<td>SCM performance</td>
<td>Sridharan et al. (2005) and Stevenson (2005)</td>
</tr>
<tr>
<td>Agreed metrics, good/integrated IT support for data capturing and reporting</td>
<td>Achieving adequate and accurate data for reporting and take decision</td>
<td>Papakirikopoulos and Pramatari (2010)</td>
</tr>
<tr>
<td>Customer focus and Customer service management</td>
<td>Time delivery and customer satisfaction</td>
<td>Forslund and Jonsson (2010)</td>
</tr>
<tr>
<td>TQM, JIT, WCM, contingency Product operations, production process</td>
<td>Competitive advantage, Production process flexibility, elimination of waste and response time</td>
<td>Ketokivi and Schroeder (2004) and Grando and Belvedere (2005)</td>
</tr>
<tr>
<td>Marketing strategy</td>
<td>Quality of relationships</td>
<td>Stevenson (2005)</td>
</tr>
<tr>
<td>Supplier and customer relationship</td>
<td>Quality of relationships</td>
<td></td>
</tr>
<tr>
<td>Information sharing, strategic supplier partnership</td>
<td>Sharing, share forecasts with customers and performance</td>
<td>Hsu et al. (2008)</td>
</tr>
<tr>
<td>Quality of information, level of information</td>
<td>Quality control, quality cost, best practices, performance</td>
<td>Ou et al. (2010), Li et al. (2006)</td>
</tr>
<tr>
<td>Collaboration</td>
<td>Developing a collaborative culture</td>
<td>Dotti et al. (2012), Ferreira et al. (2012), Ferreira et al. (2012)</td>
</tr>
<tr>
<td>Shared goals and specific targets</td>
<td>Achieving the setting targets and goals</td>
<td></td>
</tr>
</tbody>
</table>
3.1 Supply chain management practices (SCMPs)

The seven dimensions of SCMPs cover upstream (supplier collaboration (SCMP/SC)) and downstream (customer focus (SCMP/CF)) sides of a supply chain, information flow within and across a supply chain (usage of internet (SCMP/UI)), and internal supply chain process (flexibility with partners (SCMP/FwP), lean production (SCMP/LP), internal integration (SCMP/II) and quality management (SCMP/QM)) (Figure 1). Although the seven dimensions capture the major aspects of SCM practices, the conceptualisation cannot be considered an “all-inclusive list” as there may be some other factors (geographical proximity, logistics integration, cross-functional teams, etc.) that also play a significant role in the management and improvement of supply chains. A number of researchers have converged on the fact that the ultimate goal of SCMPs is to improve the performance of firms, although these have been denoted differently and from a multidisciplinary perspective (Cook et al., 2011; Collins et al., 2010; Ou et al., 2010; Reyes and Giachetti, 2010; Koh et al., 2007; Li et al., 2006). SCM practices have been defined as the approaches and activities adopted by firms to effectively and efficiently manage the coordination of their supply, demand and their relationships to meet their customers’ expectations (Li et al., 2005). Table II has listed these dimensions of SCMPs along with supporting literature.

3.1.1 Supplier collaboration (SC). SC is a phenomenon that depicts strong and close relationship between a firm and its suppliers (Li et al., 2006). It is a practice that focuses on their direct and long-term association, mutual planning and problem-solving efforts (Dirks and Verdaasdonk, 2009). The firms are able to share benefits, information and participation in one or more key areas (vendor-managed inventory, continuous replenishment, improve product quality and lead time) through strong SC (Cook et al., 2011). The main aim of SC is to improve overall firm performance, reduce cost, increase profit and improve forecast accuracy (Grudinschi et al., 2014). Dotti et al. (2012) determined SC as a best practice and highlighted its importance for the effective and successful management and improvement of supply chains.

3.1.2 Flexibility with partners (FwP). Flexibility is the degree to which a firm is able to adjust the time in which it can ship or receive goods (Tang and Tomlin, 2008). According to Koh et al. (2007), it is defined as “the firm’s ability to adapt to changes in its environment.” There is a general consensus about the fact that supply chains which are flexible, contribute to the competitiveness of firms, despite of the contradicting views, in terms of supply chain
Therefore, supply chain flexibility, and more specifically flexibility with partners, is regarded as a crucial SCM practice. These practices give the firm an opportunity to increase flexibility by generating alternative sourcing for procurement, and reducing supply chain risks (Iddris et al., 2016).

The enhancement in manufacturing flexibility does not improve performance but in some situations, it could actually lead to negative results (Chang et al., 2003). In contrast, another research stream has determined and confirmed the positive effect of flexibility on the performance of firms (Tannous, 1996).

### 3.1.3 Usage of the internet (UoI)

Rapid technological developments have enabled firms to link the operations of their departments both internally with those of other departments and externally with those of their partners. In particular, the internet has served as one of the main technological developments supporting increased coordination and collaboration among supply chain partners (Cook et al., 2011). Gimenez and Lourenco (2008) found in their study that the effect of internet on SCM has been recognized as an important topic of research through e-procurement, information flows and e-fulfillment being the fundamental areas of the research. The use of internet, its growth in recent years and the importance of this factor in supply chain research has been covered in both supply and customer sides integration practices using web technology, and web-based marketing-oriented applications (Gimenez and Lourenco, 2008). It is important to analyze this factor, its significance and its unique role as a supply chain practice in affecting a firm’s performance. Therefore, the majority of MFs have continued to adopt internet-based collaboration to let them be able to take effective decisions in regards to inventory, forecasts and customers' orders.

### 3.1.4 Customer focus (CF)

The purpose of CF is to fulfill customer’s expectations, develop customer’s loyalty to the company products and services, manage customer complaints, follow up sales after delivery, improve customer satisfaction and build long-term relationships with the customers (Reyes and Giachetti, 2010). CF is concerned with planning, implementing and evaluating successful services and relationships between providers and recipients in both upstream and downstream of SCM. It deals with the ability
to communicate delivery of the right products and services to customers locally and globally at the right time, right place and right quantity with correct invoices (Li et al., 2006; Kim et al., 2006). The studies conducted by Ou et al. (2010), Collins et al. (2010) and Reyes and Giachetti (2010) have emphasized on the importance of CF practices in supply chain operations.

3.1.5 Lean production (LP). LP is defined as manufacturing without waste, which tries to remove out the unnecessary costs, time and other wastes from the entire supply chain (Taj, 2008; Boyle and Rathje, 2009). LP mainly focuses on the identification and elimination of waste throughout the product’s entire value stream. It not only extends within the organization, but also along its entire supply chain network. It results in improved output and quality levels, and achieves this using fewer resources, such as raw materials and employee effort (Belekoukias et al., 2014; Boyle and Rathje, 2009). The lean supply chains and application of lean thinking tends to improve the logistic operations (Garza-Reyes et al., 2016; Villarreal et al., 2016).

3.1.6 Internal integration (II). Integration is now widely considered a core practice that enables the success of firms, because it allows the integration of processes across different departments that includes sourcing, manufacturing and distribution (Ellegaard and Koch, 2012). An II is defined as the extent to which separate parties work together in a cooperative manner to arrive at mutually acceptable outcomes (O’Leary-Kelly and Flores, 2002). According to Narasimhan and Jayaram (1998), II involves the coordination, cooperation and collaboration between all internal functions within the firm from raw material management through production, shipping and sales. Ellegaard and Koch (2012) have recognized the positive impact of II and considered it as an important practice for the effective management of supply chains and the successful overall performance of organizations (Ellegaard and Koch, 2012).

3.1.7 Quality management (QM). The literature shows that many MFs consider quality of products as the main factor to drive their competitiveness as it refers to the ability of a product or service to consistently meet the customer expectations (Anuar and Yusuff, 2011; Reyes and Giachetti, 2010). QM has been recognized as a competitive advantage and one of the most important ways to respond rapidly, correctly and profitably to market demands in the digital world (Ou et al., 2010). Various previous empirical studies defend the significant and positive relationship between QM and firms’ performance (Das et al., 2008; Sila et al., 2006).

3.2 MFs’ performance (MFP)
It is essential to align operations, such as those of supply chains, to financial metrics. The performance of MF refers to how well an MF achieves its market and financial-oriented goals. The market share performance and financial performance have been selected in this study as the dimensions to measure the performance of Gulf MFs as part of the MFP construct. Li et al. (2006) suggested that the short-term objectives of SCM are more operational related, e.g., reduce cycle time and inventory while increasing productivity. Moreover, its long-term objectives are more financially oriented, e.g., increase market share and profits. In addition, measuring MFP-based on market share and financial performance is also in line with the work of Zhang (2002), which also considered the market share performance of companies, besides the financial performance.

3.3 Relationship between SCMP and MFP
Wu et al. (2006) stated that higher levels of SCM capabilities (i.e. responsiveness, coordination and inter firm activity integration, etc.) can potentially improve a firm’s market and financial performance. Li et al. (2006) argued that the customer service management
practice has a greater direct impact on competitive advantage than on firm performance. According to Li et al. (2006), the performance of firm refers to how well a firm achieves its market-oriented as well as financial goals. On the other hand, Al-Shboul’s (2012) found that this practice has a greater impact on firm performance (market share and financial). Furthermore, the results of Kim’s (2006) study showed that the customer service practice has a positive and significant impact on operational performance, but it does not have a direct significant impact on firm performance. The outcome of Kannan and Tan’s (2010) study suggested that there is an overlap to some degree in the domains and practices of supply and QM. This outcome contradicted Al-Shboul’s (2012) study, which found that the total QM practice was practiced and implemented at a high level and there was no overlap between them. The findings of Ting’s (2004) study argued that internal LP is not practiced in his study since labor cost is low, and has no significant impact on total cost.

3.4 Relationship between high-performing manufacturing firms and best supply chain management practices

Significant research effort has been paid to identify the best SCMPs to support firms and achieve a high level of performance. However, most of these efforts have failed to investigate the effect of these practices on firms’ performance. Therefore, the concept of best practices refers to a technique, method, process or activity that is more effective at delivering a particular outcome than any other technique, method, process or activity. The best SCM practices are those that lead to improvement in performance, that is, they help low-performing firms to become a medium performer, medium performer become a high-performing firm and high-performer firms maintain their success (Koh et al., 2007; Davies and Kochhar, 2002; Ungan, 2004).

4. Research methodology

4.1 Questionnaire design

A web-survey was developed based on the SCM literature and consisted of three main parts. The first part comprised of respondent’s profile, SCM practices and GMFP. The second part asked the respondents to rank the degree of using important 23 SCMPs grouped in seven categories (SC, FwP, UoI, CF, LP, II and QM) (Table III). In particular, the SCM practices were ranked (i.e. measured) using a five-point Likert scale as follows: 1 = not used, 2 = slightly used, 3 = no change, 4 = highly used, 5 = strongly used. The use of these practices was considered between the period 2013 and 2015. However, in the third part, the respondents were asked to rank their firms’ performance based on seven market share and financial performance items (GMFP/MSP1, GMFP/MSP3, GMFP/MSP4, GMFP/FP2 and GMFP/FP5-7) previously established as important (Table III). These were ranked using a five-point Likert scale as follows: 1 = performance has strongly deteriorated, 2 = performance has slightly deteriorated, 3 = no change, 4 = performance has slightly improved, 5 = performance has strongly improved.

The survey was deployed in English to measure SCMPs elements, which included supplier collaboration, flexibility with partners, usage of internet, CF, LP, II and QM as the main dimensional construct of SCMPs. This measurement considers SCM practices within supply chain that included downstream, upstream and internal processes across the supply chain. Questions related to GMFs performance construct measures were adopted and developed based on the instrument main items previously used by Qrunfleh and Tarafdar (2013) and Li et al. (2005, 2006). These constructs were further tested and validated from previous studies using data collected from MFs. All items were measured based on a five-point Likert scale as a unit of measurement with response option ranging from 1 (strongly disagree) to 5 (strongly agree). The survey was pre-tested by four professionals.
<table>
<thead>
<tr>
<th>Item</th>
<th>Initial CITC</th>
<th>Final CITC</th>
<th>Initial α</th>
<th>Final α</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Supplier collaboration (SC) construct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC1: our firm share information on inventory levels with our suppliers</td>
<td>0.63</td>
<td>0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC2: our suppliers provide any assistance to improve the quality of our firm’s products</td>
<td>0.55</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC3*: our firm has continuous improvements programs that include our key suppliers</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC4: our suppliers have high level of flexibility and delivery speed</td>
<td>0.61</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC5: our suppliers share forecasts of customer demand with our firm</td>
<td>0.66</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Flexibility with partners (FwP) construct</td>
<td></td>
<td></td>
<td>0.71</td>
<td>0.87</td>
</tr>
<tr>
<td>SCMP/FwP1: our firm is able to deal with different nonstandard orders</td>
<td>0.74</td>
<td>0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP2: our firm is able to produce different features of products such as: options, sizes, and colors</td>
<td>0.64</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP3: our firm is able to offer special customer specifications</td>
<td>0.61</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP4*: our firm is able to offer/introduce new products for customers</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP5: our firm is able to adjust capacity (accelerate/decelerate) in production regarding to rapidly customer demand changes</td>
<td>0.68</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usage of internet (UoI) construct</td>
<td></td>
<td></td>
<td>0.74</td>
<td>0.80</td>
</tr>
<tr>
<td>SCMP/UoI1: exchange of information with our supply chain partners is done via internet</td>
<td>0.61</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI2: in our firm, most of purchasing processes (materials, components, items, etc.) and services are done via internet</td>
<td>0.66</td>
<td>0.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI3: to high extent of selling of products and services in our firm is done via internet</td>
<td>0.66</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI4*: promotion and marketing in our firm relies to high extent on internet</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Customer focus (CF) construct</td>
<td></td>
<td></td>
<td>0.75</td>
<td>0.84</td>
</tr>
<tr>
<td>SCMP/CF1: our firm is requently follow up and monitor our customers for quality/service feedback</td>
<td>0.72</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF2*: our firm is frequently tries to determine our future customer expectations</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF3: our firm is frequently measures and evaluates our customer satisfaction</td>
<td>0.67</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF4: our firm provides and facilitates any assistance for our customer</td>
<td>0.68</td>
<td>0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) Lean production (LP) construct</td>
<td></td>
<td></td>
<td>0.72</td>
<td>0.76</td>
</tr>
<tr>
<td>SCMP/LP1*: suppliers’ warehouses are located very close to our firm</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP2: time has been reduced for inspection of incoming materials/items/components</td>
<td>0.73</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP3: our firm encourages suppliers for shorter lead-times</td>
<td>0.61</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP4: our firm’s policy is looking for reduction in set-up times</td>
<td>0.66</td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(e) Internal integration (II) construct</td>
<td></td>
<td></td>
<td>0.79</td>
<td>0.89</td>
</tr>
<tr>
<td>SCMP/II1: there is high level of coordination between different departments in our firm</td>
<td>0.73</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/II2*: ability to handle unexpected challenges within different departments in our firm</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/II3*: there is an internal integration between logistics, production and marketing departments in our firm</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/II4: our firm formulates quality circles and cross-functional teams for solving problems and/or developing processes, products, and services</td>
<td>0.74</td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III. Item purification for SCM practices and Gulf manufacturing firm performance constructs and dimensions (continued)
Independent $t$-tests were carried out and Levene’s test for equality of variances was also applied to decide whether the data are equal or unequal variances version of the tests. Based on the results, an equal variances version of the tests have been used in the analysis.

4.2 Data collection, population and sampling

A quantitative data collection procedure was followed to facilitate the analysis and increase the validity and reliability of outcomes. A questionnaire survey has been designed due to its suitability to collect a large amount of data from a large number of respondents. The sampling frame of this study consisted of all medium- and large-sized MFs as listed in the ministry of industry and trade of each of the six Gulf countries considered for this study (i.e. Saudi Arabia, United Arab Emirates, Kuwait, Oman, Bahrain and Qatar). Due to different factors such as large amounts of trading agreements among these countries, similar level of maturity of their manufacturing sectors, similar working cultures, similar levels of economic development and geographical proximity, it was assumed that their SCM practices were very similar. This allowed the study of these practices to be concluded within a regional context as opposed to an individual national context. The data from these countries also helped us to gain a broader insight into the SCM practices adopted by Gulf MFs. A total of 1,421 surveys were distributed via electronic and post mails, which represent whole population of medium and large-sized Gulf MFs. From these, 144 complete and usable responses were obtained, giving an overall response rate of 10.1 percent. Demographic data, shown in Table II, depicted that the majority of the firms’ respondents are from textile, plastic/rubber and chemical firms, which constitute 57.05 percent of the total firm sample. In terms of job title, the majority of respondents were senior managers, which represents 53.47 percent. Majority of firms have 3–5 product lines, which represents 61.80 percent of the total firm sample. The total numbers of employees in the sample were between 251 and

<table>
<thead>
<tr>
<th>Item</th>
<th>Initial CITC</th>
<th>Final CITC</th>
<th>Initial $\alpha$</th>
<th>Final $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f) Quality management (QM) construct</td>
<td></td>
<td></td>
<td>0.82</td>
<td>0.85</td>
</tr>
<tr>
<td>SCMP/QM1: our firm has a salary promotion and incentives for encouraging employees’ participation in quality improvement</td>
<td>0.70</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/QM2: the defect rates of the primary products in our firm are decreasing</td>
<td>0.66</td>
<td>0.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/QM3: our firm has quality circles and cross-functional teams</td>
<td>0.76</td>
<td>0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/QM4: top management in our firm encourages and offers all resources required for employee education and training</td>
<td>0.64</td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/QM5*: our firm implements various inspections effectively and frequently</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/QM6*: our firm treats customer complaints based on quality criteria with top priority</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(g) Market share performance (MSP)</td>
<td></td>
<td></td>
<td>0.71</td>
<td>0.77</td>
</tr>
<tr>
<td>GMFP/MSP1: market share</td>
<td>0.57</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/MSP2: the growth of market share</td>
<td>0.63</td>
<td>0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/MSP3: the growth of sales</td>
<td>0.67</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial performance (FP)</td>
<td></td>
<td></td>
<td>0.76</td>
<td>0.79</td>
</tr>
<tr>
<td>GMFP/FP2: return on investment</td>
<td>0.67</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/FP3: growth in return on investment</td>
<td>0.70</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/FP4: profit margin on sales</td>
<td>0.73</td>
<td>0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/FP7: overall competitive position</td>
<td>0.70</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *Denoted items were deleted

Table III.
500 or greater and majority of firms were large-sized which represents (54.86 percent). Almost all the MFs in this study have 10–30 years of operational experience. In terms of annual sales, the majority of firms have sales between 6 and 50m dollars, which represents 69.44 percent from the total firm sample.

The data from these countries also helped us to gain a broader insight into the SCM practices adopted by Gulf MFs. The linkages between medium- and large-sized MFs and low- and high-performing firms are important in the context of best SCM practices as they play a crucial role in an economy. Furthermore, various researchers emphasized that both large and medium firms are interested and have more concerns than smaller firms in implementing SCMPs to enhance their performance. This indicates that best practices will come from both medium- and large-sized high-performing MFs (Sundram et al., 2016; Ungan, 2007). In particular, each medium- and large-sized firms in the sample met the following criteria: must have been in operation for at least ten years, and must have had 51 employees or more. In relation to the targeted respondents, the study included middle and high-level managers (e.g. CEO’s, presidents, purchasing managers, supplying managers, planning managers, logistics managers, IT managers, manufacturing managers, distribution/transportation/sales managers and operations managers). Similar to the studies of Al-Shboul et al. (2017), Andreadis et al. (2017) and Belekoukias et al. (2014), respondents in these functional positions were considered to have an adequate knowledge on SCM practices and their effect on the performance of their firms. The respondents came from eight manufacturing sectors, namely, food processing, 179; furniture, 156; pharmaceutical, 135; textile, 135; chemical, 263; tobacco and cigarettes, 67; paper and packaging, 73; and plastic/rubber, 268.

Gulf countries were selected as a desired sample as it is one of the fastest-growing regions in the world that has benefited from rising oil prices over the past two decades and introduced many facilitations to encourage foreign investors to invest in the Gulf region. It has attracted many international firms and most of them already have branches and offices in the Gulf countries and operate globally. The region is easy to access and more approachable for data collection process and there is a lack of empirical studies in the SCM field. Gulf and Western countries arguably share the biggest gap concerning their cultural business and acceptance to SCMPs, regional and global supply chains. Recognizing this fact, many American and western firms now have a sustainability and strategic partnerships with Gulf firms who help to create the “new supply chain.”

4.3 Measure refinement and validation
4.3.1 Assessment of reliability. Cronbach's $\alpha$ was used for each unidirectional scale along with the corrected item-to-total correlation (CITC) to assess the reliability of each construct, and their items, of the theoretical research framework proposed (Figure 1). An $\alpha$ score higher than 0.7 was considered acceptable for all constructs of this study (Nunally, 1978), whereas the cut-off values for $\alpha$ and CITC were between 0.60 and 0.89. Table III has displayed the reliability measures, for instance, the SCMP/SC construct initially included five-items. An initial $\alpha$ indicated that SCMP/SC3 item had $\alpha \leq 0.50$. After removing this item from any further analysis, all remaining items were analyzed and strongly loaded into their respective $\alpha$ with loadings $\geq 0.68$, as shown in Table III. Similarly, the SCMP/FwP dimension was initially represented by five-items. An initial $\alpha$ indicated that SCMP/FwP4 item had $\alpha \leq 0.50$. After removing this item, the remaining items were factor analyzed and strongly loaded into their respective $\alpha$ with loadings $\geq 0.71$. The SCMP/UoI dimension was initially represented by four-items. An initial $\alpha$ indicated that SCMP/UoI4 item had $\alpha \leq 0.50$. After removing this item, all remaining items were factor analyzed and the results are shown in Table III. It can be seen that all items loaded on their respective $\alpha$ with most of
loadings $\geq 0.73$. The same purification process was applied to the rest of the dimensions and their items. Table III denotes with a "*" all the items that were eliminated through this process.

To achieve a significant level of instrument validity, a five-point Likert scale was used in the questionnaire. The questionnaire instrument was reviewed and re-evaluated by five academics and six expert practitioners, who were asked to provide feedback in relation to the appropriateness of the instrument, meaningfulness and usefulness for the targeted respondents.

4.3.2 Assessment of validity. Factor analysis was performed on the remaining items from the reliability analysis to verify the dimensionality and reliability of each construct to ensure convergent validity (Nunally, 1978). Factor analysis was used to examine the multidimensionality of both SCM practices and GMFP. The multidimensionality between the produced factors was checked, which is a measure of sampling adequacy, was found to be 0.783. This value is greater than 0.5, so, it can be considered that the factor analysis test has proceeded correctly and that the sample was used adequately. This shows that the factor processes were correct and suitable for testing multidimensionality. The final analysis was performed after removing all items that have scored below 0.7. Therefore, the result found that all items were strongly loaded (loading $\geq 0.50$) on their associated factors, which suggested that there was a convergent validity. Discriminant validity was confirmed when the load of item was stronger on its associated factor than on others. Factor analysis was run to assess the discriminant validity. All items loaded as theorized and seven factors together explained 73.53 percent of the total variance. As a consequence, KMO of sampling adequacy (0.689) indicated that the data were adequate for factor analysis (Hatcher, 1994), the results shown in Table IV.

4.4 Data analysis
A total 144 complete and usable responses were obtained from the whole population of 1,421 medium- and large-sized GMFs. The survey asked the respondents to rank their firm's performance improvements within the last three years on the basis of seven performance indicators (Tables III and V). The sample was divided into two groups, high performers and low performers, to show how much level medium and large-sized GMFs adopt and use SCM practices (Tables V, VI and VIII). The analysis was performed in three steps. First, the GMFs were divided into two groups, namely, low- and high-performing GMFs. High-performing firms were those that achieved an average score $\geq 4$ on all three market share performance items (i.e. GMFP/MSP1, GMFP/MSP3 and GMFP/MSP4) and four financial performance items (i.e. GMFP/FP2, GMFP/FP5, GMFP/FP6 and GMFP/FP7). All these seven items represent the improvements in the combination of market share and financial performance items (GMFP/(MSP-FP)). On the other hand, the low-performing firms were considered those that achieved an average score $x < 4$ on all previous performance indicators, showing either a deterioration in their performance or at best maintaining an status quo, where $x$ represents to firm’s performance. This means that MFs that have an average score $1 \leq x < 3$ represent that their performance has strongly or slightly deteriorated, while firms that have an average score $3 \leq x < 4$ represent that there is no change in their performance. Therefore, merging the last two scales used in this study considers firms as low-performing GMFs, and those who already adopted SCM practices did not have any significant impact on their performance in the last three years. While, high-performing firms were considered those that achieved an average score $x \geq 4$ on all performance indicators, showing either slight or strong improvement in their performance.

Table V has indicated that a total of 75 Gulf MFs’ performance have strongly improved their market share performance during the last three years; while 69 firms indicated that their market share performance has stayed the same with no change, or even deteriorated in
the last three years. In total, 80 firms have strongly improved their financial performance over the last three years. A total of 79 Gulf MFs have strongly improved in combination of market share and financial performance during the last three years, hence, were categorized as high-performers. In total, 65 firms indicated that the combination of market share and financial performance have stayed the same with no change, or even deteriorated in the last three years, hence were categorized as low performers (Table V). The t-test analysis was performed to determine the differences in the implementation of SCM practices between high- and low-performer groups for each of the (in total) three categories (GMFP/MSP, GMFP/FP and the combination of GMFP/(MSP-FP)) (Table VI). Lastly, a multiple linear regression analysis was conducted to determine SAM practices that had significant influence on firm’s performance (Table VII).

5. Results of the study

5.1 Adoption of SCMPs in medium- and large-sized Gulf MFs

The t-test analysis has shown that there are many significant differences in the degree of implementation of SCM practices between low- and high-performing GMFs (Table VI). This indicates that most of the SCM practices have been implemented differently among low and high performers. The results show that the flexibility with partners, usage of internet, LP and II are the least implemented and used SCM practices (mean < 4) among the respondents in all categories of performance (GMFP/MSP, GMFP/FP and in the combination of GMFP/ (MSP-FP)). On the other hand, supplier collaboration, CF and QM are the highest SCM

Table IV.
Results of factor analysis for discriminant validity

<table>
<thead>
<tr>
<th>Item</th>
<th>SCMP/SC</th>
<th>SCMP/FwP</th>
<th>SCMP/UoI</th>
<th>SCMP/CF</th>
<th>SCMP/LP</th>
<th>SCMP/II</th>
<th>SCMP/QM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCMP/SC1</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC2</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC4</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC5</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP1</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP2</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP3</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP5</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI1</td>
<td></td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI2</td>
<td></td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI3</td>
<td></td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF1</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF3</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF4</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP1</td>
<td></td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP3</td>
<td></td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP4</td>
<td></td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/II1</td>
<td></td>
<td></td>
<td></td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/II4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/QM1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>SCMP/QM2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>SCMP/QM3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>SCMP/QM4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.79</td>
</tr>
</tbody>
</table>

Kaiser–Meyer–Olkin measure of sampling adequacy: 0.689; factor loadings of 0.4 and above are shown.

Table V.
Item SCMP/SC SCMP/FwP SCMP/UoI SCMP/CF SCMP/LP SCMP/II SCMP/QM

<table>
<thead>
<tr>
<th>Item</th>
<th>SCMP/SC</th>
<th>SCMP/FwP</th>
<th>SCMP/UoI</th>
<th>SCMP/CF</th>
<th>SCMP/LP</th>
<th>SCMP/II</th>
<th>SCMP/QM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCMP/SC1</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC2</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC4</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/SC5</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP1</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP2</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP3</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/FwP5</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI1</td>
<td></td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI2</td>
<td></td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/UoI3</td>
<td></td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF1</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF3</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/CF4</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP1</td>
<td></td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP3</td>
<td></td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/LP4</td>
<td></td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/II1</td>
<td></td>
<td></td>
<td></td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/II4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCMP/QM1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>SCMP/QM2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>SCMP/QM3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>SCMP/QM4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.79</td>
</tr>
</tbody>
</table>

Eigen value 3.74 3.45 2.89 2.86 2.78 2.67 2.45
Variance (percent) 14.88 13.64 12.85 11.85 11.29 10.74 9.78
Cumulative variance (percent) 14.88 28.45 40.34 51.56 64.22 56.74 73.53
<table>
<thead>
<tr>
<th>Coding</th>
<th>Improvements in GMFs Performance during the last three years</th>
<th>Mean</th>
<th>SD</th>
<th>Statistic</th>
<th>SE</th>
<th>Statistic</th>
<th>SE</th>
<th>Number of high performers</th>
<th>Number of low performers</th>
<th>Total n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Market share performance (MSP)</td>
<td>3.36</td>
<td>0.62</td>
<td>0.109</td>
<td>0.247</td>
<td>−0.091</td>
<td>0.401</td>
<td>0.379</td>
<td>0.368</td>
<td>75</td>
</tr>
<tr>
<td>GMFP/MSP1</td>
<td>Market share</td>
<td>3.30</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/MSP3</td>
<td>Growth of market share</td>
<td>3.42</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/MSP4</td>
<td>Growth of sales</td>
<td>3.40</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Financial performance (FP)</td>
<td>3.05</td>
<td>0.73</td>
<td>0.245</td>
<td>0.233</td>
<td>0.273</td>
<td>0.341</td>
<td>−0.212</td>
<td>0.454</td>
<td>80</td>
</tr>
<tr>
<td>GMFP/FP2</td>
<td>Return on investment</td>
<td>3.60</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/FP5</td>
<td>Growth in return on investment</td>
<td>2.70</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/FP6</td>
<td>Profit margin on sales</td>
<td>3.00</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/FP7</td>
<td>Overall competitive position</td>
<td>2.90</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/(MSP-FP)</td>
<td>Combination of market share and financial performance (MSP-FP)</td>
<td>3.20</td>
<td>0.42</td>
<td>0.440</td>
<td>0.409</td>
<td>0.454</td>
<td>0.401</td>
<td>0.166</td>
<td>0.247</td>
<td>79</td>
</tr>
<tr>
<td>GMFP/MSP</td>
<td>High-low performing GMFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High-performing GMFs</td>
<td>4.13</td>
<td>0.36</td>
<td>0.43</td>
<td></td>
<td>0.83</td>
<td></td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low-performing GMFs</td>
<td>2.36</td>
<td>0.23</td>
<td>−1.05</td>
<td></td>
<td>1.16</td>
<td></td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/FP</td>
<td>High-performing GMFs</td>
<td>4.00</td>
<td>0.37</td>
<td>1.12</td>
<td></td>
<td>0.46</td>
<td></td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low-performing GMFs</td>
<td>2.08</td>
<td>0.18</td>
<td>0.87</td>
<td></td>
<td>−0.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMFP/(MSP-FP)</td>
<td>High-performing GMFs</td>
<td>4.29</td>
<td>0.27</td>
<td>1.07</td>
<td></td>
<td>0.64</td>
<td></td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low-performing GMFs</td>
<td>2.05</td>
<td>0.16</td>
<td>1.16</td>
<td></td>
<td>0.72</td>
<td></td>
<td>65</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table V. Average values for the Gulf manufacturing firms’ performance criteria in the two groups of low- and high-performing GMFs.
### Table VI.

<table>
<thead>
<tr>
<th>Improvements in Gulf manufacturing firms’ performance in the last three years (2013–2015)</th>
<th>GMFP/MSP</th>
<th>GMFP/FP</th>
<th>GMFP/(MSP-FP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low/High-performing GMFs</td>
<td>Sample size (n)</td>
<td>Mean</td>
<td>Skewness/Kurtosis</td>
</tr>
<tr>
<td>GMFP/MSP</td>
<td>High-performing GMFs</td>
<td>75</td>
<td>4.401</td>
</tr>
<tr>
<td>Low-performing GMFs</td>
<td>69</td>
<td>2.310</td>
<td>2.291</td>
</tr>
<tr>
<td>GMFP/FP</td>
<td>High-performing GMFs</td>
<td>80</td>
<td>4.337</td>
</tr>
<tr>
<td>Low-performing GMFs</td>
<td>64</td>
<td>2.401</td>
<td>2.315</td>
</tr>
<tr>
<td>GMFP/(MSP-FP)</td>
<td>High-performing GMFs</td>
<td>79</td>
<td>4.361</td>
</tr>
<tr>
<td>Low-performing GMFs</td>
<td>65</td>
<td>2.311</td>
<td>2.314</td>
</tr>
</tbody>
</table>

**Notes:** 1 = no use, 5 = high use. *Coefficients are statistically Significant (two-tailed) at \( p \leq 0.05 \)

Supply chain management practices in the last three years (2013–2015)

<table>
<thead>
<tr>
<th>Improvements in Gulf manufacturing firms’ performance in the last three years (2013–2015)</th>
<th>SCMP/SC</th>
<th>SCMP/QM</th>
<th>SCMP/FwP</th>
<th>SCMP/UoI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low/High-performing GMFs</td>
<td>Sample size (n)</td>
<td>Mean</td>
<td>Skewness/Kurtosis</td>
<td>Mean</td>
</tr>
<tr>
<td>GMFP/MSP</td>
<td>High-performing GMFs</td>
<td>75</td>
<td>4.230</td>
<td>6.11*</td>
</tr>
<tr>
<td>Low-performing GMFs</td>
<td>69</td>
<td>2.391</td>
<td>2.521</td>
<td>2.401</td>
</tr>
<tr>
<td>GMFP/FP</td>
<td>High-performing GMFs</td>
<td>80</td>
<td>4.221</td>
<td>9.67*</td>
</tr>
<tr>
<td>Low-performing GMFs</td>
<td>64</td>
<td>2.561</td>
<td>2.601</td>
<td>2.591</td>
</tr>
<tr>
<td>GMFP/(MSP-FP)</td>
<td>High-performing GMFs</td>
<td>79</td>
<td>4.433</td>
<td>7.57*</td>
</tr>
<tr>
<td>Low-performing GMFs</td>
<td>65</td>
<td>2.310</td>
<td>2.521</td>
<td>2.601</td>
</tr>
</tbody>
</table>
practices implemented and used among the respondents (mean ≥ 4) in all performance categories of Gulf MF’s performance.

The difference between low- and high-performing GMFs is not significant in any of the performance categories. High performers in all categories implement SCM practices related to supplier collaboration, CF and QM to a significantly (p ≤ 0.05) higher degree than low performers. This suggests that unlike low performers, high performers adopt, apply and gain more from these SCM practices. Additionally, Gulf MFs which are high performers seem to be more consistent with the use of the practices over time. The exception to this is the difference in the use of supplier collaboration (for MSP-FP), which was found not to be significant (p > 0.05). The rest four SCM practices are also used more by high performers in the single performance categories, although the significance is lower than p ≤ 0.05.

The Pearson Correlation Coefficient (r) was used to measure the magnitude and direction of the relationship between GMFP/MSP and GMFP/FP. The results, as shown in Table VI, indicate that the correlation coefficient (r) between GMFP/MSP and GMFP/FP is 0.716 and has a strong positive correlation (r (144) = 0.003, p = 0.05, two-tailed). The researcher considers these ranges of correlations (r) for hypotheses analyses in this study as follows: if r > 0.7, correlation is considered strong; if 0.3 ≤ r ≤ 0.7, correlation is considered moderate; and if r < 0.3, correlation is considered weak. Also, the same ranges apply to negative values.

5.2 The performance effects of SCMPs
A multiple linear regression analysis was performed to reveal the performance impact of the different SCM practices and the results have been illustrated in Table VIII.

5.2.1 Market share performance of Gulf MFs. The SCMPs such as supplier collaboration, usage of internet, CF and QM practices are positively related to better market share performance in high-performing GMFs. These practices have statistically significant differences (p ≤ 0.05) with GMFP/MSP dimension. The proportion of variance explains 78.8 percent (R² = 0.788), whereas the F-value is 2.98. This means that there is a significant positive impact and strong (r > 0.7) relationship between these practices and GMFP/MSP (Cronk, 2004). On the other hand, flexibility with partners, LP and II did not show a significant effect (p > 0.05) on the market share performance of high-performing GMFs. Whereas, it is found that only usage of internet practice has positive and significant (p ≤ 0.05) effect on the market share performance of low-performing GMFs. The proportion of variance explains 35.8 percent (R² = 0.358), whereas the F-value is 18.68. This means that there is a positive impact and an existence moderate (0.3 ≤ r ≤ 0.7) relation between this practice and GMFP/MSP dimension.

<table>
<thead>
<tr>
<th>Coding</th>
<th>GMFP/MSP items</th>
<th>GMFP/FP items</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMFP/MSP items</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson correlation</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>144</td>
<td>144</td>
</tr>
<tr>
<td>GMFP/FP items</td>
<td>0.716*</td>
<td>1</td>
</tr>
<tr>
<td>Pearson correlation</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Sig. (two-tailed)</td>
<td></td>
<td>144</td>
</tr>
<tr>
<td>n</td>
<td>144</td>
<td></td>
</tr>
<tr>
<td>R² = 0.513</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *Correlation is statistically significant at 0.05 level

Table VII. Pearson correlation analysis for GMFP/MSP and GMFP/FP items relationship
## Table VIII.

Multiple linear regression analysis: Supply chain management practices in the last three years (2013–2015)

<table>
<thead>
<tr>
<th>Improvements in Gulf manufacturing firms’ performance in the last three years (2013–2015)</th>
<th>Sample size (n)</th>
<th>(Unstandardized coefficients)/standardized coefficients; t, Sig.</th>
<th>Supply chain management practices in the last three years (2013–2015)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-performing GMFs/MSP</td>
<td>75</td>
<td>(B/SE)</td>
<td>SCMP/SC</td>
</tr>
<tr>
<td>$R^2 = 0.787; F$-value $= 298.1$</td>
<td></td>
<td>$0.310$</td>
<td>$0.106$</td>
</tr>
<tr>
<td>Adjusted $R^2 = 0.697, P = 0.021$</td>
<td></td>
<td>$0.107$</td>
<td>$0.053$</td>
</tr>
<tr>
<td>Low-performing GMFs/MSP</td>
<td>69</td>
<td>(B/SE)</td>
<td>SCMP/SC</td>
</tr>
<tr>
<td>$R^2 = 0.358; F$-value $= 18.68$</td>
<td></td>
<td>$0.147$</td>
<td>$0.043$</td>
</tr>
<tr>
<td>Adjusted $R^2 = 0.286, P = 0.043$</td>
<td></td>
<td>$0.148$</td>
<td>$0.025$</td>
</tr>
<tr>
<td>High-performing GMFs/FP</td>
<td>80</td>
<td>(B/SE)</td>
<td>SCMP/SC</td>
</tr>
<tr>
<td>$R^2 = 0.601; F$-value $= 26.7$</td>
<td></td>
<td>$0.166$</td>
<td>$0.025$</td>
</tr>
<tr>
<td>Adjusted $R^2 = 0.512, P = 0.016$</td>
<td></td>
<td>$0.065$</td>
<td>$0.075$</td>
</tr>
<tr>
<td>Low-performing GMFs/FP</td>
<td>64</td>
<td>(B/SE)</td>
<td>SCMP/SC</td>
</tr>
<tr>
<td>$R^2 = 0.368; F$-value $= 9.23$</td>
<td></td>
<td>$0.015$</td>
<td>$0.087$</td>
</tr>
<tr>
<td>Adjusted $R^2 = 0.291, P = 0.033$</td>
<td></td>
<td>$0.022$</td>
<td>$0.067$</td>
</tr>
<tr>
<td>High-performing GMFs/(MSP-FP)</td>
<td>79</td>
<td>(B/SE)</td>
<td>SCMP/SC</td>
</tr>
<tr>
<td>$R^2 = 0.769; F$-value $= 278.9$</td>
<td></td>
<td>$0.144$</td>
<td>$0.091$</td>
</tr>
<tr>
<td>Adjusted $R^2 = 0.689, P = 0.042$</td>
<td></td>
<td>$0.077$</td>
<td>$0.112$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0.112$</td>
<td>$0.301$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.867$</td>
<td>$0.811$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0.021$</td>
<td>$0.031$</td>
</tr>
</tbody>
</table>

(continued)
Supply chain management practices in the last three years (2013–2015) in Gulf manufacturing firms’ performance in the last three years (2013–2015) Sample size (n) (Unstandardized coefficients/standardized coefficients; t, Sig.)

<table>
<thead>
<tr>
<th>Low-performing GMFs/(MSP-FP)</th>
<th>65</th>
<th>( B ) ( SE )</th>
<th>( \beta )</th>
<th>( t )</th>
<th>Sig.</th>
<th>SCMP/SC</th>
<th>SCMP/QM</th>
<th>SCMP/FwP</th>
<th>SCMP-UoI</th>
<th>SCMP/CF</th>
<th>SCMP/LP</th>
<th>SCMP/II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R^2 = 0.319 ) ( F )-value = 78.4</td>
<td>( 0.034 ) ( 0.107 )</td>
<td>( 0.089 ) ( 0.076 )</td>
<td>( -0.008 ) ( 0.152 )</td>
<td>( -0.071 ) ( 0.137 )</td>
<td>( -0.376 ) ( 1.408 )</td>
<td>0.450</td>
<td>0.129</td>
<td>0.949</td>
<td>0.396</td>
<td>0.665</td>
<td>0.248</td>
<td>0.264</td>
</tr>
</tbody>
</table>
| Adjusted \( R^2 = 0.257 \), \( P = 0.029 \) | \( -0.011 \) \( 0.048 \) | \( 0.112 \) \( 0.119 \) | \( 0.099 \) \( 0.400 \) | \( 0.199 \) \( 0.996 \) | \( 0.093 \) \( 0.665 \) | 0.129 | 0.248 | 0.665 | 0.264 | \( 0.147 \) \( 0.075 \) | \( 0.148 \) \( 0.120 \) | \( 0.077 \) \( 0.277 \) |}

\[ \text{Note: *Coefficients are statistically significant at } p \leq 0.05 \]
The supplier collaboration practice contributes to rely on few high-quality suppliers and provides assistance in sharing information in inventory levels to improve the quality of suppliers’ products with reliable and speed of delivery. Usage of internet has a significant influence on market share as it facilitates the exchange of information between the firm and its suppliers and customers and well. If the exchanged information is timely, accurate, complete, adequate and reliable, it will contribute to increase the market share of Gulf MF performers. The CP practice contributes in increasing the follow up and monitoring the customers’ quality/service feedback, evaluating customers’ satisfaction and providing assistance for their customers. This will build a good reputation for Gulf MFs as well as increase their sales in local, regional and global markets. The QM practice appears to have a positive influence on enhancing market share.

5.2.2 Financial performance of Gulf MFs. The implementation of supplier collaboration, CF, LP and QM practices contribute to better financial performance in high-performing GMFs. This suggests a positive significant \( (p \leq 0.05) \) relationship between these practices and financial performance except lean practice, which has a negative effect. The proportion of variance explained is 60.1 percent \( (R^2 = 0.601) \), whereas the \( F \)-value is 26.7. This means that there is a significant impact and an existence of moderate relationship between these practices and GMFP/MSP dimension. The LP practice showed a significant negative effect on the financial performance of high-performing GMFs. This surprising result may be due to the difference in targeted respondents. Moreover, the majority of targeted respondents generally ignored the concept of lean practice. Even when it is applied, it is done partially and lacks a true spirit and totality, which is supported by the study conducted by Abu-Alrejal (2007).

The flexibility with partners, usage of internet and II did not show any effect on this measure \( (p > 0.05) \). LP and II practices have significant \( (p \leq 0.05) \) positive effect on financial performance in low-performing GMFs. The proportion of variance explains 36.8 percent \( (R^2 = 0.368) \), whereas the \( F \)-value is 92.3. This means that there is a significant positive impact and an existence of moderate relationship between these practices and GMFP/FP dimension. While, all other practices do not have an effect on financial performance in low performers.

The supplier collaboration practice has positive effect on financial performance. Additionally, CF has a positive relationship and influence on financial performance, which plays an important role to increase customers’ satisfaction, follow up and monitor firm’s services and customers’ claims. On the other hand, vast empirical and theoretical evidence has shown the positive effect of LP on various performance dimensions of organizations (Belekoukias et al., 2014), including financial. However, the results obtained from this study contradict this previous evidence. Belekoukias et al. (2014) indicated that the incorrect application of some lean tools, e.g. value stream mapping (VSM), may result in negative effects on the performance of firms. The QM practice has a positive relationship and significantly contributes to better financial performance of high-performing GMFs through encouraging employees to be more involved in QM and improvement activities.

5.2.3 Combination of market share and financial performance. The results as shown in Table VIII have indicated that the Pearson coefficient \( (r) \) is 0.616 for GMFP/MSP and GMFP/FP, while the correlation has probability \( (p) \) 0.003 for two-tailed test. Hence, a moderate positive and statistically significant correlation was found. The proportion of variance explains 61.3 percent \( (R^2 = 0.613) \). A multiple linear regression analysis indicated that there is a significant \( (p \leq 0.05) \) positive relationship between the adoption and implementing of supplier collaboration, CF, QM, and improvement of high-performing GMFs. These practices have statistically significant differences \( (p \leq 0.05) \) with a combination of GMFP/(MSP-FP) dimension, whereas all other practices did not have any
significant effect. This means that there is a significant positive impact and an existence of strong ($r > 0.7$) relationship between these practices and GMFP/(MSP-FP) dimension (Cronk, 2004).

On the other hand, the analysis revealed that the II practice lead to improved market share and financial performance in low-performing GMFs. The proportion of variance explains only 31.9 percent ($R^2 = 0.319$), whereas the $F$-value is 78.4. This means that there is a significant positive impact and an existence of moderate relationship ($0.3 \leq r \leq 0.7$) between this practice and a combination of GMFP/(MSP-FP) dimension in low performers. This means that there is a significant positive impact and an existence of moderate relationship between this practice and this dimension, while all other practices did not have any significant effect. Apparently, a high degree of improvement on the two categories of performance (market share and financial performance) is associated with the implementation of SCM practices at QM, CF and supplier collaboration, respectively. Therefore, these practices are considered the best SCM practices for both medium- and large-sized firms of high-performing GMFs.

SCM practices directed toward improving supplier collaboration, CF and QM have a significant positive effect on all performance combinations involving high-performing GMFs. Whereas, flexibility with partners, usage of internet, LP and II practices have no effect on improving performance combination GMFP/(MSP-FP) in both medium and large-sized firms in low and high-performing GMFs. Figure 2 has illustrated the best SCM practices in Gulf MFs.

6. Discussion

6.1 QM, CF and supplier collaboration practices
There are positive significant relationships ($p \leq 0.05$) between the implementation of QM, CF and supplier collaboration. These practices lead to significant improvements in the combination of market share and financial performances (GMFP/(MSP-FP)). In contrast, these practices do not show any significant synergetic effect in low-performing GMFs. These findings suggest that QM, CF and supplier collaboration practices should be qualified as best practices, that is, they support high-performing GMFs achieve significant improvements in most performance areas and combinations. Furthermore, these practices seem to reinforce and complement each other. The role of QM as a best practice in the sense of contributing to performance improvement in high-performing GMFs.

None of the other SCM practices investigated in this study appears to produce any significant impact on high-performing GMFs. These practices include; flexibility with partners, usage of internet and II. It has been found that LP practice has a negative significant effect on financial performance in high-performing GMFs, while it has positive
significant effect in low-performing GMFs. There are no significant effects \((p > 0.05)\) of QM, CF and supplier collaboration practices on market share, financial and combination of market share and financial performances in low-performing GMFs.

6.2 Flexibility with partners, usage of internet, II and LP practices

Flexibility with partners practice does not have a significant impact \((p > 0.05)\) on improvements of market share, financial and combination of market share and financial performances in both low- and high-performing GMFs. Usage of internet practice has positive significant \((p \leq 0.05)\) impact on improvement of market share performance in both low and high-performing GMFs. While it does not have any significant \((p > 0.05)\) effect on any of the other performance areas. LP practice has mixed effects on GMFs performance and has negative significant \((p \leq 0.05)\) effect on improvement of market share performance in high-performing GMFs. While it has positive significant \((p \leq 0.05)\) effect on improvement of financial performance in low-performing GMFs. II practice has only one positive significant \((p \leq 0.05)\) effect on financial performance in low-performing GMFs; whereas does not have any significant \((p > 0.05)\) effect on any other performance areas. Therefore, the conclusion is that these practices does not appear as best practices from MFs’ performance perspective.

According to Al-Najem et al. (2013), the term “lean” is still a relatively unknown concept in Arab countries. This lower level of LP awareness among Gulf countries when compared, for example, to western countries may have contributed for this SCM practice to have lower level of implementation than other SCM practices. Although, the use of internet is widely spread among MFs, there is still a gap in the use of this technology between developed countries, with developed nations showing a much higher rate (Zaied et al., 2007). This low use of internet may also impede a more effective II among the departments of the Gulf firms and the improvements of their flexibility with partners. This pattern indicates that high-performing GMFs have problems gaining benefit of practices directed toward flexibility with partners, usage of internet, II and lean practices. This may be due to the fact that the concepts are not activated and/or rather new, especially in the manufacturing industry.

The differences in implementing and adopting SCMPs as addressed in this study and exploring the best practices between low- and high-performing GMFs in both medium- and large-sized firms may be due to the following:

- majority of the targeted respondents were from large-sized firms (251 employees or greater);
- large-sized firms play a crucial role in economy and many research emphasizes that large firms are interested and have more concerns in implementing a proper SCMPs for enhancing their performance than other sizes (Sundram et al., 2016);
- many large firms have growing number of franchises, trading agreements for the long-run, mergers, alliances and strategic partnerships with other international foreign investors to invest in the Gulf countries and operate globally;
- the influence of multinational firms in large-sized firms than other sizes, so, many local firms have also engaged and involved in implementing and creating such best SCMPs in this study; and
- recognizing the fact that many American and western firms now have a sustainability strategy that involves partnering with Gulf’s firms who help to create the new supply chain and best SCMPs.

The results have been analyzed in Table IX, which clarifies that supplier collaboration, CF and QM have a variety of performance effects and reinforce each other. Therefore, these four SCMPs investigated seem to represent best practices in high-performing GMFs.
The status of usage of internet, which may had been best practice in the past, has lost its status. Usage of internet appears to have a positive significant impact on market share performance in both low- and high-performing GMFs, whereas it has no significant effect on financial and the combination of market share and financial performances. The status of LP is less straightforward, which produced mixed results. Furthermore, LP has negative impact on improvement of financial performance in high-performing GMFs. While it has no significant impact on market share and the combination of market share and financial performances. The improvement in LP as operationalized in this study is not a best practice, but may develop into best SCMPs. The direct effect of this practice on GMFs performance indicators was limited, while, conversely, a firms’ QM, supplier collaboration and CF practices will greatly depend on its manufacturing performance.

Therefore, the study concluded that usage of internet and LP practices are not currently, but may be developed into, best practices in future. The other two practices (flexibility with partners and II) do not have any significant effect in high-performing GMFs and should therefore not be considered as best practices. The study has focused on a limited set of industrial sectors (eight), representing a variety of GMFs in terms of size, process and type as it affects the SCMPs and the performance of Gulf MF.

### 7. Conclusion

The study aimed to investigate the current SCMPs that are commonly implemented by high-performing GMFs in both medium- and large-sized firms. The study has also determined the best practices, which are having the most significant effects on the performance of these firms in a sample of 144 MFs. It is found that high and low Gulf MF performers differ in terms of usage and implementation “width” and “depth” of SCM practices. The study has presented an improvement in best practice research in terms of its starting point that is the high-performing GMFs are the ones that (must) have the best SCMPs. The results in this study seem valid for the seven SCMPs investigated in high-performing GMFs. The results have shown that supplier collaboration, CF and QM are very strong configuration and currently represent best practices. Usage of internet is an emerging practice, but may develop significant positive effects on market share performance in the overall Gulf MFs’ performance. LP practice, also did not currently seem best practice too; in contrast, it has a significant negative effect on financial performance; whereas, no significant effects at all in the other performance combinations. Surprisingly many SCMPs, notably flexibility with partners and II, do not have a significant impact on Gulf MFs’ performance, either negatively and/or positively. Therefore, these practices currently are not considered best SCM practices in Gulf MFs.

---

<table>
<thead>
<tr>
<th>Supply chain management practices (SCMPs)</th>
<th>Best practice</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplier collaboration (SC)</td>
<td>Yes</td>
<td>Strong manufacturing performance effects</td>
</tr>
<tr>
<td>Flexibility with partners (FwP)</td>
<td>No</td>
<td>No manufacturing performance effects</td>
</tr>
<tr>
<td>Usage of internet (UoI)</td>
<td>No longer</td>
<td>Has strongly positive significant effect on market share performance; no significant effects from other performance combinations</td>
</tr>
<tr>
<td>Customer focus (CF)</td>
<td>Yes</td>
<td>Strong manufacturing performance effects</td>
</tr>
<tr>
<td>Lean production (LP)</td>
<td>Possibly</td>
<td>Has strongly negative significant effect on financial performance; no significant effects from other performance combinations</td>
</tr>
<tr>
<td>Internal integration (II)</td>
<td>No</td>
<td>No manufacturing performance effects</td>
</tr>
<tr>
<td>Quality management (QM)</td>
<td>Yes</td>
<td>Strong manufacturing performance effects</td>
</tr>
</tbody>
</table>

Table IX. SCM practices (seven categories) and best practices (three)
7.1 Research limitations, implications, and further research

The analysis suffers from three weaknesses, each reducing the validity of the results. First, the study cannot exclude the possibility that there are additional best practices that are also used by high-performing GMFs in medium and large-sized firms. Second, this study does not allow for an estimation of the potential of emerging some SCMPs. Finally, it is not clear whether or to what extent the results also hold for Gulf non-MFs (i.e., service sector). Nevertheless, since the results of this study contradict the experience and results of other researches, a suggestion to perform further studies in relation to this aspect is recommended as part of a future research agenda.

Majority of the Gulf MFs recognized and emphasized on the importance of using and implementing SCMPs, but unfortunately, some of them still do not know exactly which practices should be implemented effectively. Therefore, the researcher tested a list of pre-defined SCM practices but cannot exclude the possibility that there are additional practices explaining the best Gulf MFs’ performance. Moreover, the interviews could have provided further insights on the best SCM practices used by GMFs. Therefore, future research studies should follow a mixed methods approach (questionnaire survey and interviews) for data collection.

Third, the use of single respondent from each organization may not be enough to generate accurate data about the SCMPs in Gulf MFs’ and may lead to some measurement and/or results inaccuracy. Therefore, the future research studies should involve more respondents from each sample/targeted firm for data collection. In addition, there are some limitations linked to the sample size. A larger sample size will give a clearer picture and more accurate data for generalization of the results about SCMPs that are already adopted by low and high-performing GMFs and the best practices that were identified by high-performing GMFs. From practical implication viewpoint the managers and practitioners should focus on some areas that need to be improved to overcome the weaknesses of SCMPs such as flexibility with partners, usage of internet, LP, and II to enhance the firms’ performance.

Future research can expand the domain of SCMPs by considering additional dimensions such as geographical proximity, JIT, outsourcing, external integration, product innovation, E-procurement, order planning, bullwhip, cycle time, inventory management, production level, power/dependence, lead-time management and others, which have not been discussed in this study. In addition, future research can also focus on splitting the population of the study into sub-sectors based on industry type (metal, food processing, electrical, etc.) and size (medium and large) of the GMFs. It can then examine the best SCMPs in each sub-sector alone and conduct a comparative study between sub-sectors their effects on manufacturing performance.

References


Further reading


Fraser, J. (2006), “Metrics that matter: uncovering KPIs that justify operational improvements”, research project carried out by Manufacturing Enterprise Solutions Association and presented at Plant2Enterprise Conference in Orlando, FL.


Further reading


Fraser, J. (2006), “Metrics that matter: uncovering KPIs that justify operational improvements”, research project carried out by Manufacturing Enterprise Solutions Association and presented at Plant2Enterprise Conference in Orlando, FL.


**About the authors**

Dr Moh’d Anwer AL-Shboul has been Assistant Professor in the Business Administration Department at Princess Sumaya University for Technology (PSUT), Jordan, since 2018, and was previously working in the Logistics Sciences Department at German-Jordanian University (GJU), Jordan, from 2013 to 2018. He has published few articles in leading international Journals in the areas of E-Government and its implementation in Jordan. Moh’d AL-Shboul received the PhD Degree from Bradford University, UK (2012); MBA Degree from Yarmouk University, Jordan (2002); additionally, AL-Shboul has double major (BSc) in the areas of Electrical Engineering/Telecommunication from JUST (1992) and Business Administration from Yarmouk University (2000), Jordan. Dr AL-Shboul has long academic and professional experience in the area of teaching as Instructor in Business and Electrical Engineering; had worked at Jordan Customs for eight years and more than ten years in Electrical Engineering in several Jordanian companies. His current research interests include lean and agile supply chain, supply chain management, Logistics, supply chain responsiveness and operations management.

Dr Jose Arturo Garza-Reyes is Professor of Operations Management and Head of the Centre for Supply Chain Improvement at the College of Business, Law and Social Sciences, University of Derby, UK. He is actively involved in industrial projects where he combines his knowledge, expertise and industrial experience in operations management to help organizations achieve excellence in their internal functions and supply chains. He has also led and managed international research projects funded by the British Academy, British Council and Mexico’s National Council of Science and Technology (CONACYT). As a leading academic, he has published over 100 articles in leading scientific journals, international conferences and four books in the areas of operations management and innovation, manufacturing performance measurement and quality management systems. Areas of expertise and interest for Professor Garza-Reyes include general aspects of operations and manufacturing management, business excellence, quality improvement and performance measurement. He is Chartered Engineer (CEng), a certified Six Sigma-Green Belt and has over eight years of industrial experience working as Production Manager, Production Engineer and Operations Manager for several international and local companies in both the UK and Mexico. He is also fellow member of the Higher Education Academy (FHEA) and a member of the Institution of Engineering Technology (MIET). Dr Jose Arturo Garza-Reyes is the corresponding author and can be contacted at: J.Reyes@derby.ac.uk

Dr Vikas Kumar is Associate Professor in Enterprise Operations Management at Bristol Business School, University of West of England (UWE), UK. Prior to joining UWE, he was working as Lecturer in Management at Dublin City University Business School in Ireland. He holds the PhD Degree in Management Studies from Exeter Business School, UK, and the Bachelor of Technology (first class distinction) Degree in Metallurgy and Material Science engineering from NITFF, India. He has published more than 100 articles in leading International journals and International conferences including the *International Journal of Production Research*, *International Journal of Production Economics*, *Expert System with Applications, TQM & Business Excellence, Computers & Industrial Engineering* and *Production Planning & Control*. He serves on the editorial board of six international journals and is co-founder and Editor of the *International Journal of Supply Chain and Operations Resilience* (Inderscience). His current research interests include sustainability, supply chain management, short food supply chains, process modeling and service operations management.

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: permissions@emeraldinsight.com
Interpersonal antecedents to employee upward voice: mediating role of psychological safety

Sriji Edakkat Subhakaran and Lata Dyaram

Department of Management Studies, Indian Institute of Technology Madras, Chennai, India

Abstract

Purpose – Despite the increasing prominence of employee voice in organizational innovation and productivity, employees continue to struggle to influence matters that affect them at work. The purpose of this paper is to model work group context and manager behavior as the predictors of employee upward voice. Further, a mediating role of employee psychological safety is examined in this link.

Design/methodology/approach – With data from 575 employees representing various technology firms in India, the authors test the hypothesized relationships using covariance-based structural equation modeling.

Findings – Results indicate coworkers upward voice and manager pro-voice behavior to significantly impact employee upward voice with a mediating impact of psychological safety. This implies that perceived psychological safety plays a significant role in explaining the impact, coworkers and manager behavior would have on regulating employee upward voice.

Originality/value – This study contributes to the employee voice literature from an Indian context, where upward communication is culturally discouraged.

Keywords Employees, Managers

Paper type Research paper

1. Introduction

Employees continually face situations at work where they often exercise their discretion on whether to speak up (voice) or not in giving out their honest opinion and views. Getting employees to voice, it has long been recognized to facilitate the innovation and early detection of job-related problems (Morrison, 2014; Nemeth, 1997), whereas withholding voice (silence) (Morrison and Milliken, 2000) is cited as backing most organizational fiascos (Greenberg and Edwards, 2009; Morrison, 2011). Organizations are often found not enabling voice and lagging in providing trusting environment to speak up. Employees who speak up often run the risk of being perceived as critical and self-promoting which can affect interpersonal relationships at work and trigger employee isolation/retaliation. Unlike other forms of contextual job performance enablers such as positive organizational culture and cooperative behavior, voice is unique in that it is inherently risky for those engaging in it (LePine and Van Dyne, 1998, 2001). Voice is usually perceived as involving implicit or explicit criticism of the status quo and draws attention to the problems at workplace eliciting various forms of retributions from the target.

While employees are often attuned to the cues readily available from their immediate work setup, it is managers and coworkers who become the key informants shaping their ideas. A ready interpretation of things based on what employees see and hear from them, often vicariously guide their discretion to speak up in a confronting situation. Further, speaking up gets difficult as the targets of upward voice (managers) have the power to evaluate, resulting in holding back on everything from good ideas to great questions.

This study examines coworkers and managers as impacting employee upward voice in a non-western context. The study model as shown in Figure 1 suggests coworkers upward voice and manager pro-voice behavior to enhance employee upward voice with a mediating role of psychological safety.
1.1 Why examine employee upward voice in India?
Extant voice literature is predominantly western. It is argued that most Indian management literature often overdraws and adopts theories and frameworks from west tending toward global homogenization and convergence of management research (Khatri et al., 2012). Although some of these concepts have some relevance in the Indian context, India is known to differ for its socio-cultural and institutional diversity from that of west (Sparrow and Budhwar, 1997). Indians culturally approve and unresent high-power distance. There is often a perceived and felt comfort with personal ties established with the higher ups and authority as large power distance plus weak uncertainty avoidance leads to viewing organization as a family (Hofstede, 1985). Societal approval and significance to relationships, dependence proneness and sneh-shraddha (superior’s affection for subordinates and a subordinate’s deference to the superior) encourage many in India to maintain harmony and amicable relationships in personal and professional lives (Sinha, 2014). Hence, dissent and conflict with those in authority is perceived as risky and culturally discouraged.

There is a need to develop management research with indigenous data from Indian firms to draw employee views on their voice experiences. As ignoring the research context can be misleading (Schuler, et al., 2002), we examine how the perception of favorable interpersonal context at work can enhance psychological safety against potential negative implications for employee upward communication.

2. Theory and hypotheses
Systematic studies on employee voice date back to seminal work by Hirschman (1970). Exit Voice Loyalty Neglect (EVLN) framework alludes to a wide range of voice behavior which help to moderate employee dissatisfaction (Farrell, 1983; Hirschman, 1970). Subsequent work treated voice as employee input and behavior; discounting other facilitating mechanisms enabling voice (Burris et al., 2008; LePine and Van Dyne, 1998; Ma, 2016; Morrison, 2011; Van Dyne and LePine, 1998). Employee voicing work-related concerns within the organization is seen as distinct from communication avenues to voice.

We see upward voice behavior as “employee expression of opinions, concerns or ideas about work-related issues to their manager” (Tangirala and Ramanujam, 2012). We refer to this upward voice as any communication directed to someone higher in the organizational

![Proposed model of employee upward voice](image-url)
hierarchy with the perceived power or authority to take action on the problem or suggestion (reporting manager in the present study). When voice involves challenging the status quo or bringing up tough issues for discussion, it often elicits negative outcomes such as being labeled negatively, damaged relationships, retaliation/punishment, including fear of losing the job, not getting promoted to the next level and other unfavorable reactions from organizational members (Milliken et al., 2003; Morrison, 2014; Premeaux and Bedeian, 2003). As a result, voicing entails substantial risk for employees. Most jobs do not prescribe voice as in-role requirement (Van Dyne and LePine, 1998). This makes employees to use their discretion to speak up or not.

2.1 Co-worker’s upward voice

Research on group dynamics highlights social information processing, where individuals adapt their attitudes, behavior and beliefs to that of their social context (Liu et al., 2015; Salancik and Pfeffer 1978; Xu et al., 2017). Likewise, Bandura (1986) notes how vicarious learning takes place in a variety of social contexts. In other words, people learn acceptable and normative behavior from observing others. Often, individual behavior is a function of their attitude and favoring norms (Ajzen, 1991). Group norms govern work group members’ behavior. Observation and interaction with others at work aids in employees getting a ready interpretation of things in terms of what is acceptable and normative. This triggers employees to model coworkers’ behavior which appears to be normative within their group context. In a voice context, employees often turn to their immediate interpersonal network (i.e. coworkers and managers) to check how welcoming/unfavorable it is to freely express themselves at work (Milliken et al., 2003). When others voice upwards, it may indicate that upward voicing is acceptable; as what others do is a statement of descriptive norm. An increase in the frequency of upward voice by coworkers is more likely to make individual members to initiate and project a competent image of themselves to the managers. It is also likely that some of such coworkers become a ready reference in a voice context. Further, peers upward voice might help individual members to view dyadic context as favorable and manager as approachable. This prompts them to speak up, facilitating positive attributions about upward communication. Hence, group members’ upward voice may set a behavioral standard for individual members to emulate. Therefore, we hypothesize:

H1. Co-workers’ upward voice positively relates to employee upward voice.

2.2 Manager behavior

Most workforces across the world work in hierarchical organizations and with their agents (supervisors/managers) to seek economic support and well-being. Hence, fear of displeasing and offending those in authority is widespread and more so in cultures with high-power distance. While there are myriad situations in which employees may feel challenged in dealings with those above them, a common way to get in trouble seems to simply speak up in ways perceived as challenging the authority, status quo or critical of manager’s valued project/programs. Given this, it is only natural to anticipate employees to withhold and clam up if the cues seem unfavorable. Often managers have substantial influence on subordinate behavior as they are entrusted with the integrative function of overseeing subordinate’s key concerns on work assignment, compensation and progression (Detert and Trevino, 2010). Manager characteristics and varied styles of conduct have a significant impact on the subordinate’s behavior and development (Avey et al., 2012; Walumbwa et al., 2012).

Further, most employees consider managers as representatives of their organization’s management. Their perception of extent to which their manager values their contribution and cares about their well-being impacts they meeting their socio-emotional needs and benefits of their work effort (organization support theory, Eisenberger, et al., 1986).
Higher the perception of managerial/organizational support, greater would be the employee reciprocity in terms of desired work behavior. Behavioral outcomes of such perceived managerial support are noted to increase not only in-role, but also extra-role performance (voice in this case) (Kurtessis et al., 2017).

Specific to voice, positive affect and consideration behavior by managers were found to encourage subordinate initiation and voice (Liu et al., 2010, 2017). This indicates that supportive managerial behavior helps to build a sense of comfort and approachability among subordinates for potential risk taking. Conversely, manager disapproval of subordinate feedback, indifference, egoism and autocratic leadership styles were found to significantly discount employee voice (Burris et al., 2008; de Vries et al., 2012; Roberto, 2002). Likewise, de Vries et al. (2012) found subordinate’s perception of speaking up as pseudo-voice to moderate their voice behavior. Managers are inherently important to the improvement-oriented voice processes as they are often the targets of voice. Hence, with manager’s latent cues that they are open, and willing to act on subordinate voice, it is reasonable to anticipate subordinate’s motivation to voice to increase; conversely, where subordinates perceive the manager’s behavior as unsafe for initiation, they may less likely to voice.

Given this, we believe that manager’s pro-voice behavior – favoring subordinate voice in the team – will significantly impact subordinates’ propensity to voice. Most managers provide voice favoring cues by offering their time, providing input opportunities, soliciting suggestions, actively consulting subordinates and seeking clarifications. Such cues help subordinates to view things as fair, inclusive, respectful, allowing a sense of decision control. This may assist in being less fearful of the negative consequences for speaking up. Hence, manager’s voice favorable behavior is more likely to enhance perceived control and influence among the subordinates for initiation and upward communication. Therefore, we hypothesize manager’s pro-voice behavior to positively influence the subordinate’s upward voice:

\[ H2. \text{ Manager's pro-voice behavior positively relates to employee upward voice.} \]

2.3 Psychological safety
Psychological safety is an individual’s belief that the context is safe for interpersonal risk taking (Edmondson, 1999). In a work context, psychological safety might mean that individuals can engage themselves in work-related activities without fear of negative consequences from coworkers or people above them.

Voice often becomes more challenging than harmonious. To this extent, it becomes riskier for the initiator or speaker. Those who initiate are more likely to have direct or indirect professional or interpersonal implications as a result of voice construed as challenging the status quo than aligning to managers or organizational expectations. Further, it is cited that even when voice is intended to help others, it gets viewed as unacceptably challenging the authority, attention seeking, complaining and wasting time (LePine and Van Dyne, 1998). Hence, it is important to note that voice is perceived as personally risky for initiation (Detert and Burris, 2007; LePine and Van Dyne, 1998). It is widespread, only when those who initiate, perceive the negative consequences to be minimal or at least feel psychologically safe to speak as when doing so is not career damaging. Thus, considering the risky nature of voice behavior, we believe that psychological safety becomes more pertinent for employee upward communication to ensure whether signals and the situation are favorable.

Although work group members support often helps resolve several generic work-related issues in terms of carrying out tasks as expected and efficiently (Mayo et al., 2012), such support may not have a direct significant bearing on individual’s propensity to initiate upward voice. Hence, in a voice context, we argue that it is coworker’s upward voice that acts as a direct reference and helps provide essential psychological safety and support to alleviate stress associated with upward voice. Besides coworkers, as discussed above, manager
behavior is particularly salient for employees for the variety of reasons. Manager behavior is established to be pivotal across employee lifecycle. However, scant empirical work citing the effect of manager behavior on voice is inconclusive. In a voice context, it is important to move beyond mere supportive and cordial relationship to increase subordinate voice behavior. It is the interaction characterized by mutual respect between managers and subordinates that helps subordinates to comfortably express dissent. It is an environment where the manager’s behavior specifically signals openness to change, willingness to act on input from below, assures no negative repercussions for speaking up, that helps build comfort and psychological safety to voice and freely express themselves. Therefore, we hypothesize:

H3. Psychological safety mediates the relationship between the co-workers’ upward voice and employee upward voice.

H4. Psychological safety mediates the relationship between manager pro-voice behavior and employee upward voice.

3. Method
3.1 Participants and procedures
We chose a work setup which is communication and coordination intensive such as information technology (IT) firms in India, with ample scope and avenues for engaging in various formal/informal networks through several dynamic communication channels. Challenging the authority is culturally discouraged in countries like India, where power relationships (superior/subordinate relationships) are accepted as typical. This implies that although IT firms provide ample avenues for employee voice, employees engaging in upward voice are devious. With negligible recognition for worker union in this sector, we would expect interpersonal context at work and manager behavior to shape indirect and devious approaches employees employ in their potential upward voice.

We contacted non-managerial employees from various multinational IT services firms to examine their propensity toward upward communication and sought their consent to administer a self-reporting questionnaire. Purposive sampling helped to contact 1,685 employees, of which 949 employees agreed to participate and 625 employees responded, yielding a response rate of 65.86 percent. After data screening for incomplete or missing data, we retained 575 valid responses. Study sample represented varied educational background, gender and tenure of service. Respondents had an average work tenure of 4.2 years (SD = 2.9), with an average age of 26.9 years (SD = 3.5) and with greater representation from male respondents (67 percent).

3.2 Measures
3.2.1 Coworkers’ upward voice. Five items were adapted from the Van Dyne and LePine’s (1998) voice scale to measure coworkers’ upward voice. A sample item includes “My colleagues express their views and ideas to our manager about any work-related issues.” Participants respond to a five-point scale (1 = never, 5 = always). Exploratory factor analysis (EFA) results indicate the loading of all items on one factor. Confirmatory factor analysis (CFA) resulted in dropping of one item. The Cronbach’s $\alpha$ for the remaining four items was 0.73.

3.2.2 Manager pro-voice behavior. Six items were directed to assess how non-managerial employees perceive their manager’s demonstration of voice favoring behavior. This included behaviors such as involving team members in key decisions, consulting with employees before taking decisions that affect them, encouraging to express ideas/suggestions (Kim and Yukl, 1995). Sample items include: “My manager actively involves team members in key decisions” and “My manager encourages us to express our ideas/suggestions.”
Participants respond to a five-point scale (1 = never, 5 = always). To assess the scale’s structure, an EFA was conducted on the items, which pointed to an acceptable fit (KMO = 0.87; Bartlett’s sphericity = 1,366.85, p < 0.001). An examination of the component matrix with varimax rotation showed a one-factor solution where all factor loadings were greater than 0.50. The manager pro-voice scale was developed to assess employees’ evaluations of manager voice friendly behavior, and the factor structure supported the theoretical underpinnings since the items fit the corresponding factor. CFA resulted in dropping of one item. The Cronbach’s α for the remaining items was 0.88.

3.2.3 Psychological safety. Six items were adapted from previous studies (Liang et al., 2012). A sample item is “I can freely express my thoughts in my team.” Participants respond to a five-point scale (1 = strongly disagree, 5 = strongly agree). EFA results indicate the loading of all items on one factor. CFA resulted in the dropping of two items. The Cronbach’s α for the remaining items was 0.79.

3.2.4 Employee upward voice. Seven items were adapted from the Van Dyne and LePine’s (1998) voice scale. A sample item is “I express my opinions on any work-related issues to our manager even when others in the team disagree/hold a different opinion.” Participants respond to a five-point scale (1 = never, 5 = always). EFA results indicate the loading of all items on one factor. CFA resulted in the dropping of two items. The Cronbach’s α for the remaining five items was 0.84.

3.2.5 Control variables. Employee age and longer tenure can aid in developing greater familiarity with people and work systems to ward off potential inhibitions in upward voice. Hence, we controlled for age and work tenure (Tangirala and Ramanujam, 2008).

As we carried out this study in a context where speaking up is culturally discouraged, we focus on individual employee perception about their voice or silence, their sense of safety to voice and about their manager behavior and coworkers voice. Although multi-rater environment such as supervisors and peer rating on the outcome variable of interest – employee voice – appears ideal, confidentiality and identity concerns of the employees would be under served, adversely affecting the very key variables of the study such as employee psychological safety. Hence, respondents themselves would be ideal to report their perceived and felt psychological safety to speak up. Thus, a self-report measure may be more appropriate than supervisor or peer ratings. Thereby, we assured participant anonymity and random ordered the self-report measures. Further, such self-reported measures of employee voice are not uncommon in the voice and management literature (Axtell, et al., 2000; Ng and Feldman, 2013; Parker et al., 2006).

As self-reported data often render to common method bias (CMB), we employed multiple procedural remedies suggested by Podsakoff et al. (2003) to limit this bias such as ensuring respondent anonymity and randomizing the item order. Further, we determined the extent to which the self-reported measures may have inflated the relationships between the study variables by conducting post-hoc test. We conducted Harman’s single-factor test analyzing all variables in an EFA and examining the unrotated factor solution. CMB becomes an issue if a single factor emerges from the solution, or one general factor emerges accounting for the majority of the variance. Using the commonly accepted eigenvalue cutoff of 1.0, four factors have emerged, with the first factor accounting for approximately 38.78 percent of the total variance. These results suggest that CMB is a non-issue.

4. Results
4.1 Descriptive statistics
Table I provides the descriptive statistics (means and standard deviations), bivariate correlations and internal consistency reliability for the variables of interest in the study. We note employee upward voice having a positive correlation with coworkers’ upward voice
4.2 Structural equation modeling (SEM) We used AMOS 22 for SEM to test the proposed study model as SEM allows examining multiple dependent relationships simultaneously and incorporates both unobserved (latent) and observed variables in the same model. It also accounts measurement errors and indicates the extent to which data support study model than regression analyses (Gefen et al., 2000). Following the two-step analysis procedure, we first examined the measurement model and then the structural model. Measurement model helps us to infer the appropriateness and the strength of the relationships between observed variables and latent variables. We conducted CFA to examine the measurement model and to assess the discriminant validity of the latent variables. The theoretical four-factor model (co-worker’s upward voice, manager pro-voice behavior, psychological safety and employee upward voice) fits the data well with $\chi^2(\text{df} = 125) = 364.46$, $\chi^2/\text{df} = 2.92$, SRMR $= 0.05$, GFI $= 0.93$, CFI $= 0.95$, RMSEA $= 0.06$. All indicators had statistically significant factor loadings ($p < 0.01$), suggesting the convergent validity (Anderson and Gerbing, 1988). Also, a series of $\chi^2$ difference tests revealed that the four-factor model fits the data significantly better than several alternative measurement models (Table II). In all comparisons, alternative models yielded a significantly poorer fit. Taken together, these results favor the theoretical four-factor model, thus supporting discriminant validity among the measures.

We then examined the structural model to analyze the path relationship between constructs (latent variables), i.e., to test hypothesized relationships. The fit indices are initially checked to assess how well the hypothesized model fits the data set. The fit indices of the structural model were $\chi^2 = 364.46$, $\chi^2/\text{df} = 2.92$, SRMR $= 0.05$, GFI $= 0.93$, CFI $= 0.95$, RMSEA $= 0.06$, ensuring a good fit. The structural model analysis results with path

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$</th>
<th>SD</th>
<th>1 2 3 4 5 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>26.92</td>
<td>3.50</td>
<td></td>
</tr>
<tr>
<td>Tenure (years)</td>
<td>4.22</td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>Coworkers upward voice</td>
<td>3.47</td>
<td>0.71</td>
<td>0.09*</td>
</tr>
<tr>
<td>Manager pro-voice behavior</td>
<td>3.59</td>
<td>0.88</td>
<td>0.01</td>
</tr>
<tr>
<td>Psychological safety</td>
<td>3.70</td>
<td>0.71</td>
<td>0.05</td>
</tr>
<tr>
<td>Employee upward voice</td>
<td>3.55</td>
<td>0.76</td>
<td>0.17**</td>
</tr>
</tbody>
</table>

Notes: $n = 575$. Italic numbers on the diagonal are Cronbach’s $\alpha$. *, **Significant at the 0.05 and 0.01 levels (two-tailed), respectively.

Table II. Results for measurement model comparisons

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2(\text{df})$</th>
<th>$\chi^2/\text{df}$</th>
<th>SRMR</th>
<th>GFI</th>
<th>CFI</th>
<th>RMSEA</th>
<th>$\Delta\chi^2/\Delta\text{df}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical 4-factor model (PV, MP, PS, EV)</td>
<td>364.46 (125)</td>
<td>2.92</td>
<td>0.05</td>
<td>0.93</td>
<td>0.95</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>3-factor model 1 (PV and MP merged, PS, EV)</td>
<td>558.61 (128)</td>
<td>4.36</td>
<td>0.07</td>
<td>0.89</td>
<td>0.90</td>
<td>0.08</td>
<td>194.15 (3)**</td>
</tr>
<tr>
<td>3-factor model 2 (PV, MP and PS merged, EV)</td>
<td>578.67 (128)</td>
<td>4.52</td>
<td>0.06</td>
<td>0.89</td>
<td>0.89</td>
<td>0.08</td>
<td>214.21 (3)**</td>
</tr>
<tr>
<td>3-factor model 2 (PV and PS merged, MP, EV)</td>
<td>466.54 (128)</td>
<td>3.65</td>
<td>0.06</td>
<td>0.91</td>
<td>0.92</td>
<td>0.07</td>
<td>102.08 (3)**</td>
</tr>
<tr>
<td>2-factor model (PV, MP and PS merged, EV)</td>
<td>721.05 (130)</td>
<td>5.55</td>
<td>0.07</td>
<td>0.86</td>
<td>0.86</td>
<td>0.09</td>
<td>356.59 (5)**</td>
</tr>
<tr>
<td>1-factor model</td>
<td>1135.06 (131)</td>
<td>8.67</td>
<td>0.08</td>
<td>0.77</td>
<td>0.77</td>
<td>0.12</td>
<td>770.60 (6)**</td>
</tr>
</tbody>
</table>

Notes: PV, peer upward voice; MP, manager pro-voice behavior; PS, perceived psychological safety; EV, employee upward voice; SRMR, root-mean-square residual; GFI, goodness of fit index; CFI, comparative fit index; RMSEA, root-mean-square error of approximation. $\chi^2$ difference of each model reflects its deviation from the four-factor model. **$p < 0.01$.
coefficients are given in Figure 2. This model explained 61 percent variance of psychological safety and 57 percent variance of employee upward voice. Coworkers upward voice is positively related to employee upward voice ($\beta = 0.34$, $p < 0.001$) supporting for $H1$. However, manager pro-voice behavior is not significantly related to employee upward voice ($\beta = 0.04$, $p > 0.10$), not supporting $H2$.

4.2.1 Mediation tests. $H3$ and $H4$ make predictions about two mediating relationships. $H3$ predicts psychological safety to mediate the effects of coworkers’ voice on employee upward voice. $H4$ predicts psychological safety to mediate the effects of manager pro-voice behavior on employee speaking up. We followed the procedures outlined by Williams et al. (2009) for testing mediation. They suggest that a mediation effect can be concluded when the product of the path between the independent variable and the mediator (called path a) and the path between the mediator and the dependent variable (called path b) is statistically significant. This product approach has been recommended over Baron and Kenny’s (1986) method for its greater power and higher control for Type I errors (MacKinnon et al., 2002). Another merit of this approach is that it allows researchers to generate confidence intervals for indirect effects by using bootstrapping which is agnostic to sample size.

Basis this, we employed bootstrapping method in SEM and found both mediating relationships ($H3$ and $H4$) to be statistically significant. Psychological safety significantly mediated the effect of co-workers upward voice on employee upward voice. The estimated indirect effect was 0.10 with standard error 0.04 (95% bias corrected confidence interval 0.03, 0.18). Further, psychological safety significantly mediated the effects of manager pro-voice behavior on employee upward voice. The estimated indirect effect was 0.15 with standard error 0.06 (95% bias corrected confidence interval 0.04, 0.27). Zero being outside the confidence interval indicates the significance of mediating effect (see Table III).

4.3 Alternative models
To examine the potential alternative models to explain employee upward voice, we tested four alternative models as discussed below.

4.3.1 Alternative Model 1. Here, we posit coworkers upward voice and manager pro-voice behavior to have direct impact on employee upward voice with no mediating
processes or relationships specified among the latter constructs. The results in Table IV (row 3) demonstrate that this model has poorer fit than the original model. Particularly, the increase in $\chi^2$ value (which indicates badness of fit) in Alternative Model 1 is statistically significant ($\chi^2 = 10.57, \Delta df = 1, p < 0.01$), suggesting that the proposed model fits the data better than this model with no mediation effects. This model explained 65 percent variance of psychological safety and 56 percent variance of employee voice.

4.3.2 Alternative Model 2. Here, we posit that coworkers upward voice and manager pro-voice behavior impact employee upward voice only through their felt psychological safety with no direct relationships specified. The results in Table IV (row 4) demonstrate that Alternative Model 2 has poorer fit than the original model. Specifically, the increase in $\chi^2$ value (which indicates badness of fit) in Alternative Model 2 was statistically significant ($\chi^2 = 46.8, \Delta df = 2, p < 0.01$), thus suggesting that the study model fits the data better than the alternative model with only indirect effects. This model explained 70 percent variance of psychological safety and 51 percent variance of employee voice.

4.3.3 Alternative Model 3. Here, we controlled the effect of age and tenure on employee upward voice as they have an important correlation with employee upward voice. While age was not significantly associated with employee upward voice ($\beta = 0.02, p > 0.10$), we found job tenure to have a low magnitude positive association with employee upward voice ($\beta = 0.11, p < 0.10$). The high correlation between age and job tenure ($r = 0.84$) perhaps nullified the impact age would have on employee upward voice. To confirm the same, we controlled age in the absence of job tenure and found a positive association with employee upward voice ($\beta = 0.12, p < 0.01$). However, this model was not statistically different from proposed model ($\chi^2 = 38.62, \Delta df = 30, p > 0.10$). This suggests that the proposed mediation model is unaffected by the variation in the demographic characteristics of the respondents despite small changes in the $\beta$ values. This model explained 61 percent variance of psychological safety and 57 percent variance of employee voice.

4.3.4 Alternative Model 4. Here, we dropped the non-significant path (of the proposed model) between manager pro-voice behavior and employee upward voice. Though the results in Table IV (row 6) appear better than the proposed study model, it was not statistically significant ($\chi^2 = 0.36, \Delta df = 1, p > 0.10$), suggesting that the proposed

### Table III.

Bootstrapping results for mediation with 5,000 resamples at 95% confidence interval

<table>
<thead>
<tr>
<th>Path</th>
<th>Indirect effect (ab)</th>
<th>SE</th>
<th>Bias corrected percentile method LB</th>
<th>Bias corrected percentile method UB</th>
<th>Percentile method LB</th>
<th>Percentile method UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coworkers upward voice $\rightarrow$ psychological safety $\rightarrow$ employee upward voice</td>
<td>0.10</td>
<td>0.04</td>
<td>0.03</td>
<td>0.18</td>
<td>0.03</td>
<td>0.17</td>
</tr>
<tr>
<td>Manager pro-voice behavior $\rightarrow$ psychological safety $\rightarrow$ employee upward voice</td>
<td>0.15</td>
<td>0.06</td>
<td>0.04</td>
<td>0.27</td>
<td>0.04</td>
<td>0.27</td>
</tr>
</tbody>
</table>

**Notes:** LB, lower bound; UB, upper bound

### Table IV.

Alternative model comparison

<table>
<thead>
<tr>
<th></th>
<th>$\chi^2$</th>
<th>df</th>
<th>$\chi^2$/df</th>
<th>SRMR</th>
<th>GFI</th>
<th>CFI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>364.46</td>
<td>125</td>
<td>2.92</td>
<td>0.05</td>
<td>0.93</td>
<td>0.95</td>
<td>0.06</td>
</tr>
<tr>
<td>Model 1</td>
<td>181.47</td>
<td>70</td>
<td>2.59</td>
<td>0.04</td>
<td>0.96</td>
<td>0.97</td>
<td>0.05</td>
</tr>
<tr>
<td>Model 2</td>
<td>411.26</td>
<td>127</td>
<td>3.23</td>
<td>0.05</td>
<td>0.92</td>
<td>0.94</td>
<td>0.06</td>
</tr>
<tr>
<td>Model 3</td>
<td>403.08</td>
<td>155</td>
<td>2.60</td>
<td>0.05</td>
<td>0.93</td>
<td>0.95</td>
<td>0.05</td>
</tr>
<tr>
<td>Model 4</td>
<td>364.82</td>
<td>126</td>
<td>2.89</td>
<td>0.05</td>
<td>0.93</td>
<td>0.95</td>
<td>0.06</td>
</tr>
</tbody>
</table>
mediation model and Alternative Model 4 are similar. This model explained 61 percent variance of psychological safety and 56 percent variance of employee voice.

Taken together, these results support the study model over all alternative models to explain employee upward voice.

5. Discussion
Most progressive firms desire their employees to express their ideas and concerns in a timely manner as it aids them in early detection of errors and prepares them for potential adversities. Given the critical role of employee input, there is growing attention from scholars and practitioners alike, toward understanding enablers of employee upward voice.

While extant literature predominantly cites the influence of individual level antecedents and leadership styles to explain employee voice behavior, we set out to investigate the antecedents of employee upward voice by commissioning the assumptions of social information processing and social adjustment. Although interpersonal context and informal/networked social systems at work organize and shape tacit knowledge toward human action, given the risky nature and non-specification of voice as in-role job requirement, employees often embark on this tacit knowledge toward assurance and psychological safety to voice. Accordingly, perceived psychological safety was examined as an explanatory mechanism in employee voice studies (Liu et al., 2010, 2015). While in search for what ensures safety, extant work independently examines the interpersonal aspects of coworkers or managers, whereas we examined a common effect of coworker’s voice and manager pro-voice behavior to infuse necessary psychological safety to voice. Through this study, we highlight the role of immediate work group setup which perhaps helps create a sense of acceptable range of action in the organization to better explain risky upward voice and detailed on the application of social information processing and social learning in employee upward voice studies. The results of the study suggest the employee perception of coworkers’ upward voice as well as manager pro-voice behavior to influence their upward voice via their perceived psychological safety. This highlights how employees learn to voice or remain silent vicariously, through their observation and modeling of others/other’s cues at work (Bandura, 1986).

The study contributes to the scant but a growing body of research on factors that influence the employee upward voice. By examining the role of interpersonal context, we significantly expand the knowledge about how peers and managers can foster psychological safety for upward voice. Barring some of the work by Liu et al. (2015), Xu et al. (2017) and Walumbwa et al. (2012), not much is examined on how interpersonal dynamics involving both peers and managers concurrently determine employee voice and voice standards. The present study in this direction makes important assertions. The study results highlight how coworkers voice behavior determines employee vicarious input, initiation and voice. Observation and interactions with others at work seem to provide a ready interpretation of things in terms of what and when to speak up. Coworkers upward voice became a ready reference to gauge the favorableness and approachability of manager-subordinate dyadic context. These have implications not only in terms of member’s attitude toward and evaluation of upward communication, but also in their propensity to raise relevant issues with the management.

The study results do not support any significant direct relationship between manager pro-voice behavior and employee upward voice. This indicates that manager pro-voice behavior is not a sufficient factor to facilitate subordinates upward voice. Given that managers facilitate subordinates’ development and growth, subordinates often refrain from speaking up for, it may be perceived as confronting them. Moreover, the finding that coworkers’ upward voice alone has a significant and positive direct impact on employee upward voice reinforces the assertion that an employee models their coworkers’
voice behavior. This indicates that subordinates relate to their work group better than to the power relationship of subordinate-supervisor interactions. People develop proximity, affiliation and quality relationships more in a group and interdependent work context. This aids in peers and coworkers having a strong social influence on employee initiation behavior. Thus, co-workers upward voice behavior seemed to have provided a sense of safety and positive signals, aiding employee upward voice. Given that the model was tested in a culture where confronting those in authority is discouraged, respondents seem to maintain a distance with their managers or perhaps were seeking lesser managerial intervention.

Further, the study situates psychological safety as a key mechanism between interpersonal context and upward communication. The results indicate that psychological safety mediates the relationship between coworkers’ upward voice, manager pro-voice behavior and employee upward voice. We found the indirect effect of psychological safety in the link between coworker’s upward voice and employee upward voice, to be of lesser magnitude. Beyond serving as normative reference point, coworkers seem to have a limited role in facilitating a sense of psychological safety to speak up. Further, peers and coworkers may only influence to an extent, for most subordinates often have or aspire to have a unique dyadic relationship with their supervisors. Also, since opportunities for collective or representational voice are limited in these firms, most employees depend on their one-on-one direct communication. Hence, we found a limited role of coworkers in creating a sense of psychological safety for an employee to voice upward.

While no direct effect, we found a significant indirect effect of psychological safety in the link between manager pro-voice behavior and employee upward voice. This clearly indicates that it is only through perceived and felt safety without fear of negative consequences from the manager, that the manager’s pro-voice behavior would be leveraged by the subordinates. It is the interactions characterized by mutual respect between managers and subordinates that help subordinates to comfortably express dissent. It is manager pro-voice behavior that specifically signals safety assuring no negative repercussions for speaking up, helping build comfort to voice and freely express themselves.

The mediating role of psychological safety in the interpersonal context and upward voice can also be viewed from national cultural standpoint. Relationships and respect for authority have immense significance in Indian society. Consideration and maintaining social ties and warm relationships with others is important for collectivist society. Perceived and felt safety ensures endured relationships with both superiors and coworkers. As employee upward voice has several important employee level and organizational benefits, the study findings highlight how the perceptions of favorable interpersonal context impact the employee sense of psychological safety motivating them for honest upward communication. Further, with the increasing number of global firms’ business operations in India, an understanding of what may be emphasized in a hierarchal and relational society is an important contribution of this study.

6. Managerial implications, limitations and future scope
The concept model of this study has important implications for organizations. Basis findings, we see manager behavior as key to employee psychological safety. Voice can be optimized to the extent that employees feel safe to engage in it. Given that there is reluctance for employee upward input in certain contexts, managerial/organizational efforts must focus toward counterbalancing the restraining forces by enabling psychological safety. Hence, to build the perceptions of favorable interpersonal context among employees, organizations must train and develop their managers on how to give and receive feedback, listen and communicate in a non-defensive manner. Organizations may also recognize and incentivize favorable manager behavior to encourage positive
work attitudes in the firm. This can effectively bring in the perceptions of positive organizational and managerial support. Also, practices that promote the fair treatment of employees constitute a source of psychological safety.

It is important to move beyond manager influence in general, to more specific actions they consciously or unconsciously conduct, toward comprehensively understanding employee reluctance in providing their input. Often, managers build unseen walls between them and their team members not allowing alternate viewpoints in key decisions. Grounded in empirical research evidence, we highlighted that there are many different ways that managers can enable upward communication and input. Managers may recognize several situational and social factors influencing employee sense of safety toward upward input. With a work culture of honest communication, voice can become a positive experience for employees. Candor and honest communication allow greater trust which leads employees to feel more valued and connected to the organization and its members. This not only facilitates upward input, but also communication truly distributed across the levels of the organization. As upward input is essential for organizational growth, firms must find a way to encourage and sustain employee voice across the levels of the organization. Further, employees, particularly in collectivist countries such as India, look up to their organization for support and care. Supervisors play a critical role in the employee–organization relationship. People in decision-making capacities have an important role in fostering a fair and healthy work environment. Therefore, organizations must ensure managerial effectiveness toward subordinate’s fair treatment and dignity. It is important for organizational management to break psychological barriers and unseen walls, toward engaged workforce freely exchanging views to promote and sustain organizational trust.

We suggest firms to nurture voice friendly culture promoting candor. This ensures employee safe interpersonal risk taking in terms of expressing dissent, despite hierarchy. This aids organizations in translating vocal culture into their competitive advantage.

6.1 Limitations and future scope
This study is not without limitations. The cross-sectional design of the study does not render itself to examine dynamic and causal inferences. Future work may entail time series or longitudinal studies toward this end. We used employee perceptual data as managers influence, rest on the subordinates’ perception of the same, even though such perceptions can be factually wrong. Further, study data represent a hierarchical cultural setup in technology firms in India. Future studies must test employee voice behavior in diverse cultural and occupational settings with multisource, multi method data to enhance external validity. Besides, as study highlights vicarious social learning in emulating voice/silence, it is important to note how employee voice behavior is malleable. Future studies may examine how group dynamics shift voice standards or influence voice frequency.

7. Conclusion
Extant management literature often cites the need to study any phenomenon within a specific context toward enabling its comprehensive understanding. South Asian countries, specifically India for its diverse culture, demand the creation of pluralistic and indigenous management scholarship (Khatri et al., 2012). The present study provides evidence that psychological safety is important toward enabling upward input among the employee base in India. Voice favoring organizational culture, specifically as enabled by immediate interpersonal context, such as managers and coworkers, can enhance interpersonal risk taking and a sense of safety to speak up. The results may be useful for guiding the future theory development of employee upward voice in a cross-cultural context.
References


Appendix

<table>
<thead>
<tr>
<th>Construct</th>
<th>References</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-worker upward voice</td>
<td>Van Dyne and LePine (1998)</td>
<td>My colleagues express their views and ideas to our manager about any work-related issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My colleagues express their opinions on work-related issues even when manager disagrees/holds a different opinion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My colleagues express their opinions on work-related issues even when others in the team disagree/hold a different opinion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My colleagues encourage others in the team to speak to manager about work-related issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;My colleagues get involved in issues that affect the quality of work life in our team</td>
</tr>
<tr>
<td>Manager pro voice</td>
<td>Kin and Yukl (1995)</td>
<td>My manager actively involves team members in key decisions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My manager consults with us to get our opinions and suggestions, before taking any decision that affects us</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My manager encourages us to express our ideas/suggestions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My manager invites us to participate in our team meetings/discussions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My manager welcomes open discussion of ideas from team members</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;My manager actively seeks our views on various aspects of job</td>
</tr>
<tr>
<td>Psychological safety</td>
<td>Liang et al. (2012)</td>
<td>I feel secure to express my opinions to my manager, in my team</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I can freely express my thoughts in my team</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I feel safe to challenge my manager's viewpoints/decisions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There is a supportive climate in my team to express our views</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;I am worried that expressing true feelings in my team would do me harm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;Nobody will pick on me even if I differ with my teams' opinions</td>
</tr>
<tr>
<td>Employee upward voice</td>
<td>Van Dyne and LePine (1998)</td>
<td>I convey my ideas to manager on work-related issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I give my opinions on work-related issues to manager even when manager disagrees/holds a different opinion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I express my opinions on work-related issues to manager even when others in the team disagree/hold a different opinion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I give suggestions to my manager on any work-related issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I express divergent views with my manager in the work-related meetings/discussions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;I speak up for new projects or changes in the procedures, in our team</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;I get involved in issues that affect the quality of work life in our team</td>
</tr>
</tbody>
</table>

Note: "CFA of each construct resulted in the dropping of these items

Table AI. Items of study constructs

About the authors

Sriji Edakkat Subhakaran is PhD Scholar (UGC-SRF) at the Department of Management Studies, Indian Institute of Technology Madras, Chennai, India. Her research interests are in organizational behavior and human resource management. Her current research topic is determinants of employee voice. Sriji Edakkat Subhakaran is the corresponding author and can be contacted at: srijies@gmail.com

Lata Dyaram is Associate Professor at the Department of Management Studies, IIT Madras. Specialising in Industrial/Organizational psychology, she currently teaches organizational theory, leadership and organization development. Her research and publication interests include cognition, motivation emotion and behavior in organizations.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldsight.com
A performance analysis of Brazilian public health: TOPSIS and neural networks application

Claudia Affonso Silva Araujo, Peter Wanke and Marina Martins Siqueira
COPPEAD Graduate School of Business, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Abstract

Purpose – The purpose of this paper is to estimate the performance of Brazilian hospitals’ services and to examine contextual variables in the socioeconomic, demographic and institutional domains as predictors of the performance levels attained.

Design/methodology/approach – The paper applied a two-stage approach of the technique for order preference by similarity to the ideal solution (TOPSIS) in public hospitals in 92 Rio de Janeiro municipalities, covering the 2008–2013 period. First, TOPSIS is used to estimate the relative performance of hospitals in each municipality. Next, TOPSIS results are combined with neural networks in an effort to originate a performance model with predictive ability. Data refer to hospitals’ outpatient and inpatient services, based on frequent indicators adopted by the healthcare literature.

Findings – Despite a slight performance increase over the period, substantial room for improvement is observed. The most important performance predictors were related to the demographic and socioeconomic status (area in square feet and GDP per capita) and to the juridical nature and type of ownership of the healthcare facilities (number of federal and private hospitals).

Practical implications – The results provide managerial insights regarding the performance of public hospitals and opportunities for better resource allocation in the healthcare sector. The paper also considers the impact of external socioeconomic, demographic and institutional factors on hospitals’ performance, indicating the importance of integrative public health policies.

Originality/value – This study displays an innovative context for applying the two-stage TOPSIS technique, with similar efforts not having been identified in the healthcare literature.

Keywords Performance, Hospitals, Brazil, TOPSIS, Neural networks, Health services

1. Introduction

The availability of resources within healthcare systems worldwide and their equitable and efficient allocation are central issues for the delivery of universal health coverage. However, inefficiencies are responsible for wasting 20–40 percent of all global health spending (World Health Organization, 2010). While the aging population and increased incidence of chronic disease put pressure on resource utilization, there is also a widespread increase in healthcare costs, related to the rapid adoption of technology in the sector, making more sophisticated and expensive treatments and equipment available (Sun and Luo, 2017).

Besides these challenges, Brazil, as a developing country of continental dimensions, with a population over 224m people and widespread inequalities, faces many structural problems at the national and regional levels that increase the complexity and interest in the efficient allocation of resources to provide health services to the population (Paim et al., 2011).

In 2014, Brazil was number 54 out of 55 countries in terms of efficiency, according to the Bloomberg Healthcare Efficiency Index, which assesses life expectancy, healthcare spending per capita and healthcare spending as a share of GDP (Bloomberg, 2014). Also, a World Bank Group report on public spending identified possible efficiency gains of 37 percent in public primary care (with the potential to save $2.7bn) and 71 percent in public secondary and tertiary care (with the potential to save an additional $3.6bn). These potential
savings in health spending showed great variation among municipalities, depending on the geographic location and population size and came from different sources, such as inefficient scale of hospital services, non-coordination among healthcare providers, lack of incentives to improve the productivity of healthcare professionals and hospital payments misaligned with the cost and quality of healthcare provided (World Bank, 2017).

The healthcare sector in Brazil is divided into the public and private systems. The Unified Health System (SUS) is one of the largest and most comprehensive public health systems in the world (Marinho, 2017) and offers free health services to all Brazilian citizens through public institutions at the federal, state and municipal levels, as well as private institutions, which provide services to SUS (beds and assistance) through agreements with public agencies (Araújo et al., 2013). Despite the universal coverage, as a constitutional right, the Brazilian population commonly resorts to the private system because of long queues and poor service quality. In 2016, approximately 30 percent of the Brazilian population had private health insurance and 3.5 percent of the people spent more than 25 percent of household income on out-of-pocket healthcare expenditure (DATASUS, 2016).

SUS funding comes mostly from the federal government and each state and municipality must spend at least 15 percent of its annual tax income on health, being responsible for controlling its health financial resources, actions and services (LC, 8080/90). SUS expenditure surpasses $123bn, representing about 8 percent of national GDP (Brazilian Institute of Geography and Statistics, 2013). Even with expenditure comparable to Latin American countries of similar GDP per capita, such as Peru and Colombia, Brazilian healthcare spending is below many European countries with similar health proposals, like Spain and the UK (World Health Organization, 2014). In 2011, the Ministry of Health created the “Best at Home” program (BHP) to expand public home care for sick patients, reducing costs and the demand for hospital public services (Ministério da Saúde, 2017). Besides operating within tight budgets, the ongoing economic and political crisis (Bahia et al., 2016) leads to pressure on public health, with SUS accommodating millions of unemployed users who have lost access to private health plans (Watts, 2016).

Another particular issue amplifying SUS management complexity is the size and heterogeneity of Brazilian territory. Among the 27 Brazilian states, Rio de Janeiro (RJ) stands out for its socioeconomic relevance — representing 8.4 percent of all population and the second biggest GDP per capita — and the intensity and complexity of its healthcare services. The state accounts for a significant portion of the nation’s healthcare patients, hospitals and public spending; in 2016, RJ state had the sixth and the third largest expenditure with SUS inpatient and outpatient services, respectively, representing a budget of more than $2m (DATASUS, 2016). In this same year, there were 499 hospitals in the state, accounting for 7.5 percent of all Brazilian hospitals (Brazilian National Register of Health Facilities, 2016). The great socioeconomic and health inequalities within this state also feature an interesting characteristic for performance measurement: the state contains municipalities with GDP per capita and Human Development Index (IDH) values varying from $12,385.17 and 0.837 at Niterói to $5,120.62 and 0.668 at Laje do Muriaé (Brazilian Institute of Geography and Statistics, 2010). Table I presents some demographic, socioeconomic and health indicators from Brazil and RJ, also portraying the public–private arrangement in the healthcare sector.

Despite its managerial complexity, the national literature on health performance and its underlying factors is still scarce. To date, no study in Brazil has addressed the following questions:

(1) What is the performance of Brazilian hospitals in RJ state, aggregated at municipality level, applying a two-stage technique for order preference by similarity to the ideal solution (TOPSIS)?

(2) How has performance changed over the years?
Can socioeconomic, demographic and institutional variables predict the performance scores attained?

The specific objectives are to assess the performance of inpatient and outpatient public services at RJ hospitals from 2008 to 2013, aggregated at municipality level, and to examine contextual variables in the socioeconomic, demographic and institutional domains as predictors of performance levels.

The aim of this study is to contribute to the healthcare performance literature in a number of ways: providing new empirical evidence on hospital services’ performance and bringing a methodological approach not yet explored in this field. Also, this study includes the complete population of municipalities in RJ, which avoids sample bias, and analyses high-quality data obtained from an official source – the National Register of Health Facilities (CNES). It employs a two-stage TOPSIS – a widely accepted multiple criteria

<table>
<thead>
<tr>
<th>Brazilian Demographic Census (IBGE, 2013)</th>
<th>Brazil (Rank: World)</th>
<th>Rio de Janeiro (Rank: Brazil)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area (sq. ft.)</td>
<td>9,162,965 (56th)</td>
<td>47,107 (25th)</td>
</tr>
<tr>
<td>Population</td>
<td>224,554,744 (6th)</td>
<td>17,989,596 (3rd)</td>
</tr>
<tr>
<td>GDP per capita (USD)</td>
<td>15,128 (73th)</td>
<td>12,875 (83rd)</td>
</tr>
<tr>
<td>Human Development Index</td>
<td>0.754</td>
<td>0.761</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>World Bank Indicators</th>
<th>Brazil</th>
<th>Latin American and Caribbean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Births attended by skilled health staff (% of total) (2013)</td>
<td>99.1</td>
<td>94.18</td>
</tr>
<tr>
<td>Physicians (per 1,000 people) (2013)</td>
<td>1.8</td>
<td>1.98</td>
</tr>
<tr>
<td>Population without access to an improved water source (%) (2015)</td>
<td>1.9</td>
<td>5.4</td>
</tr>
<tr>
<td>Prevalence of children underweight for age (% under age 5) (2007)</td>
<td>2.2</td>
<td>2.9</td>
</tr>
<tr>
<td>Life expectancy at birth, total (years)</td>
<td>75.2</td>
<td>75.3</td>
</tr>
<tr>
<td>GINI index (World Bank estimate) (2015)</td>
<td>51.3</td>
<td>-</td>
</tr>
<tr>
<td>Poverty headcount ratio at national poverty lines (% of population) (2015)</td>
<td>8.7</td>
<td>-</td>
</tr>
<tr>
<td>Poverty headcount ratio at $5.50 a day (2011 PPP) (% of population)</td>
<td>22.1</td>
<td>27.3</td>
</tr>
<tr>
<td>People without basic sanitation services (% population) (2015)</td>
<td>13.9</td>
<td>14.3</td>
</tr>
<tr>
<td>Mortality rate, maternal (per 100,000 live births) (2015)</td>
<td>44</td>
<td>67</td>
</tr>
<tr>
<td>Mortality rate, neonatal (per 1,000 live births) (2016)</td>
<td>8.2</td>
<td>9.2</td>
</tr>
<tr>
<td>Cause of death, by non-communicable diseases (% of total) (2016)</td>
<td>73.3</td>
<td>73.9</td>
</tr>
<tr>
<td>Population ages 65 and above (% of total) (2017)</td>
<td>8.2</td>
<td>7.8</td>
</tr>
<tr>
<td>Unemployment, total (% of total labor force) (2017)</td>
<td>13.6</td>
<td>8.7</td>
</tr>
<tr>
<td>Health expenditure, total (% of GDP) (2015)</td>
<td>8.3</td>
<td>7.1</td>
</tr>
<tr>
<td>Health expenditure, public (% of GDP) (2015)</td>
<td>3.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Health expenditure, public (% of total health expenditure) (2015)</td>
<td>46.0</td>
<td>51.2</td>
</tr>
<tr>
<td>Health expenditure, private (% of GDP) (2015)</td>
<td>4.5</td>
<td>3.5</td>
</tr>
<tr>
<td>Health expenditure per capita (current USD) (2015)</td>
<td>947.4</td>
<td>712.6</td>
</tr>
<tr>
<td>Out-of-pocket health expenditure (% of total expenditure on health) (2015)</td>
<td>25.5</td>
<td>31.7</td>
</tr>
<tr>
<td>SUS Data System (DATASUS, 2016)</td>
<td>Brazil</td>
<td>Rio de Janeiro</td>
</tr>
<tr>
<td>Hospitals, total</td>
<td>6,820 (100%)</td>
<td>499 (100%)</td>
</tr>
<tr>
<td>Public hospitals</td>
<td>2,423 (35.5%)</td>
<td>162 (32.5%)</td>
</tr>
<tr>
<td>Federal hospitals</td>
<td>86 (1.3%)</td>
<td>27 (5.4%)</td>
</tr>
<tr>
<td>State hospitals</td>
<td>628 (9.2%)</td>
<td>34 (6.8%)</td>
</tr>
<tr>
<td>Municipal hospitals</td>
<td>1,705 (25%)</td>
<td>101 (20.2%)</td>
</tr>
<tr>
<td>Private hospitals</td>
<td>2,540 (37.2%)</td>
<td>261 (52.3%)</td>
</tr>
<tr>
<td>Non-profit hospitals</td>
<td>1,857 (27.3%)</td>
<td>76 (15.2%)</td>
</tr>
<tr>
<td>Health services registered on the Brazilian Ministry of Health</td>
<td>313,622 (100%)</td>
<td></td>
</tr>
<tr>
<td>Hospital beds</td>
<td>492,208 (100%)</td>
<td>46,558 (100%)</td>
</tr>
<tr>
<td>Unified Health System (SUS)</td>
<td>337,244 (68.4%)</td>
<td>26,791 (57.5%)</td>
</tr>
<tr>
<td>Non-SUS</td>
<td>155,964 (31.6%)</td>
<td>19,767 (42.5%)</td>
</tr>
<tr>
<td>Private Health Plans (no. of firms operating)</td>
<td>1,276</td>
<td>-</td>
</tr>
</tbody>
</table>

Table I. Brazil socioeconomic and hospital features

decision-making technique – combined in the second stage with neural network analysis, to assess the prediction of performance levels through contextual variables that help discriminate the contexts wherein each municipality operates (Das, 2010).

TOPSIS methodology, developed by Hwang and Yoon in 1981, employs the notion of positive and negative ideal solutions as references to estimate and rank performance among units of analysis. In this study, the performance refers to the productive process in which human and material resources are employed for the delivery of the hospital services analyzed, and the positive ideal solution represents the concurrent minimization of input consumption and maximization of output production, while the negative ideal solution represents the concurrent maximization of input consumption and minimization of output production. TOPSIS has been successfully applied across diverse research areas, including supply chain, logistics, engineering, manufacturing, business and marketing systems. However, regardless of the methodological strength and potentiality in terms of systematically assessing the performance and its determinants across diverse sectors, TOPSIS has a limited number of applications in the healthcare activities (Behzadian et al., 2012), with no TOPSIS applications regarding the performance of hospital services identified so far.

Thus, in addition to providing new empirical evidence on hospital services in Brazil, academically this study enriches this scarce literature and stimulates new applications of TOPSIS methodology to analyze the performance of healthcare services. The innovative aspect of this study consists in combining TOPSIS and neural networks in a new context.

Besides the academic relevance, in the Brazilian hospitals’ reality of scarce resources and poor infrastructure, this approach may help public managers and decision-makers identify waste and inefficiencies in the health services. Measuring performance within a specific sector or activity allows the comparison of relevant peers and the acknowledgment of which units are poorly performing, compared to the others in the sample, also presenting an opportunity to examine the successful practices carried out in the best performing units (Bogetoft and Otto, 2011).

The study is structured as follows. This introduction presented a contextual setting of the Brazilian health system. The next section provides a brief review of previous performance studies, including TOPSIS applications. The last sections exhibit the data and empirical results, as well as their discussion and conclusion.

2. Related literature

The healthcare literature on performance measurement has increased significantly in the last decades, providing valuable information for policy makers (Hollingsworth, 2008). The analyses are conducted at different levels of aggregation, such as healthcare units (Ketabi, 2011; Kirigia et al., 2007) and hospitals (Chowdhury et al., 2011; Narci et al., 2015; Rezaei et al., 2016; van Ineveld et al., 2016). The health services are also aggregated at municipality (Halkos and Tzeremes, 2011; Varela et al., 2012), state and country levels (Hadad et al., 2013). They encompass several scientific methods, with a predominance of tools such as data envelopment analysis (DEA), stochastic frontier analysis (SFA), Malmquist index (MI) and several indicators formed by the quotient of one input and one output. Additionally, hospitals constitute the predominant unit of analysis (Hollingsworth, 2003, 2008; Hollingsworth et al., 1999; Hussey et al., 2009).

In most studies, outputs comprise health services (i.e. hospital procedures) (Hollingsworth, 2003, 2008; Hussey et al., 2009) and inputs focus on several types of resources and costs (Hussey et al., 2009). Besides measuring performance, the literature inquires about its determinants, including socioeconomic (i.e. income, education) and structural characteristics (i.e. type of ownership, market concentration, presence of a given technology) (Ancarani et al., 2009; Lindlbauer and Schreyögg, 2014).
The articles reviewed used mostly frontier methods as SFA (Lindlbauer and Schreyögg, 2014; Meyer, 2015; Rezaei et al., 2016; Xu et al., 2015) and DEA (e.g. Kirigia et al., 2007; Lindlbauer and Schreyögg, 2014; Narci et al., 2015; Tiemann and Schreyögg, 2012; van Ineveld et al., 2016; Xu et al., 2015). Many studies also take a two-stage approach, using methods such as Tobit (Ancarani et al., 2009; Dimas et al., 2012; Narci et al., 2015), bootstrap (Lindlbauer and Schreyögg, 2014) and translog regression (Almeida and Cima, 2015; Azevedo and Mateus, 2014). Other techniques combined to enrich the models are MI (Dimas et al., 2012; Kirigia et al., 2007; van Ineveld et al., 2016) and principal components analysis (PCA) (Xu et al., 2015). MI allows longitudinal performance analysis, noting productivity changes over time, while PCA sums up input or output variables in principal components that represent much of the variation in the original data, allowing the increase of discrimination power in the model. Table II presents a summary of these research efforts. The variables vary considerably among articles, although often seeking to represent healthcare industry characteristics. Despite a growing number of empirical studies in the hospital performance literature, there is no consensus about what is a desirable performance level, considering the different national realities and the basic dissimilarities in the studies' design and samples. Among the few studies with a broad picture of hospital performance are the works of Hollingsworth (2008) and Varabyova and Schreyögg (2013). Hollingsworth (2008) performed a systematic review of 317 published papers on frontier efficiency measurement in the health services sector, with 52 percent of applications in hospitals. Most studies are from the USA, where the average efficiency is 0.826. Next comes the European sample with an average efficiency of 0.86. In turn, Varabyova and Schreyögg (2013) used parametric and non-parametric approaches to compare the technical efficiency in the hospital sector of 31 countries in the Organization for Economic Co-operation and Development. Using nine model variations of DEA and three of SFA, the authors found average efficiency scores ranging from 0.77 to 0.86.

Some academic studies have analyzed the performance of Brazilian public health services at the municipality level. Marinho (2003) employed DEA to assess the performance of inpatient and outpatient services in the municipalities of RJ state in 1998, finding room for improvement – the average efficiency score was 0.83 – and great discrepancy in the efficiency of municipalities. Varela et al. (2012) measured the performance of municipalities in São Paulo (SP) state in 2006, also using DEA, and found great discrepancies among the municipalities in the sample, with an average score of only 0.45. In turn, Santos et al. (2014) applied DEA to assess the performance of public primary care services in Brazilian municipalities located in the Southeastern Region, including RJ and SP states, from 2007 to 2010, and found a low overall average score (0.61) and just a slight improvement within the period analyzed.

Regarding TOPSIS, the healthcare sector is a recent field for the use of this methodology (Behzadian et al., 2012). Among the few applications, the technique has been used for assessing health technology projects in Iran, helping to set the priority of its practical application (Mobinizehdeh et al., 2016), and for ranking cultural aspects of patient safety in hospitals emergency wards in Iran (Tourani et al., 2015). A fuzzy-TOPSIS – TOPSIS combined with fuzzy theory, allowing non-limited numbers of criteria and alternatives – has also been applied to rank the failure risks in outpatient service for elderly patients in Taiwan (Kuo et al., 2012), to measure the performance of hospital managers in Iran (Shafii et al., 2015) and to evaluate risks associated with human health in coalmines in Iran (Mahdevari et al., 2014). No previous study has focused on the productive performance of hospital services, regarding the conversion of human, physical and infrastructure resources into expected health outcomes.

In addition, although two-stage models combining efficiency assessment with efficiency prediction are emerging in the literature, they mainly use DEA. The applications of multiple
<table>
<thead>
<tr>
<th>Authors</th>
<th>Unit</th>
<th>Method</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marinho (2003)</td>
<td>Public hospitals in 23</td>
<td>DEA-CCR combined with Tobit regression</td>
<td>Beds, accredited establishments and outpatient installed capacity per</td>
<td>Visits with hospitalization in accredited hospitals and outpatient</td>
<td>Population and GDP of municipalities; average length of stay</td>
</tr>
<tr>
<td></td>
<td>municipalities of Rio de</td>
<td></td>
<td>capita; average cost of outpatient procedures and hospitalization</td>
<td>visits (per capita); mortality rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Janeiro</td>
<td></td>
<td>No. of doctor and nurse hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kirigia et al.</td>
<td>17 public health centers of</td>
<td>DEA-VRS</td>
<td>Hospitalizations; cases treated, no. of children immunized and those</td>
<td>No. of patients dressed, domiciliary cases treated, no. of children</td>
<td>Presence of computerized patient management; conflict among physicians;</td>
</tr>
<tr>
<td></td>
<td>Seychelles</td>
<td></td>
<td>participating in school health program</td>
<td>immunized and those participating in school health program</td>
<td>importance of profit maximization; importance of hospital prestige's</td>
</tr>
<tr>
<td>Ancarani et al.</td>
<td>48 Italian hospitals</td>
<td>DEA-CCR, BBC, NIRS combined with Tobit</td>
<td>No. of Hospital beds; physicians; non-medical staff; turnover of</td>
<td>Hospital discharges; patients treated per day; outpatient care</td>
<td>presence of surgical unit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>regression</td>
<td>operating rooms; maintenance costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chowdhury et al.</td>
<td>113 acute care hospitals in</td>
<td>MI combined with bootstrap technique</td>
<td>No. of staffed beds, FTE nurses and administrative workers; purchased</td>
<td>No. of inpatient and outpatient patient visits</td>
<td>Not employed</td>
</tr>
<tr>
<td></td>
<td>Ontario</td>
<td></td>
<td>services and supplies, equipment expenses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimas et al.</td>
<td>27 Greek hospitals</td>
<td>DEA–BBC combined with Tobit regression and</td>
<td>No. of hospital beds; wages expenditure; total expenditure</td>
<td>Days of hospitalization; care without hospitalization;</td>
<td>Bed occupancy rate; average length of stay of patients; average cost of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MI</td>
<td></td>
<td>emergency care</td>
<td>patient; average no. of daily patients</td>
</tr>
<tr>
<td>Tiemann and</td>
<td>548 public hospitals of acute</td>
<td>DEA–BBC combined with linear regression</td>
<td>Amount spent on supplies per year; no. of FTE physicians, nursing,</td>
<td>No. of inpatient cases per year, average in-hospital mortality</td>
<td>No. of licensed and staffed beds, proportion of all hospital beds that</td>
</tr>
<tr>
<td>Schreyögg (2012)</td>
<td>care in German</td>
<td>model for truncated longitudinal data</td>
<td>other clinical, administrative and nonclinical staff</td>
<td>rate per year</td>
<td>had been leased, presence of teaching activities for the training of</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>non-medical staff, Hirschman-Herfindahl index (HHI), market area,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Elixhauser Comorbidity Index, pre-privatization efficiency,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>competitive environmental</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Authors</th>
<th>Unit</th>
<th>Method</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Varela et al. (2012)</td>
<td>Primary care services from 603 municipalities of São Paulo</td>
<td>DEA–BBC combined with regression analysis</td>
<td>Primary care expenditure</td>
<td>Nursing actions and other health procedures; basic medical actions; basic actions in dentistry; professionals registered by the PSF/PACS (health family program/program of community health agents)</td>
<td>Urban population; age range (0–18); Age group (60 and over); population density; range of health facilities; % of own resources invested in health</td>
</tr>
<tr>
<td>Varabyova and Schreyögg (2013)</td>
<td>31 OECD countries</td>
<td>SFA, DEA (input-oriented, BBC and CRC, scores corrected for bootstrap bias) combined with bootstrap regression</td>
<td>No. of beds, nurses, physicians, hospital employees</td>
<td>No. of discharges, mortality</td>
<td>Health expenditure, hospital expenditure, private sector expenditure, income inequality, hospital per million population, % of publicly owned hospitals, average length of stay (inpatient care), % population with 65 years old or more, % population with upper secondary education, life expectancy at birth, infant mortality, incidence of full-time employment</td>
</tr>
<tr>
<td>Gok and Sezen (2013)</td>
<td>348 Turkish hospitals</td>
<td>DEA–CRC and BBC combined with stepwise multiple regression</td>
<td>No. of hospital beds; physicians and non-specialist physicians</td>
<td>Bed utilization rate; bed turnover rate; total surgical operations; no. of births; total outpatient visits; average facility inpatient days, and no. of discharge</td>
<td>Structural quality, hospital efficiency, hospital size, average length of stay, % of physicians</td>
</tr>
<tr>
<td>Azevedo and Mateus (2014)</td>
<td>69 general acute care public hospitals in Portugal</td>
<td>Translog cost function, with a parametric difference-in-difference approach</td>
<td>Hospital Mergers</td>
<td>Total variable costs</td>
<td>Inpatient days, outpatient visits, emergency episodes, staff costs, staff, salary, beds, beds occupancy rate, number of operating beds, beds occupancy rate, GDP deflator, hospital case-mix index, type of hospital (district or county), time-years (pre-post hospital merges)</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Authors</th>
<th>Unit</th>
<th>Method</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lindlbauer and Schreyögg (2014)</td>
<td>1,239 acute care German hospitals</td>
<td>SFA in one-stage approach and DEA–BBC combined with bootstrapping regression</td>
<td>No. of beds, physicians, nurses and other staff members</td>
<td>No. of weighted inpatient cases</td>
<td>Ownership (private, non-profit, public), market concentration, % of leased beds from the total of beds, year, hospital specialization, mortality rate, occupancy rate, average length of stay per hospital</td>
</tr>
<tr>
<td>Santos et al. (2014)</td>
<td>1,097 Brazil’s Southeastern cities</td>
<td>DEA–BBC output oriented and Malmquist Index (MI)</td>
<td>No. of family health teams, health establishments providing primary care services</td>
<td>No. of people registered by the family health team, home visits of the family health team, ambulatory production in primary care</td>
<td>Market density in the outpatient sector; managed care penetration</td>
</tr>
<tr>
<td>Meyer (2015)</td>
<td>121 public hospitals in Switzerland</td>
<td>Stochastic cost frontier with Cobb-Douglas form, using the method of maximum likelihood</td>
<td>Total costs; price of labor; price of capital; no. of hospital wards and services; no. of internship categories; binary variable (1,0) for university hospital and for specialized surgical clinic</td>
<td>Inpatient discharges; outpatient revenue</td>
<td></td>
</tr>
<tr>
<td>Xu et al. (2015)</td>
<td>50 tertiary public hospitals in Beijing</td>
<td>Ratio analysis (RA), SFA, and DEA, with index screening and PCA</td>
<td>Medical personnel and equipment index (medical and non-medical staff, beds and equipment); construction index (floor space and the area of work place); financial index (total assets, subsidies, estimated revenue charges)</td>
<td>Hospital revenue; bed utilization; no. of outpatient admissions, hospital discharges, operations for inpatients, cured cases utilization rate, turnover times of actual bed, working days of hospital beds, etc.</td>
<td>Not applied</td>
</tr>
<tr>
<td>Narci et al. (2015)</td>
<td>1,224 public and private hospitals in Turkey</td>
<td>DEA combined with Tobit regression</td>
<td>No. of beds, part-time and FTE physicians, general practitioners; nurses and other staff</td>
<td>No. of discharged patients; outpatient visits; emergency cases; daycare patients; surgeries</td>
<td>Population/km² in a province; % of uninsured population in a province; physician/1,000 population in a province</td>
</tr>
<tr>
<td>Almeida and Cima (2015)</td>
<td>43 Portuguese public hospitals</td>
<td>Translog cost function</td>
<td>Annual total variable cost (average wage, Other costs)</td>
<td>Admissions, outpatient surgeries, average length of stay (days), visits to emergency services, outpatient emergency visits</td>
<td>Occupancy rate, no. of beds, medical-surgical emergency, polyvalent emergency, teaching hospital</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Authors</th>
<th>Unit</th>
<th>Method</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Widmer (2015)</td>
<td>90 public financed Swiss hospitals</td>
<td>Cost efficiency using standard frontier model (SFM) and random parameter frontier model (RPFM)</td>
<td>Labor input price; price of other production inputs; no. of beds; no. of internship categories; no. of specialties</td>
<td>Inpatient admissions; revenues from outpatient care</td>
<td>Types of prospective payment; global budget; payment per patient case; payment for diagnostic-related group</td>
</tr>
<tr>
<td>van Ineveld et al. (2016)</td>
<td>65 Dutch hospitals</td>
<td>DEA (BBC, CRC and FPR – fixed proportion ratio model, input oriented) combined with linear regression</td>
<td>No. of beds, operational expenses, FTE physicians, FTE non-physicians</td>
<td>No. of inpatient admissions, primary outpatient visits, day care treatment</td>
<td>% canceled elective surgeries, % inpatients having decubitus (ulcers)</td>
</tr>
<tr>
<td>Rezaei et al. (2016)</td>
<td>12 Kurdistan teaching hospital</td>
<td>Stochastic frontier analysis with Cobb-Douglas production function</td>
<td>No. of active beds, nurses, physicians, and other staff members</td>
<td>Inpatient admission</td>
<td>Not applied</td>
</tr>
</tbody>
</table>
criteria decision analysis techniques like TOPSIS are still limited. Within this scarce literature, recently Wanke et al. (2016) built a two-stage efficiency analysis model based on TOPSIS and artificial neural networks to assess the performance of Malaysian Islamic banks. The strength of such an approach lies in not only identifying the most efficient units of analysis in terms of similarity to the ideal solution, but also identifying the most critical contextual variables to improve efficiency. However, no previous study using a two-stage TOPSIS approach has been identified so far in the healthcare literature, representing an unexplored method.

2.1 Hospital performance determinants

Hospital performance studies have attracted growing interest from researchers, mainly focusing on the estimation of hospital efficiency levels and their variation in time, also investigating the determinants of efficiency regarding social, economic, demographic and organizational factors subjacent to the hospital’s operation (Colombi et al., 2017; Samut and Cafrì, 2016). Regarding Brazil’s hospital sector, however, the literature of efficiency determinants is still scarce. Araujo et al. (2013) approached this topic, but they focused on the efficiency determinants of private, for-profit hospitals and did not compare different characteristics at hospital, socioeconomic and demographic levels.

Aspects like hospital ownership and hospital size are frequently discussed in the literature and generate non-consensual evidence regarding their impact on efficiency. Some studies indicate that public hospitals perform better than private hospitals. The argument is that public hospitals face more intense resource limitations than private ones, leading to a greater focus on cost reduction and on the efficiency of input consumption (Hollingsworth, 2003, 2008; Tiemann and Schreyögg, 2012). In turn, the opposite result is also found, with public hospitals performing worse than private ones or with a higher number of public hospitals in a given geographical region worsening the overall performance in the area. This effect is associated with fewer investments in medical technologies, lack of competition and legal and bureaucratic restrictions hindering performance improvements (Colombi et al., 2017; Samut and Cafrì, 2016). Regarding hospital size, most studies in the literature indicate better performance by large-sized hospitals, when compared to their smaller peers. The explanation refers to a higher number of health procedures and service deliveries, leading to higher use and better exploitation of available resources, while small- and medium-size hospitals share their fixed costs across a relatively small pool of patients (Colombi et al., 2017; Ugá and López, 2007). However, after a certain size, often represented by the number of hospital beds, large units may jeopardize performance improvements, due to greater managerial complexity, often dealing with a less focused scope of operation (Giancotti et al., 2017).

GDP per capita has also been acknowledged as an explanatory variable in the healthcare performance literature and results indicate that it is positively associated with hospitals’ performance, since wealthier geographic areas present healthier individuals and improved health outcomes (Halkos and Tzeremes, 2011; Marinho 2003; Samut and Cafrì, 2016; Santos et al., 2014; Varela et al., 2012).

In contrast, the population size and territorial area variables do not present such unanimity regarding their effect on performance. While areas with higher population density lead a greater demand for hospital services – attracting more health resources and leading to increased competition – areas with a very large population and great geographical distances may present increased complexity in the management of the health demand and allocation of health resources, attracting residents of other regions (Allin et al., 2016; Marinho, 2003). Such capacity issues are commonly addressed in the health management literature through health-consortiums among hospitals or geographic areas (Deegan and Mathews-Schultz, 2016). In Brazil, where small municipalities have great
difficulty meeting their health demands, given the poor health infrastructure and the restrictions on revenue, the intermunicipal health consortium is an increasingly adopted alternative: municipalities located in contiguous geographic areas form a legal association to provide health services to their population within SUS (Mello et al., 2017).

3. Methods

3.1 TOPSIS method

TOPSIS is a widely accepted multiple criteria decision-making technique based on the ideal and negative ideal solution concepts. The first is the one that maximizes the benefit and minimizes the total costs, while the second one minimizes the benefit and maximizes the cost. The TOPSIS basic principle assumes that the chosen alternative should simultaneously have the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution (Hwang and Yoon, 1981).

TOPSIS employs analytical methods of Euclidean distance functions on normalized vectors of positive (outputs) and negative (inputs) criteria. The determination of weights shaping the importance of each criterion (benefits and costs or simply outputs and inputs) is a cornerstone step, being previously defined by the research decision making. Another characteristic is that TOPSIS does not impose any functional form on the data nor make assumptions on data convexity or the distribution of calculated scores (Ertaşğurul, and Karakaşğlu, 2009). Thus, TOPSIS allows the adoption of multiple variables and does not require prior specification of the production function, fitting the complexity of healthcare services (La Forgía and Couttolenc, 2009). The technique can be explained through its major analytic steps:

(1) An evaluation matrix consisting of m alternatives and n criteria is developed, with the intersection of each alternative and criteria given as \( x_{ij} \), therefore one obtains a matrix \((x_{ij})_{m \times n}\).

(2) The matrix \((x_{ij})_{m \times n}\) is normalized to from a regulated matrix \( R^* = (r_{ij}) \). In this study, the authors use the vector normalization method as demonstrated in the following equation:

\[
r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}, \quad i = 1, 2, \ldots m \text{ and } j = 1, 2, \ldots, n.
\]

(3) The weighted normalized decision matrix for performance assessment is calculated by the following equation:

\[
W = (w_{ij})_{m \times n} = (w_{ij} \cdot r_{ij})_{m \times n},
\]

where \( w_j \) is the weight given to the criteria \( j \) and \( \sum_{j=1}^{n} w_j = 1 \). In our study, six attributes have been given the same weight with a fair consideration.

(4) The worst alternative (the negative ideal assessment unit) \( A_a \) and the best alternative (the positive ideal assessment unit) \( A_b \) are determined by using the following equations:

\[
A_a = \{ \langle \min(w_{ij} | i = 1, 2, \ldots, m) | j \in f_+ \rangle, \langle \max(w_{ij} | i = 1, 2, \ldots, m) | j \in f_- \rangle \}
\]

\[
= \{ x_{ij} | j = 1, 2, \ldots, n \},
\]

(3)

\[
A_b = \{ \langle \max(w_{ij} | i = 1, 2, \ldots, m) | j \in f_+ \rangle, \langle \min(w_{ij} | i = 1, 2, \ldots, m) | j \in f_- \rangle \}
\]

\[
= \{ x_{ij} | j = 1, 2, \ldots, n \},
\]

(4)
where \( J^+ = \{ j | j \in \text{positive} \} \) and \( J^- = \{ j | j \in \text{negative} \} \), which are a set of positive (benefit) and negative (cost) attributes, respectively.

(5) The distance \( d_{ia} \) between the target alternative \( i \) and the worst condition \( A_a \) is calculated by Equation (5), and the distance \( d_{ib} \) between the alternative \( i \) and the best condition \( A_b \) by Equation (6):

\[
d_{ia} = \sqrt{\sum_{j=1}^{n} (w_j - x_{aj})^2}, \quad i = 1, 2, \ldots, m, \tag{5}
\]

\[
d_{ib} = \sqrt{\sum_{j=1}^{n} (w_j - x_{bj})^2}, \quad i = 1, 2, \ldots, m, \tag{6}
\]

where \( d_{ia} \) and \( d_{ib} \) are the Euclidean distance from the target alternative \( i \) to the worst and best conditions, respectively.

(6) The similarity of alternative \( i \) to the worst condition (the inefficient best conditions) is calculated by the following equation:

\[
S_i = d_{ia} / (d_{ia} + d_{ib}), \tag{7}
\]

where \( 0 \leq S_i \leq 1, \ i = 1, 2, \ldots, m; \ S_i = 0, \) if and only if the alternative solution has the worst condition; \( S_i = 1, \) if and only if the alternative solution has the best condition.

(7) The alternatives according to \( S_i \) are ranked, with a higher value of \( S_i \) indicating a better solution with respect to higher performance levels, allowing the subsequent assessment of the impact of contextual variables.

3.2 The data, performance rank and performance prediction

The input and output data were obtained from CNES (http://cnes2.datasus.gov.br/), covering the years 2008 to 2013. The TOPSIS alternatives consisted of each one of the 92 RJ municipalities for the six years covered in the study, forming a sample of 552 units of analysis.

The variables choice observed the previous literature and data availability. The inputs are: the number of surgical beds, clinical beds, equipment for life support, equipment for diagnostic imaging, physicians, nurses and the total number of healthcare professionals working for the SUS. The outputs are: the total number of diagnostic, clinical and surgical procedures performed in outpatient and inpatient settings, and the number of domiciliary medical consultations through the Family Health Program (PSF), that is a national strategy to provide primary care for defined populations in the domiciliary environment, allowing a broader understanding of the health/disease process, health promotion and prevention.

The clinical and surgery beds, as well as the life support and diagnostic imaging equipment, depict a basic infrastructure for the care of outpatient and inpatient patients, while healthcare professionals, such as physicians and nurses, portray the imperative workforce at any public or private hospital within any specialty. They enable the hospitals in each municipality to provide the health service. Given its importance, use-intensity and significant amount within hospital services, these variables are commonly present in the hospital performance literature. Likewise, medical procedures and consultations are often used in the related literature, comprising the main forms of outpatient and inpatients services (see Table II – literature review conducted by the authors). The relevance of these variables is also seen in Hollingsworth (2008) and Hussey et al. (2009), systematic literature reviews on efficiency measurement in the healthcare sector. Both reviews
indicate that most articles analyzed used health services, like health procedures, medical consultations and patient discharges, as output variables, and that the most common input types are equipment, hospital beds and the number of human resources, like nurses and physicians.

As presented in Table III, the contextual variables include demographic and socioeconomic indicators (population, area in square feet, GDP per capita), variables that characterize the health facilities' juridical nature (number of federal, state and municipal hospitals) and type of ownership (number of private hospitals). Regarding the reliability and validity issues, the selection of contextual variables for predicting efficiency scores observed the prescription in Moro et al. (2014). Given that we are working with a relatively small number of contextual variables, we selected them manually, by using the problem domain knowledge, i.e. by having a clear understanding of the qualitative importance of contextual variables that affect efficiency in the public health system under study. Their impact was measured one at a time and variables were incorporated based on changes of the response variables. Therefore, a one-dimensional sensitivity analysis was performed on the TOPSIS efficiency estimates for the best neural network model, as described in Cortez and Embrechts (2013).

In short, the input and output variables were used to compute the performance scores through TOPSIS in the municipalities throughout the years of analysis. Subsequently, these efficiency scores were linked to the respective contextual variables through neural networks, allowing the performance of sensitivity analyses to determine the most relevant variables. The analyses were performed using R software, with R packages “Topsis,” “nnet” and “Caret” being used, respectively, for the TOPSIS computations, for neural network computations and for extracting relative importance of contextual variables.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inputs – TOPSIS criteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical beds</td>
<td>0.00</td>
<td>8,220.00</td>
<td>135,829</td>
<td>757.69</td>
</tr>
<tr>
<td>Clinical beds</td>
<td>0.00</td>
<td>7,934.00</td>
<td>159,138</td>
<td>730.45</td>
</tr>
<tr>
<td>Equipment for life support</td>
<td>0.00</td>
<td>42,000.00</td>
<td>589.13</td>
<td>3,594.59</td>
</tr>
<tr>
<td>Equipment of image diagnosis</td>
<td>0.00</td>
<td>4,796.00</td>
<td>9,843</td>
<td>444.73</td>
</tr>
<tr>
<td>No. of nurses</td>
<td>2.00</td>
<td>10,015.00</td>
<td>141,998</td>
<td>790.00</td>
</tr>
<tr>
<td>No. of physicians</td>
<td>0.00</td>
<td>18,581.00</td>
<td>239,072</td>
<td>1,429.04</td>
</tr>
<tr>
<td>Total no. of professionals (SUS)</td>
<td>70.00</td>
<td>100,963.00</td>
<td>1,967,463</td>
<td>8,597.54</td>
</tr>
<tr>
<td><strong>Outputs – TOPSIS criteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic procedures (inpatient)</td>
<td>0.00</td>
<td>806.00</td>
<td>9,275</td>
<td>65.31</td>
</tr>
<tr>
<td>Diagnostic procedures (outpatient)</td>
<td>4,582.00</td>
<td>29,513,854.00</td>
<td>765,007,536</td>
<td>2,613,455.13</td>
</tr>
<tr>
<td>Clinical procedures (inpatient)</td>
<td>0.00</td>
<td>143,399.00</td>
<td>4,536,217</td>
<td>14,686.53</td>
</tr>
<tr>
<td>Clinical procedures (outpatient)</td>
<td>24,797.00</td>
<td>33,334,541.00</td>
<td>1,184,884,891</td>
<td>2,991,444.70</td>
</tr>
<tr>
<td>Surgical procedures (inpatient)</td>
<td>0.00</td>
<td>110,970.00</td>
<td>2,395,295</td>
<td>10,411.63</td>
</tr>
<tr>
<td>Surgical procedures (outpatient)</td>
<td>0.00</td>
<td>967,420.00</td>
<td>56,552,592</td>
<td>129,446.01</td>
</tr>
<tr>
<td>Patients attended (Family Health Program)</td>
<td>0.00</td>
<td>934,470.00</td>
<td>23,366,549</td>
<td>63,754.57</td>
</tr>
<tr>
<td><strong>Contextual characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area (km²)</td>
<td>7,653.00</td>
<td>4,026,696.00</td>
<td>455,283,348</td>
<td>497,586.04</td>
</tr>
<tr>
<td>Municipal hospitals</td>
<td>4.00</td>
<td>316.00</td>
<td>33,69</td>
<td>37.41</td>
</tr>
<tr>
<td>State hospitals</td>
<td>0.00</td>
<td>109.00</td>
<td>1,361</td>
<td>8.97</td>
</tr>
<tr>
<td>Federal hospitals</td>
<td>0.00</td>
<td>58.00</td>
<td>0,596</td>
<td>5.03</td>
</tr>
<tr>
<td>Private hospitals</td>
<td>0.00</td>
<td>4,139.00</td>
<td>120,828</td>
<td>364.53</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>5,124.49</td>
<td>299,321.57</td>
<td>23,775,916</td>
<td>34,250.57</td>
</tr>
<tr>
<td>Population</td>
<td>5,269.00</td>
<td>6,429,922.00</td>
<td>174,974,453</td>
<td>668,268.46</td>
</tr>
</tbody>
</table>

Table III. Descriptive statistics for the TOPSIS criteria and the contextual variables.
3.3 Predicting performance levels using neural networks

Artificial Neural Networks are learning models that combine mathematical and artificial intelligence techniques for prediction purposes. They reproduce the behavior of biological neural networks, learning from experience how to distinguish underlying patterns (Zou et al., 2008).

Given their ability to identify complex nonlinear interactions among predictors and a health-related predicted variable, neural networks have been effectively applied in the healthcare field (Baxt, 1995; Eller-Vainicher et al., 2011). They are used to predict variables such as hospital length of stay (Launay et al., 2015; Tsai et al., 2016), in-hospital mortality after a surgical procedure (Shi et al., 2013; Wise et al., 2015), hospital emergency admission (Handly et al., 2015; Leegon et al., 2006) and care charges related to specific cancer types (Lee et al., 2004; Wang et al., 2009). Predictive models in healthcare are also integrated with computerized systems and used as decision tools (Kaplan, 2001), helping to ground the judgment of physicians and hospital managers (Marlais et al., 2011).

Models for systematically predicting performance in health facilities present extensive management value. They allow for anticipating low service quality and narrowing its negative outcomes – as deterioration of patients’ health status, material waste, patient dissatisfaction and ineffective treatments. Nevertheless, the absence of predictive analysis in the healthcare performance literature reviewed indicates a potential gap.

Lastly, it is worth observing a trade-off traditionally faced by predictive models: a high level of accuracy implying increased complexity of interpretation. One way to face this is to prioritize the prediction accuracy, since, while a validated prediction model may be highly useful, even when it comes to a difficult interpretation, there is little use in a predictive model of easy interpretation but no statistical validation (Kuhn and Johnson, 2013). The predictive model of this paper meets Kuhn and Johnson’s (2013) perspective.

4. Results

The TOPSIS results provide performance scores for all 92 municipalities in the selected years, summarized in Table IV and Figure 1 – TOPSIS performance grouped by year. The performance scores show an overall average of 0.506, with little variation from 2008 to 2011 and a slight improvement after 2011. From 2008 to 2011, about 60 percent of municipalities obtained performance scores below 0.5; in 2012 and 2013 this number dropped to nearly 28 percent of municipalities and no municipality is deemed efficient in any of the years analyzed.

The maximum and minimum scores, along with the standard deviation values, indicate a significant variation among municipalities (Table IV). Comendador Levy Gasparian is first in the rankings, with a score of 0.649 in 2012, while Mesquita is in the last position, with a score of 0.382 in 2008, representing an inefficiency level of 0.618 (1–0.382). Mesquita remained in the last 20 positions from 2008 to 2011 (average score of 0.42), and Comendador Levy Gasparian stood in the first 25 positions from 2008 to 2013 (average score of 0.62).

Regarding the neural network analysis, Figure 2 exhibits the apparent root mean squared error (RMSE) grouped by the number of hidden layers in the neural network. The graphic shows that the RMSE tends to decline with the increase of hidden layers,

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.506</td>
<td>0.494</td>
<td>0.491</td>
<td>0.490</td>
<td>0.492</td>
<td>0.532</td>
<td>0.533</td>
</tr>
<tr>
<td>Median</td>
<td>0.504</td>
<td>0.495</td>
<td>0.491</td>
<td>0.494</td>
<td>0.496</td>
<td>0.534</td>
<td>0.544</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.649</td>
<td>0.629</td>
<td>0.622</td>
<td>0.610</td>
<td>0.613</td>
<td>0.649</td>
<td>0.647</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.382</td>
<td>0.382</td>
<td>0.389</td>
<td>0.389</td>
<td>0.385</td>
<td>0.399</td>
<td>0.398</td>
</tr>
<tr>
<td>SD</td>
<td>0.055</td>
<td>0.051</td>
<td>0.051</td>
<td>0.049</td>
<td>0.048</td>
<td>0.054</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Table IV.
Descriptive statistics for performance scores
suggesting a response bias toward a large number of hidden layers. The RMSE assessment across different numbers of hidden layers helps to outline a portrait of the response bias or the overfitting within each predictive technique. In contrast, Figure 3 shows that a common pattern within the cross-validation methods is seen where the RMSE peaks at lower hidden
layer values, decays and stays roughly constant within the window between 0.045 and 0.040. The most accurate neural network obtained (RMSE = 0.038) used the bootstrap 632 technique, for 15 hidden layers and a decay rate of 0.0090.

Figure 4 indicates the relative importance of each contextual variable as a predictor of TOPSIS performance scores. The most important are the variables related to the demography and socioeconomic status of municipalities (area in square feet and GDP per capita), and to the juridical nature and type of ownership of the healthcare facilities in each
municipality (number of federal and private hospitals). Figure 5 presents the impact direction of these variables on performance, showing a sensitivity analysis of TOPSIS scores. As recommended by Faraway (2005), the contextual variables were standardized.

The territorial area (first graph in the first row – Figure 5) presents the greatest negative relative impact on municipalities’ performance scores, followed by the number of federal hospitals (first graph in the second row – Figure 5). This means that the larger the area of the municipality and the higher the number of federal hospitals, the lower its performance. Similarly, the number of municipal and state hospitals (second and third graphs in the first row of Figure 5) and the population size (graph in the third row – Figure 5) has a negative impact on municipalities’ performance scores. In contrast, the number of private hospitals and the GDP per capita (second and third graphs in the second row – Figure 5) produce a positive effect on performance, with the former having a relative impact greater than the GDP per capita on performance.

5. Discussion and conclusion
This paper applied a two-stage TOPSIS and neural network techniques to analyze the performance of Brazilian hospitals in RJ municipalities, to estimate the change in performance between 2008 and 2013 and to reveal the relative importance of socioeconomic, demographic and institutional variables as predictors of the performance scores.

Regarding the hospitals performance, the results reveal a low average performance score of 0.506 and a significant variation among municipalities’ performance scores.
Therefore, the benchmarking approach shows that there is substantial room for improvement, and it is feasible to increase the number of outputs achieved without necessarily increasing the consumption of resources. This may be chased through the analysis of the production process, and investigation of the best practices used in the municipalities considered as a reference.

This result is aligned with other academic studies that analyzed the performance of Brazilian public health services at the municipality level, applying DEA (Marinho, 2003; Santos et al., 2014; Varela et al., 2012). Despite the differences among the studies, including variables, method, year and municipalities, they all point out inefficiencies in the performance of Brazilian public health services.

Concerning the performance change over the years, the results indicate little variation from 2008 to 2011 and a slight improvement from 2011 to 2012 and 2013. This slight increase may be related to the BHP, launched by the Ministry of Health in 2011 to reduce costs and the demand for hospital services. However, even with this slight improvement in performance, the average scores remain low (around 0.53), indicating much room for improving the efficiency of public health services in RJ. This result is consistent with the findings of Santos et al. (2014), who assessed the performance of public primary care services in Brazilian municipalities located in the Southeastern Region, including RJ and SP states, from 2007 to 2010, and found an overall average score of 0.61.

Regarding contextual variables, this study analyzed socioeconomic (GDP per capita), demographic (area in square feet; population) and institutional variables (facilities juridical nature – number of federal, state and municipal hospitals) as predictors of performance scores. Consistent with the findings of Allin et al. (2016), this research confirms that areas with a very large population and great geographical distances present increased complexity in the management of health demands and allocation of health resources, since the results indicate a negative relationship between these demographic variables and the municipalities’ performance. Also, this study indicates that territorial area is the variable that has the greatest negative impact on municipalities’ performance. One possible reason is that bigger municipalities in Brazil usually serve as a support channel for the smallest and poorest ones, compromising the performance due to the attraction of severe cases from other locations and this transfer process may be an importation mechanism of external inefficiencies (Marinho, 2003). This phenomenon occurs because small municipalities have constant difficulties in meeting their health demands in proportion to their financial income and often operate at a production scale incompatible with its population demand, leading to capacity problems, as indicated by Deegan and Mathews-Schultz (2016) and Mello et al. (2017).

In relation to the socioeconomic variable, the results indicate a positive relationship between GDP per capita and the municipalities’ performance, consistent with previous studies (Halkos and Tzeremes, 2011; Marinho, 2003; Samut and Cafri, 2016; Santos et al., 2014; Varela et al., 2012), reinforcing the understanding that wealthier geographic areas have healthier individuals and improved health outcomes.

Concerning the facilities’ juridical nature, the results reveal a negative relationship between the number of public hospitals (federal, municipal and state hospitals) and the performance of the municipalities. In turn, the number of private hospitals is positively related to the performance of the municipalities. This result strengthens the argument that a higher number of public hospitals in a given geographical region worsen the overall performance of the area and that private hospitals have fewer bureaucratic and political requirements than their public counterparts, as advocated by Colombi et al. (2017) and Samut and Cafri (2016), among others.

It is also worth noting that RJ is one of the Brazilian states with the greatest concentration of public hospitals. Further, over 60 percent of the public hospitals are of small size – with a maximum of 50 beds – suggesting a suboptimal operation volume (Ugá and López, 2007). This may help understanding the low average performance levels attained and the negative
impact of a higher number of public hospitals on municipalities’ performance. That is, possibly the inefficiency is not in the scarcity of public hospitals, but in the scale of operation at the existing ones. Different aspects associate economies of scale and scope with higher performance and better outcomes in health services: the learning curve resulting from increased volumes of activity may lead to higher productivity and less wastage of resources, and bigger hospitals may decrease the proportion of fixed costs per health service and increase the bargaining power with suppliers (Lindlbauer and Schreyögg, 2014; Narci et al., 2015; Rezaei et al., 2016). Also, in Brazil a significant number of private hospitals sign agreements with the SUS to provide beds and assistance to the population, decreasing the demand pressure on public services, as mentioned by Araujo et al. (2013).

In short, the results suggest that municipalities with larger territorial areas and bigger population sizes attained worse hospital performance levels, as did those municipalities with a greater number of public hospitals. In contrast, the number of private hospitals and the GDP per capita in the municipalities seem to affect hospitals’ performance levels positively. The socioeconomic and demographic variables analyzed relate to the magnitude of infrastructure and demand for health services, to economies of scale and to the availability of economic resources, while the facilities’ juridical nature relates to the form of hospital management, influencing management practices, forms of funding and managerial autonomy.

For public health managers, this study shows that there is substantial room to increase the outputs in public health services, improving the results for the population, without increasing the consumption of resources. In addition, for policymakers, these results emphasize that health policies and performance analyses should be carried out with caution and in a broad perspective, considering the demographic and socioeconomic context in which such policies are embedded. It is not enough to investigate the resources available, or try to maximize the results obtained: the performance of public health services transcends the adequate use of resources.

Besides the practical relevance of empirically discussing health services’ performance, this study benefits the academic community by enriching the health management literature and encouraging new theoretical and empirical debates on the topic. Also, this study displays an innovative context for applying two-stage TOPSIS and neural networks, no similar effort having been identified in the healthcare literature so far. The dissemination of TOPSIS applications and its variants may help to improve the method and bring a new and academically valuable tool for performance measurement in the healthcare literature, which currently relies on a limited number of methods.

Despite the practical and academic insights and methodological advances highlighted, the findings of this study should be interpreted with prudence. Although robust and widely used, TOPSIS has some inherent shortcomings. The technique measures relative performance, meaning that municipalities are compared to others in the sample, not to external or international standards. Additionally, the results’ accuracy depends greatly on the variables used. Omitted variables can lead to measurement errors, if relevant aspects of the production process are not properly covered. To minimize this limitation, we selected variables previously used in the literature. Moreover, given the importance of management aspects for public health services, we believe that the benefits obtained by the systematic application and cautious interpretation of this performance analysis overcome such limitations.

To deepen this analysis, future research could further the performance analysis by including municipalities of other Brazilian states, enlarging the scope of the benchmark and searching for national health references that could stimulate the dissemination of good operational practices. Another opportunity would be the further investigation of performance predictors, covering other socioeconomic and infrastructure dimensions of hospitals’ operation. A longitudinal assessment contemplating longer periods would also facilitate monitoring results in response to specific public health initiatives, such as educational campaigns, quality programs for hospital accreditation and professional training.
References


About the authors
Claudia Affonso Silva Araujo is Associate Professor of COPPEAD/UFRJ, Area Head of Operations, Technology and Logistics of COPPEAD/UFRJ and Researcher of the Center for Studies in Management of Healthcare – CESS. Claudia received Doctor of Sciences in Business Administration (COPPEAD/UFRJ), Master degree in Business Administration (COPPEAD/UFRJ) and is graduate in Economics from the Catholic University of Rio de Janeiro – PUC/RJ. Claudia was Visiting Researcher at Maryland University (USA) in 2004.

Peter Wanke received Doctor of Sciences in Product Engineering from Engineering Program of Federal University of Rio De Janeiro (COPPE/UFRJ) and is Visiting Scholar of the Marketing and Logistics Department of Ohio State University. Peter received Master degree in Product Engineering from COPPE/UFRJ and Bachelor degree in Product Engineering from UFRJ. Peter is Adjunct Professor at COPPEAD and Coordinator of the Center for Studies in Logistics. He works in teaching, research and consultancy in the fields of logistics and supply chain and has more than 60 papers published in journals such as International Journal of Physical Distribution and Logistics Management and International Journal of Operations and Production Management.

Marina Martins Siqueira is PhD Candidate at COPPEAD/UFRJ and Researcher of the Center for Studies in Management of Healthcare – CESS. Marina received Master degree in Business Administration (COPPEAD/UFRJ) and is graduate in Administration from the Federal University of Piauí – UFRJ. Marina is Visiting Scholar (USA) at Duke University (2017–2018). Marina Martins Siqueira is the corresponding author and can be contacted at: marina.martins.siqueira01@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
The relationship between cost stickiness and financial reporting quality in Tehran Stock Exchange

Mahdi Salehi  
Ferdowsi University of Mashhad, Mashhad, Iran

Nasrin Ziba  
Faculty of Accountancy and Management, Qaenat Branch,  
Islamic Azad University, Qaenat, Iran, and

Ali Daemi Gah  
Imamreza International University of Mashhad, Mashhad, Iran

Abstract

Purpose – The purpose of this paper is to investigate the relationship between financial reporting and cost stickiness in companies listed on the Tehran Stock Exchange.

Design/methodology/approach – Data of all Iranian manufacturing listed companies gathered for testing hypotheses during 2010-2016 and R statistical software are employed in order to analyzing data.

Findings – The results of this study indicate that there is a significant relationship between administrative, sale, material, labor and overhead costs and the financial reporting qualities of the companies under study.

Originality/value – The study focuses on relationship between financial reporting and cost stickiness in companies listed on the Tehran Stock Exchange, which is the first study of its type in Iran.

Keywords Financial reporting quality, Distribution and sales cost stickiness, Administrative cost stickiness, Cost stickiness, Material-labor-overhead

Paper type Research paper

1. Introduction

Today, accounting information systems play a very important role in operational circles of organizations; they also play an important role in economic environment of the countries. Many of economic decisions are significantly made based on extracted information from these reporting systems, and a major degree of securities and bounds exchanges are dedicated for buying and selling companies’ stocks, which, in turn, can be influenced by the figures and accounting information. Any research about how accounting information influences the wide range of stakeholders’ decision about the companies will help having a better understanding about role of this information and the need for more and higher disclosure (Alfred, 2014). The predictive role of the cost structure at the firm level is studied in the managerial accounting literature (Banker and Chen, 2006; Weiss, 2010). However, the role of cost behavior in predicting the microeconomic indicators at firm level is paid little attention. In this paper, we explore the effect of cost stickiness in financial statements on financial reporting quality.

Financial reports are known as final outcome of the reporting process. The financial reporting process involves the creation, publication, reassurance and application of financial information by the different users. In other words, the scope of financial reporting begins from the implementation of the financial reporting standards for providing financial statements, which is extended to assurance, publication and their usage. Information users in order to make decision in capital market need useful information; they also need such data in order to find out the contractual context and the level of management’s presidency. Information quality refers to the content of financial disclosure, which means how much financial reports present the company’s economical infrastructure fairly (DeFond and Zhang, 2014). The more explicit figures
in financial statements of companies, the more transparency will be achieved. The quality of financial reports is met by increasing transparency and also publishing high-quality annual reports through comprehensive disclosure. The quality of financial reports has always been an interesting issue for board directors, shareholders, researchers and professional accountants.

Recently, the great scandals in international financial communities raise more questions and concerns about the quality of financial reporting (Brown et al., 2010). Previous studies (Healy and Palepu, 2001; Bushman and Smith, 2013; Lambert et al., 2007) point out that high-quality financial reporting can provide important economic outcomes, including the increase in investment efficiency. Managerial accountants focus on cost behavior in their earnings’ analyses and presentations which are provided for managers. Financial analysts anticipate costs in order to predict future earnings. Therefore, cost behavior is an important issue for both internal and external decision making (Chen et al., 2012). The cognition of cost behavior is one of the most important issues of cost and management accounting. In traditional models, the cost behavior varies with the change in cost drivers; moreover, the directions of changes (increase or decrease) in the volume of activity do not have any significant effect on the magnitude of costs changes. But the results of some researchers (Calleja et al., 2006) in recent years indicate that an increase in costs is due to a decrease in the volume of firm’s activities. This behavior is named cost stickiness.

Following research aimed to investigate the relationship between financial reporting quality and cost stickiness. Although many studies investigate the quality of financial reporting and cost stickiness, an empirical study that examines the relationship between financial reporting quality and cost stickiness is not applied. Therefore, the main objective of this study is to investigate the relationship between financial reporting quality and cost stickiness. In Section 2, we will discuss about the literature of the study. Section 3 explains the research methodology. The research findings are presented in Section 4. Finally, the results and suggestions are presented in Section 5.

2. Theoretical framework

When the cost stickiness phenomenon is raised, some experts refer it as an alternative to cost behavior (Anderson et al., 2003). In accounting literature, cost stickiness is the difference between costs which will change automatically with changing the volume of activities and costs which are occurred by management through resource consuming. The cost sticky behavior of management refers to managers’ decisions which are made deliberately to use resources of the organization inappropriately. In other worlds, inappropriate usage of resources by the organization arises due to the reduction in the volume of firm’s activities (Venieris et al., 2015). Cost sticky behavior of management can take place in different classifications such as manufacturing and trading costs (Chen et al., 2012; Kama and Weiss, 2013).

2.1 Creative accounting

In these days, the quality of financial reporting is given more attention by business practitioners. One of the most important challenges in financial reporting is creative accounting phenomenon. When managers employ certain accounting procedures for better monitoring rather than provide a fair picture of financial condition for information users, it leads to an incorrect presentation of the performance and position of the entity which is called “creative accounting” (Peter and Mclaney, 2011). Mathew (2006) believes that there is no generally accepted definition of creative accounting. However, the word creative has a negative meaning and creative accounting is used in case of misleading investors, creditors and other financial information users. This type of accounting almost appears as booking fraud, fraudulent accountancy, fictitious reporting and account arrangement (Mathew, 2006).
Many studies are conducted on the financial reporting quality issues (Biddle et al., 2009; Chen et al., 2011; Gomariz and Ballesta, 2013), and based on their findings, increasing the quality of financial reporting is a divisional decision. Several studies indicate that companies can reduce information asymmetry by increasing the quality of financial reporting. Chen et al. (2011) imply to the effects of incremental influences of reducing reversal and ethical risks and allowing managers to identify top investment opportunities as well as higher financial reporting quality about investment efficiency. Also, many studies examine the costs elements related to financial reporting quality, for example, lower costs, longer debt maturities and lower security cost for bank financing. High-quality financial reporting enables managers to be more accountable throughout providing better presidency; furthermore, possibly, it may reduce the risk of information asymmetry and subsequently reverse selection and moral hazard. It can also solve the under/over investment problems. On the other hand, the quality of financial reporting can also improve the investment efficiency by raising the possibility of better decision making about better investments for managers and better recognizing the projects and providing fairer accounting figures for domestic decision makers (Gomariz and Ballesta, 2013). Chen et al. (2011), using data from emerging market’s companies examine the impact of financial reporting quality on under/over investment, conclude that higher financial reporting quality helps companies to have more investment with insufficient investment opportunities and also helps them to lower their investment level with over investment problems.

2.2 The concept of cost stickiness
Cost stickiness reflects the economic asymmetry in cost response to increase and decrease in sales (Anderson et al., 2003). In other words, cost stickiness means that increasing in costs at the time of sales are increasing is more than reducing the same amount of costs when sales are reduced. For example, if sales revenue increases by 15 percent, costs will increase by 10 percent, but if the same sales revenue is reduced by 15 percent, costs will fall less than 10 percent. In such situation, the behavior of costs is sticky (Weiss, 2010). The cost stickiness is defined by manager’s decision to bear the cost of unused resources in sales decrease situation. Companies with cost stickiness are divided into two groups of efficient companies (with positive news) and inefficient companies (with negative news). Companies that have efficient stickiness are appointed to the group which current sales have fallen, but their sales are expected to be raised up in near future. Therefore, in this regard, these companies are efficient because of bearing additional costs (unused resources), by avoiding resource adjustment costs (decreasing and increasing again), so it will generate benefits for the company’s future. Companies with inefficient cost stickiness refer to the group of companies which current sales have fallen, but their fallen sales are expected to be permanent (long-term); therefore, the companies are inefficient in terms of bearing additional costs (unused resources) which will not generate benefits for the company’s present and future. Consequently, regardless of whether the company belongs to the efficient or inefficient cost stickiness subgroup, it has a negative impact on the current profit (decline in sales year-by-year), because the decline in sales is not compensated by the same amount of reduction in costs. But the cost stickiness of efficient companies will have a positive impact on future profits (due to a rebound in sales as positive news), and inefficient firms have a negative impact on future profits (due to the permanent decline in sales as negative news) (Homburg and Nasev, 2010).

2.3 Cost stickiness reasons
For the explanation of cost stickiness occurrence, the below items are presented:

(1) the existence of adjustment cost;

(2) change in amount of economic activity;
The first factor relates to resource adjustment cost are categorized as hiring and firing cost for labor or installation and disposal cost for equipment. Specially, more significance in adjustment cost leads to more noticeable cost stickiness because the firm’s behavior under optimal decision making is asymmetric; in this regard, in case of decreasing sales, managers are indifferent to fire employees up because retaining the unused resources can help managers to avoid large current costs as well as future costs of training new employees in the future. Vice versa, when the sales increase, managers may be indifferent to hiring more workers because of the adjustment costs: the increase in current sales can only be achieved if additional workers are hired; thus, the effect of indifference is likely to be getting less. Therefore, the high level of adjustment cost prevents managers from deciding to reduce the resource amounts proportionally to the decline in the level of economic activity (Banker and Byzalov, 2014). Balakrishnan and Gruca (2008) suggest that adjustment costs (and, accordingly, the degree of cost stickiness) are higher in core services (patient care) than in support services in hospitals, and also find results supporting the adjustment cost explanation for sticky costs.

The second factor in cost stickiness generation is making a change in economic activity. The relatively large changes in sales will distort the linear pattern of cost behavior. In this regard, Krishnan et al. (2007) find that at the level of activity managers do not consider small changes in working hours of labors (3 percent or less) as an important issue. In contrast, they usually show significant response to major changes. The investigation of Saberamaniam and Wiedemmer shows that income-based changes which is more than 10 percent will show stickiness in administrative, general and sale costs and also in the cost of sold goods.

The third affecting factor on the cost stickiness is the manager expectation of predicted sales which is driven by future demand for the firms’ products. It can influence informed decision making about resource adjustment and cost behavior. The optimistic (pessimistic) expectations of a constant decline in sales lead to an increase (decrease) in cost stickiness (Banker and Byzalov, 2014). The probable reason that can lead to a relationship between managers’ expectations of sales and cost stickiness is due to the manager’s points of view about sustainability of demand’s change. In case of optimism about growth demand in future, managers expect a higher level of future resource requirements. Therefore, they are supposed to keep unused resources even in case of experiencing decreases in current sales; such activities provide them better situation to reduce current and future adjustment costs during the predicted future resource expansion. On the other hand, in case of pessimism about growth demand in future, managers are supposed to fire extra employees when meet decline in sales rate and are also cautious about hiring additional employees if they face current increases in demand. Thus, managerial optimism alleviates cost response to decreases in current sales and strengthens the cost response to the increases in current sales, thereby resulting in increased cost stickiness.

The fourth affecting factor on cost stickiness is agency costs (Anderson et al., 2003; Chen et al., 2012), which refers to managers’ trends to grow the firm according to its optimal size or to maintain unused resources with the purpose of increasing personal utility from status, power, compensation and prestige. In other words, agency problem’s effect on cost stickiness is decision makers’ decision to maintain unfavorable resources which is derived by personal considerations. In this regard, Chen et al. (2012) refer to ambitious behavior of managers as an important factor in cost stickiness behavior.

2.4 Construction costs stickiness
Construction costs include manufacturing costs, such as materials, labor and overhead costs of production. The issue of unemployment rate is one of the most important economic...
matters which is highly considered by researchers in Iran and other countries. The publication of Keynesian’s general theory suggests that the nominal wages stickiness is the reason for unemployment rate. In fact, nominal and real wages are quite insensitive to production fluctuations (Huang et al., 1998). Involuntarily unemployment generates a sustainable labor market for the future, and the existence of such unemployment rate raises the question: why wages do not fall in the labor market, while involuntarily unemployment can be explained as a phenomenon of equilibrium? (Shapiro and Stiglitz, 1984). Amongst the theories that explain the stickiness of wages and involuntary unemployment is the wages efficiency theory that assumes that the positive labor productivity depends on the real wages rates (Huang et al., 1998). The efficiency wages theory suggests that from employers’ point of view the effort or efficiency (and therefore the productivity) of workers are not independent from their wages and there is an interaction between these two factors, and therefore they are not inclined to reduce their workers’ wages. The wages efficiency theory is explained by Yellen (1984), Katz (1986) and Weiss (2010). A theoretical foundation for stickiness of labor about the idea of wage efficiency is properly defined; it suggests that higher wages would increase the efficiency of workers and would increase the effectiveness of the workforces. The study of Balakrishnan et al. (2011), in which panel data are applied for analysis, suggests that real employment and wages have a positive effect on labor productivity. Also, the results of Mahlberg et al. (2012) and Tamasauskiene and Stankaityte (2013) suggest that there is a positive correlation between the share of older employees and productivity, while there is no evidence that there is a significant relationship between the share of well-trained employees and productivity, and the difference in wages is higher than labor productivity.

The study of Soderstrom and Noreen (1997) on the behavior of hospital overhead costs suggests that for a more accurate prediction of cost stickiness, it is better to assume that costs will not change proportionally with change in the level of activity. In another study, Krishnan et al. examine the validity of relative costs pattern in hospitals and the results show that by increasing the capacity utilization, the variable cost of nursing unit increases and managers increase the use of capacity in order to compensate the high wage rates. There are also differences in cost behavior due to the different ownership of hospitals. In another related study, Balakrishnan et al. (2011), by the title of “cost structure and cost stickiness,” provide evidence that, in the long run, decision making about cost instruction is affected by our ability to identify short-term decision making about cost management.

3. Related literature and hypotheses development

Recent literatures including Cook et al. (2017) imply that prior research provides evidence that a firm’s sales and general administrative expense is negatively correlated to its capital structure and earning quality, and yet positively correlated with its stock returns. They suggest that these results are driven by the effect of a firm’s operating leases on its asset volatility. Using Merton’s model, they extend these effects to the pricing of corporate debt. As a result, not only they provide an overlooked rationale for the above correlations, but also a different way of understanding the effects of similar difficult to terminate contracts (e.g. union contracts) on the stock and bond prices of a firm. Banker et al. (2016) suggest that a decrease in sales shows more effect on earnings than sales increase which is a consequence of cost stickiness; in this regard, they hypothesize that the correlated omitted variable constitutes a confounding effect in standard asymmetric timeliness models. Controlling for a piecewise linear effect of sales changes in these models decreases the measured asymmetric timeliness significantly and changes inferences about the average level of conservatism and the extent of cross-sectional variation in conservatism. Validation tests confirm that the asymmetry for sales changes is consistent with sticky costs and is distinct from conditional conservatism. Ibrahim (2015) examines whether costs respond asymmetrically to demand change, and
examines the influence of economic growth on cost stickiness, in the pre- and post-2008 financial crisis periods in Egypt. He uses multiple regression models to investigate the behavior of three costs: selling, general and administrative (SG & A), cost of goods sold (COGS) and operating costs (OCs) for the 2004–2011 period. Moreover, he compares cost stickiness during the economic prosperity period (2006–2008) with cost stickiness during the economic recession period (2009–2011). His study results reveal that SG & A increased by 0.38 percent but decreased by 0.08 percent, and COGS increased by 1.02 percent but decreased by 0.57 percent for a 1 percent demand change, which proves cost stickiness. However, OC increased by 0.91 percent, but decreased by 1.03 percent for a 1 percent demand change, which proves cost anti-stickiness. COGSs were sticky in both periods; however, the extent of cost stickiness is larger in the prosperity period. In contrast, OCs were statistically insignificant in both periods. Banker and Byzalov (2014) formulate an integrated framework of asymmetric cost behavior and review the empirical evidence in support of this framework and its implications for both cost and financial accounting research. They clarify empirical issues and show that recent contrary claims about the validity of findings in the literature are unwarranted because of econometric errors. They present new comprehensive evidence from Global Compustat, which demonstrates that asymmetric cost behavior is a pervasive global phenomenon. Mahlberg et al. (2013) study the relationship between the age structure of employees, labor productivity and wages. Their findings indicate that firm productivity is not negatively related to the share of older employees it employs. They also suggest no evidence for overpayment of older employees and the results do not show any association between wages and the share of older employees. Furthermore, they find a negative relationship between the share of young employees and labor productivity as well as wages, which is more prevalent in the industry and construction sector. Yasukata (2011) use sales forecasts issued by managers as a proxy for managers’ prospect of future sales and examine the hypothesis that costs stickiness is the results of the deliberate decision of managers by empirically testing whether the managers’ sales forecasts have an impact on costs stickiness. They find that the prospect of future sales is related to current level of costs stickiness. They document that strong evidence that costs stickiness is the results of deliberate decision of managers. Aktas and Kargin (2011) design a study to investigate the effects of sector, reporting type and income on firms’ timely annual financial reporting practices listed on Istanbul Stock Exchange. They examine the effects of sector (financial and non-financial firms), financial statement type (consolidated-non-consolidated firms) and income (positive-negative income) for the years from 2005–2008. Their results reveal that sector, financial statement type and income have significant impact on timely reporting financial statements of selected firms. The coefficient estimates for sector, financial statement type and income are statistically significant. Effects of sector and financial statement type on lead time are positive while income is negative. Based on the results, non-financial firms publish their financial statements later than others. Similarly, consolidated firms report their financial statements later than non-consolidated firms. Finally, firms that report positive income release financial statements earlier than others. Homburg and Nasev (2010) interpret cost stickiness, i.e. the manager’s decision to bear the costs of unutilized resources when the sales decline, as a risky project and examine its impact on conditional conservatism. They show that cost stickiness increases the asymmetric timeliness of earnings by weakening the timeliness of earnings for good news firms and, at the same time, intensifying the timeliness of earnings for bad news firms. Additionally, their findings suggest that the asymmetric timeliness of earnings for cost sticky firms is more strongly driven through accounting factors, as reflected in accruals than through non-accounting factors, as reflected in cash flow. Our results imply that cost stickiness is costlier due to conditional conservatism and that the market separates the efficient from the inefficient cost sticky firms, indicating that information asymmetry is low. Weiss (2010) examines how firms’ asymmetric cost behavior influences analysts’ earnings
forecasts, primarily the accuracy of analysts’ consensus earnings forecasts. His results indicate that firms with stickier cost behavior have less accurate analysts’ earnings forecasts than firms with less sticky cost behavior. Furthermore, he finds that cost stickiness influences analysts’ coverage priorities and investors appear to consider sticky cost behavior in forming their beliefs about the value of firms.

On the other hand, in early studies by using data from hospitals in the state of Washington, Soderstrom and Noreen (1997) examine the time-series behavior of overhead costs. They find that more accurate predictions of changes in costs are usually generated by assuming a cost will not change at all (except for inflation) rather than assuming that the cost will change in proportion to changes in activity. They also document that nearly all of the effect of a change in activity on costs appears to occur in the same year as the change in activity. Finally, using a multi-period regression model, they find that the proportion of variable costs in the hospital overhead accounts is apparently very modest. These results suggest that costing systems, such as activity-based costing, where assume costs are proportional to activity, will grossly overstate relevant (i.e. incremental) overhead costs for decision making and performance evaluation purposes. In another investigation, Beatty and Harris (1999) compare public and private banks’ realizations of securities gains and losses to determine how their earnings management differs. They find that public banks consistently engage in more earnings management than private banks, and that the portion of their current period securities gains and losses attributable to earnings management is more positively associated with next period’s earnings before securities gains and losses. Their findings are consistent with earnings management occurring due to greater information asymmetry in public firms, and suggest that earnings management may not necessarily lead to the erosion in the quality of earnings.

According to literature review and above discussion, we developed our hypotheses as follows:

H1. Administrative cost stickiness (ACA) has significant impact on financial reporting quality.

H2. Sale cost stickiness has significant impact on financial reporting quality.

H3. Material cost stickiness (MCA) has significant impact on financial reporting quality.

H4. Labor cost stickiness has significant impact on financial reporting quality.

H5. Overhead cost stickiness has significant impact on financial reporting quality.

In this paper, we investigate the relationship between cost stickiness and firms’ financial reporting quality in the Tehran Stock Exchange. Our choice is dedicated by several reasons. Firstly, the growing interest by Iranian law makers related to financial reporting quality, motivate us to explore the economic consequences of cost stickiness in this context. Furthermore, the absence of International Financial Reporting Standards imposed in Iranian accounting system motivates us to study the different influential factors on financial quality, which is expected to help us to explore the reasons for applying IFRSs in Iran economy. It is required that Iranian firms provide financial statements in accordance with the four qualitative characteristics of financial information, namely, reliability, relevance, intelligibility and comparability, to enhance its credibility to the various users of financial statements in which cost stickiness has a very effective role in these qualitative characteristics. Besides that, investment is the cornerstone of value creation in the economy; Iranian standard setters recognize in the framework of the accounting system that one of the most important of financial statements’ objectives is to provide useful information for decision makers, investors, creditors and regulators. This shows the interest of addressing the topic associated with the economic consequences of financial reporting quality in Tehran stock exchange. On the other hand, prior research has focused on predictions at the firm level on developed economies (Anderson et al., 2003; Anderson et al., 2007; Balakrishnan et al., 2004; Balakrishnan and Gruca, 2008;
Weiss, 2010; Kama and Weiss, 2013); we conduct our investigation by applying managerial accounting information including different types of production costs in emerging market. In addition, we demonstrate a novel application of a cost accounting concept to a qualitative variable in order to figure out their interactions. In compare to developed markets, agency problems and asymmetric information are more noticeable in emerging markets, particularly because of a poor investor protection and a poorly regulated disclosure policy. The Iranian market shows that, in particular, a very low level of investors’ legal protection. Therefore, it is suggested that financial reporting quality could have a noticeable impact on mitigating information asymmetry and agency problems between investors and managers, which finally leads to improvement in Iran’s financial environment. On the other hand, Iran market is specified by a low level of voluntary information disclosure and poor tracking by financial analysts; therefore, low financial information quality is present, and, undoubtedly, it has a considerable impact on improving firms’ financial condition. The results of this research can help managers because they indicate that eliminating additional resources and reducing costs do not always increase future profits, and if managers expect to increase future sales, it is better to apply unused resources in decreased sales periods to reduce company costs in long run and, consequently, increase firms’ profits. Taken together, the effective element such as cost stickiness in financial quality might be remarkable which is examined by present paper.

4. Research design

The purpose of the research methodology is to dedicate which statistical method is appropriate to examine the subject matter. This is a practical research in terms of purpose. The objective of a practical research is implementing applied knowledge in a particular background and putting more emphasis on the most influential elements; on the other hand, it does not take detailed reasons into full consideration. In a descriptive study, regardless of the events’ foundations, only how events really act and their intensity is taken into consideration. But in correlational studies, the primary purpose is to dedicate whether is there any relationship between two or more variables, and, if so, how intensified is this? The present research is based on the liner correlation due to the time interval over a period of six years, and it can be categorized as post-event (based on prior information) type in accounting literature because the applied data are real and historical.

4.1 Sample selection

This study population contains listed companies on the Tehran Stock Exchange for the period of 2010–2016. The main reason for selecting this period is exposed sanctions at 2008 on Iran by the United Nations Security Council which carried out considerable extra costs on Iranian’s companies in following years. Since the first affected parties in economies are companies in which accounting figures explore related side-effects. Listed firms on the Tehran Stock Exchange are dedicated for several reasons: first, they have a regular and coherent system for recording information, and second, the financial statements of these firms have higher reliability than non-listed firms in stock exchange. This is due to the government of the Stock Exchange and other oversight bodies. Accounting and financial data used in our research are collected manually from the audited financial statements provided by Research, Development and Islamic Studies Center[1] and the official website of the Tehran Stock Exchange[2]. According to the subject, we applied removal sampling in which the selected companies are chosen in accordance with above restrictions among the Tehran Stock Exchange. The exclusive features are presented as follows:

(1) Companies are not a sub-industry of the financial intermediation, holding and banks industries. This is due to the fact that such companies differ in terms of the nature of the activities and the classification of items of financial statements in compare with other companies.
The stock trading of companies should not be completely stopped during the research period.

Companies have been listed on the Tehran Stock Exchange since the beginning of 2010.

All required research data for those companies will be available during the research period.

4.2 Measurement of variables

4.2.1 Dependent variable. The dependent variable of this research is financial quality. To model a measurement instrument, we follow prior literature which defines financial reporting quality in terms of the fundamental and enhancing qualitative characteristics underlying decision usefulness as defined in the ED. The fundamental qualitative characteristics (i.e., relevance and faithful representation) are most important and determine the content of financial reporting information. The enhancing qualitative characteristics (i.e., understandability, comparability, verifiability and timeliness) can improve decision usefulness when the fundamental qualitative characteristics are established. However, they cannot determine financial reporting quality on their own. Many researchers apply predictive value as the ability of past earnings to predict future earnings (e.g., Francis et al., 2004; Lipe, 1990; Schipper and Vincent, 2003), which refers to the firm's ability to generate future cash flows. The predictive value is considered as the most remarkable indicator of relevance in respect to decision usefulness. The following model is employed in order to measure financial reporting quality:

\[ \text{CFO}_{t+1} = a_0 + \beta_1 \text{CFO}_t + \beta_2 \Delta \text{AR}_t + \beta_3 \Delta \text{INV}_t + \beta_4 \Delta \text{AP}_t + \beta_5 \Delta \text{DEPR}_t + \beta_6 \text{OTHER}_t + \epsilon_{t+1}, \]

where \( \text{CFO} \) is cash flow from operations, \( \Delta \text{AR} \) are changes in accounts receivable, \( \Delta \text{INV} \) is change in inventories, \( \Delta \text{AP} \) are changes in accounts payable and deferred liabilities, \( \Delta \text{DEPR} \) is the depreciation of fixed and intangible assets and \( \text{OTHER} \) is the net of other accruals that are calculated as follows:

\[ \text{OTHER} = \frac{\text{OP}}{\text{CFO}} - (\text{CFO} + \Delta \text{AR} + \Delta \text{INV} - \Delta \text{AP} - \Delta \text{DEPR}), \]

where \( \text{OP} \) is the operating profit and \( \epsilon \) is the error value which is assumed of having a mean of zero and a fixed variance.

4.2.2 Independent variables. To measure the stickiness of costs, we use Anderson et al.'s (2003) approach that estimates the magnitude of variation in administrative, sale, material and labor costs. The independent variables are presented as follows: \( \text{ACA} \) is equal to the difference in administrative costs in years \( t \) and \( t-1 \) divided by the sales revenue of the previous year sales cost stickiness; it is equal to the difference in sales costs in years \( t \) and \( t-1 \) divided by sales revenue of the previous year \( MCA \), which equals the cost difference of materials in years \( t \) and \( t-1 \) divided by sales revenue of the previous year. Labor cost stickiness (\( \text{LCA} \)) equals the difference in wage costs in years \( t \) and \( t-1 \) divided by sales revenue of the previous year. Overhead cost stickiness (\( \text{OCA} \)) equals the difference between overheads in years \( t \) and \( t-1 \) divided by the sales revenue of the previous year.

4.2.3 Control variables. The used control variables of this paper are defined as follows: Return on investments (\( \text{ROI} \)) equals to the ratio of operating income to investment Return on Sales (\( \text{ROS} \)), which equals the ratio of operating income to sales revenue. Return on assets
(ROA) equals to the ratio of profit to total assets price-earnings (P/E), which is equal to the ratio of current share price relative to its per-share earnings. Company size (SIZE) is the logarithm of the total value of the company’s assets. Leverage (LEV); are total assets divided by total assets minus corporate debt.

4.3 Empirical model
In order to examine the subject matters of this paper, five hypotheses are presented as follows:

\[
FRQ_{it} = a_0 + a_2 ACA_{it} + a_3 CSA_{it} + a_5 WCA_{it} + a_6 OCA_{it} + a_7 ROI_{it} + a_8 ROS_{it} + a_9 P/E_{it} + a_{10} SIZE_{it} + a_{11} LEV_{it} + \epsilon_{it}.
\]

All abbreviations are defined in above sections (dependent, independent and control variables).

5. Empirical results

5.1 Descriptive statistics
To review more about the nature of the research variables, it is required to describe them before data analysis, as the statistical description of the data is prior to statistical inference and assists us to recognize patterns which govern data. Thus, results of descriptive statistics are given in Table I. It is noted that data are adjusted in order to provide comparability feature between variables.

According to presented results of Table I, the average of FRQ shifts to 0 and its mean and standard deviation are shifted to 1. Generally, as our independent variables are mostly scaled by firm’s sale, the results are between 1 and 0. Other descriptive statistic results are presented in Table I.

5.2 The results of inferential statistics
Chow test is applied to select proper method of panel and hybrid data analysis. Also, the Jarque–Bera test is applied to test the goodness-of-fit. In other words, it tests whether sample data have the skewness and kurtosis matching a normal distribution. The results suggest abnormal distribution. Furthermore, the Durbin–Watson test is used to indicate the presence of autocorrelation (a relationship between values separated from each other) in the residuals of sample data. We find that data are not autocorrelated. Finally, Breusch–Pagan test is exerted to test the heteroskedasticity in our linear regression model, which tests whether the variance of the errors from the regression is dependent on the values of the independent variables. The results of the most important assumption indicate the presence of heteroskedasticity.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Median</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRQ</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ACA</td>
<td>0.32</td>
<td>0</td>
<td>0.468</td>
</tr>
<tr>
<td>SCA</td>
<td>0.36</td>
<td>0</td>
<td>0.48</td>
</tr>
<tr>
<td>MCA</td>
<td>0.45</td>
<td>0</td>
<td>0.498</td>
</tr>
<tr>
<td>LCA</td>
<td>0.4</td>
<td>0</td>
<td>0.491</td>
</tr>
<tr>
<td>OCA</td>
<td>0.41</td>
<td>0</td>
<td>0.492</td>
</tr>
<tr>
<td>ROI</td>
<td>63.0</td>
<td>0</td>
<td>1.299</td>
</tr>
<tr>
<td>ROS</td>
<td>1.33</td>
<td>0.73</td>
<td>2.165</td>
</tr>
<tr>
<td>ROA</td>
<td>3.38</td>
<td>1.5</td>
<td>10.43</td>
</tr>
<tr>
<td>P/E</td>
<td>0.20</td>
<td>0</td>
<td>0.642</td>
</tr>
<tr>
<td>SIZE</td>
<td>5.484</td>
<td>1.15</td>
<td>6.358</td>
</tr>
<tr>
<td>LEV</td>
<td>98.69</td>
<td>0.37</td>
<td>59.11</td>
</tr>
</tbody>
</table>

Table I. The results of descriptive statistics of the research variables
Regression results of primary model are presented in Table III. Column (1) presents regression results for all model’s variables including our cost stickiness variables. Applying cost stickiness analysis is consistent with prior findings in pervasive literature, such as Anderson et al. (2003). The adjusted $R^2$ of the model is 0.58. The overall model is significant at the 0.001 level. Our finding on parameter estimate is consistent with those found by Anderson et al. (2003) and Venieris et al. (2015), indicating that cost is sticky with respect to revenue. Moreover, the coefficient and $p$-value on administrative cost stickiness, ACA, are 21,270 and 0.000, respectively, which implies this is statistically significant at the 0.01 level, indicating that on average, the administrative cost in our sample firms has positive impact on financial quality of firms’ reports. This analysis suggests that firms with more administrative cost reports high-quality statements. The second hypothesis suggests a positive association between sale cost stickiness and financial reporting quality. The findings also suggest that our expectations are met and such a relationship is existed between those variables. Furthermore, the results of other independent variables such as, material, labor and overhead cost are negative and statistically significant at 0.01 levels, indicating greater production, and the sale costs lead to lower financial report quality. The main justification of these findings imply to firms’ effort to state less production costs, even if the financial reporting quality will be decreased. In the competitive market of Iran and a poor financial condition which is governed due to international sanctions, companies have a tendency to show their financial condition better than actual level in order to draw investors’ attention.

<table>
<thead>
<tr>
<th>Test</th>
<th>t-statistic</th>
<th>p-value</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-Limer</td>
<td>1.128</td>
<td>0.1821</td>
<td>Prefer random effect</td>
</tr>
<tr>
<td>Jarque–Bera</td>
<td>657.760</td>
<td>0.0000</td>
<td>Presence of abnormal</td>
</tr>
<tr>
<td>Dorbin–Watson</td>
<td>2.33</td>
<td>0.0000</td>
<td>Absence of autocorrelation</td>
</tr>
<tr>
<td>Breusch–Pagan</td>
<td>13.177</td>
<td>0.7243</td>
<td>Presence of heteroskedasticity</td>
</tr>
</tbody>
</table>

Table II. The results of inferential statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>SD</th>
<th>t-statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1,826</td>
<td>1,607</td>
<td>4.136</td>
<td>0.000***</td>
</tr>
<tr>
<td>ACA</td>
<td>21,270</td>
<td>9,238</td>
<td>6.303</td>
<td>0.000***</td>
</tr>
<tr>
<td>SCA</td>
<td>−2,948</td>
<td>20,710</td>
<td>−5.142</td>
<td>0.000***</td>
</tr>
<tr>
<td>MCA</td>
<td>−4,235</td>
<td>9,129</td>
<td>−5.464</td>
<td>0.000***</td>
</tr>
<tr>
<td>LCA</td>
<td>−1,925</td>
<td>1,946</td>
<td>6.003</td>
<td>0.000***</td>
</tr>
<tr>
<td>OCA</td>
<td>−1,364</td>
<td>869</td>
<td>−4.57</td>
<td>0.000***</td>
</tr>
<tr>
<td>ROI</td>
<td>38</td>
<td>42.04</td>
<td>19.904</td>
<td>0.000***</td>
</tr>
<tr>
<td>ROS</td>
<td>1,993</td>
<td>20,430</td>
<td>8.985</td>
<td>0.000***</td>
</tr>
<tr>
<td>ROA</td>
<td>32,780</td>
<td>38,560</td>
<td>1.83</td>
<td>0.2373</td>
</tr>
<tr>
<td>P/E</td>
<td>3,549</td>
<td>74,690</td>
<td>6.048</td>
<td>0.000***</td>
</tr>
<tr>
<td>LEV</td>
<td>0.0925</td>
<td>0.078</td>
<td>5.409</td>
<td>0.000***</td>
</tr>
<tr>
<td>SIZE</td>
<td>−41,830</td>
<td>9,726</td>
<td>−4.3</td>
<td></td>
</tr>
</tbody>
</table>

Table III. Regression results financial reporting quality and cost stickiness

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>117.554</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.58</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000***</td>
</tr>
</tbody>
</table>

Note: The results are obtained regarding to statistical test at 0.05 levels.
Therefore, by employing creative accounting procedures, they try to cover the high operational cost. Our control variables’ results also suggest that firms with more ROI tend to state high-quality reports. We also find the coefficient ROA is not statistically significant, which means the level of return on asset has no effect on firms’ financial reporting quality. The coefficient on P/E and LEV are statistically significant at 0.01 level, suggesting that firm’s P/E and LEV have positive impact on firm’s FRQ, and that more P/E and LEV lead to financial reports with higher quality. Finally, the obtained results suggest that big size firms tend to report less quality financial statements compared to small firms, the negative association between financial reporting quality and SIZE suggest this analysis. Taken together, our findings suggest that firms with a reasonable financial condition have more motivation to report high-quality financial statements.

6. Conclusion
The aim of this study is to investigate the association between financial reporting quality and cost stickiness in Tehran Stock Exchange. We use an equation to test all hypotheses. Our sample consist the companies’ list of Tehran Stock Exchange for the period 2010–2016. We excluded from a list of all companies with untraded stocks, banks and financial service firms due to their difference in asset structure, which is in accordance to what has been done in previous research as well (Alexeyeva, 2012). The current study attempts to provide empirical and statistical evidences about the effect of firms’ operational factors on predicting financial quality reports. Therefore, this investigation is aim to indicates the role of comprehensive operational factors on the level of qualitative features. The results of examining the first and second hypothesis suggest that administrative and sales costs are sticky; means that the reaction of costs is not proportional to the changes of sales and the amount of economic activities. In other words, when sales are increased, the amount of administrative and cost increase, but in the case of decreasing sales, the amount of decrease in cost is not as same as the increase in sales. These results are related to link between administrative, general and sales costs and various aspects of management strategies and the characteristics of companies which are consistent with the results of Venieris et al. (2015) and Chen et al. (2012). These results show that, in fact, in the long run, managers with the hope of generate earning in future try to maintain their resources, due to preventing the deprivation of future possible earnings. Therefore, this issue cause that when the firm’s sale decrease, managers do not decrease the cost proportionally. As a result, when the stickiness of administrative cost is intensified, the quality of financial reports increases and vice versa. The results of this study also suggest that the manager’s positive expectations about the future company’s condition modifies the effect of incremental or decreasing sales on the asymmetric cost behavior, which means that when the manager has a positive expectation toward the continuation of the increase in sales, the decrease in sales of the past period is less effective in the symmetry of cost behavior, while the manager’s negative expectation toward passing sales volatility leads to more symmetry of cost behavior when the company experienced an increase in sales in the past period. These findings are consistent with findings of Anderson and Lee (2016) and Banker and Byzalov (2014). In general, considering the trading cost stickiness, it is concluded that the investment of Iranian companies in OCs is more likely to be for future sales and less attention is paid to reducing cost of future sale. A possible interpretation could be extra attention of Iranian managers to marketing. In this sense, managers of Iranian companies pay more attention to investing in marketing and sales than in optimizing processes and increasing operational efficiency; moreover, they consider investing in OCs with the aim of increase of selling in future. In other words, the positive impact of investing in sales, general and administrative activities is only recognized in companies that, first, their OC is managed efficiently and, second, has plenty of opportunities to improve their manufacturing processes. The related hypotheses to the
impact of material, labor and overhead costs stickiness on financial reporting quality are still unexplored. But, generally, the results of these hypotheses show that most of shareholders in Tehran Stock Exchange, due to their long-run prospective, manage a company’s investments in an efficient manner and plays the monitoring role in reducing the agency’s costs. It is not possible to make sensible economic decisions to allocate limited resources optimally to superior activities without timely and reliable information. Prior studies show that companies with better performance present their reports timelier. So that the more the companies can make more return in applied assets, the more quickly they state reports. According to the results of this research, managers consider the sticky behavior of costs in making decisions for maximizing the company’s profit in the long run. Our results are consistent with the results of Gaynor et al. (2016), Gomariz and Ballesta (2013), Ramalingegowda et al. (2013), Shehu (2012), Aktas and Kargin (2011), Homburg and Nasev (2010) and Weiss (2010).

According to the results of this research and other similar studies, the Tehran Stock Exchange can explore more comprehensive information about the relationship between the various performance criteria for shareholders. Since, the increase in level of different standards of corporate performance and financial reporting may have a significant effect on investors’ decisions; in this order, comprehensive disclosing and transparent information, which is provided by managers about the impact of these standards on corporate leverage, could be very effective. It is suggested that financial analysts and investment counselors in the stock markets apply specific analyses based on the condition of corporate performance indicators, effective factors and financial reporting structure of the company; these analytical methods are according to accounting standards along with regular analyzes and techniques.

As an overall conclusion for the impact of the above-discussed association between cost stickiness and financial reporting quality, the following suggestions are provided for the future researchers: the measurement of financial quality and presence of other determinants in this regard could be investigated in future studies. This research examines the relationship between cost structures and financial reporting quality. It is suggested that in future investigations, the effect of cost stickiness on the conditional and unconditional conservatism could be evaluated in emerging markets.

In the process of an empirical study, usually some conditions are out of researchers control and including limitations for this study are. The stickiness of cost is effected by several items such as expert labor force and external factors which are not included in current study as they are qualitative and no measurement is not adopted in this regard. In this study, we only investigate listed firms in the Tehran Stock Exchange market; therefore, OTG firms are not concluded because the financial data of non-exchange market firms are not available in Iran.

Notes
1. www.rdis.ir
2. www.tse.ir

References


Further reading

About the authors
Nasrin Ziba received her MSc in Accounting From Gaenat Branch, Islamic Azad University. Her research interest includes accounting, auditing and finance.
Ali Daemi Gah received his MA degree from Imamreza International University of Mashhad in 2015. He is working as Practitioner in auditing field and has a professional experience of more than four years. He also possesses advanced skills in English language and has also started learning French.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Determinants of liquidity in Malaysian SMEs: a quantile regression approach

Shaista Wasiuzzaman
Faculty of Management, Multimedia University, Cyberjaya Campus, Cyberjaya, Malaysia

Abstract

Purpose – The management of liquidity has always been seen as a critical but often ignored issue in finance. Despite the abundance of studies on liquidity management, these studies mainly focus on developed countries and on large firms. Liquidity is critical for the small firm but studies on liquidity management in small and medium enterprises (SMEs) are lacking. The purpose of this paper is to examine the firm-level determinants of liquidity of SMEs in Malaysia.

Design/methodology/approach – Data are collected for a total of 986 small firms in Malaysia from 2011 to 2014, resulting in a total of 2,683 observations. Firm-specific variables and the effect of the economy are considered as the possible determinants of liquidity. Ordinary least squares (OLS) regression analysis with standard errors adjusted for firm-level clustering and quantile regression analysis are used for this purpose.

Findings – Analysis using OLS regression technique indicates that a firm’s profitability, its growth, asset tangibility, size, age and firm status are significant factors in influencing its liquidity decision. Leverage and economic condition are not found to have any significant influence on liquidity. However, quantile regression analysis provides a different picture especially for SMEs with liquidity at the quantile levels of \( \theta = 0.10 \) and \( \theta = 0.90 \). At \( \theta = 0.10 \), only profitability, tangibility and firm status are significant, while at \( \theta = 0.90 \), tangibility, size, firm status and, to some extent, age are significant in influencing liquidity levels.

Originality/value – To the author’s knowledge, this is the first study analyzing the liquidity decision of SMEs in an emerging market such as Malaysia. Most studies on liquidity management of SMEs are focused on developed countries due to data availability but these studies are also only a handful. Additionally, this study uses quantile regression analysis which highlights the need to analyze financial decisions at different levels rather than at the aggregate level as done in OLS regression analysis.

Keywords Malaysia, Small- to medium-sized enterprises, Quantile regression, Liquidity management

1. Introduction

This study investigates the liquidity of small and medium enterprises (SMEs) in an emerging market, specifically Malaysia. Prior studies on the determinants of liquidity have mainly focused on large firms and developed markets (Banos-Caballero et al., 2009; Bruinshoofd and Kool, 2002; Chiou et al., 2006; D’Mello et al., 2008; Hill et al., 2010; Kim et al., 1998; Opler et al., 1999) with relatively few focusing on small firms in emerging markets (Nazir and Azza, 2009; Smith and Fletcher, 2009; Wasiuzzaman and Arumugam, 2013; Zariyawati et al., 2010). While large firms do face liquidity problems from time to time especially during times of economic crisis (Wasiuzzaman and Arumugam, 2013), for the small firm, although it is still possible to survive without profitability, its inability to meet its short-term obligations results in liquidity being a matter of life and death (McMahon and Stanger, 1996). The small firm is more vulnerable to bad debts due to its small customer base (Young et al., 2003) and has limited access to the capital market due to its basic nature (McMahon and Stanger, 1996); therefore, cash is a critical resource for which liquidity is very much needed. Liquidity management is one of the most important and unfortunately difficult aspects of a small firm’s financial management which takes up a majority of the owner/manager’s time (Ang, 1991). Yet, studies on the liquidity management of small firms are still lacking possibly due to the unavailability of data on these firms especially in emerging/developing markets.
Firms in emerging markets are purported to be ill-governed, with management being more interested in pursuing their own interests rather than maximizing shareholder wealth (Chen and Chang, 2013). The management of liquidity is therefore not well-monitored and likely to be below optimal level. In addition, the financial systems and capital markets in Asian economies are not well developed and often bank-dominated (Yoshino and Taghizadeh-Hesary, 2015) resulting in banks being the main source of external financing for SMEs (Beck et al., 2008). SMEs in general tend to rely on bank credit and trade credit in order to finance their working capital requirement (Tran et al., 2017). Similarly, in Malaysia, 75 percent of the external financing needs of SMEs are for working capital (Haron and Ibrahim, 2016). However, the 2011 Economic Census in Malaysia finds that majority of the SMEs use internally generated funds for their financing and less than half manage to secure funds from financial institutions. Unfortunately, Malaysian firms’ liquidity performance is still lacking with working capital management treated as an ad hoc issue and in a decentralized way (www.pwc.com/my/en/services/deals/working-capital-management.html). Price-Waterhouse Coopers’ 2017 Working Capital survey of 376 Malaysian listed firms shows RM71bn of cash trapped in the working capital cycle of those firms (www.pwc.com/my/en/assets/publications/cash-for-growth.pdf). The report adds that improving working capital performance could be the answer to a company’s pursuit for additional funds. Since SMEs have difficulty in accessing external funds, the management of liquidity therefore becomes a critical issue for SMEs in Malaysia. There is a greater need for SMEs to efficiently manage liquidity so that funds are not unnecessarily tied up in short-term assets and can be utilized efficiently for the firm’s long-term financing needs, enabling the SME to grow and achieve its potential. To improve the SME’s liquidity management, it is first important to understand the factors that influence the liquidity decision. This study therefore focuses on the factors that determine the liquidity of SMEs in Malaysia.

This study uses financial data of 986 private and public-listed Malaysian SMEs taken over a three-year period to understand the firm-specific variables that may be able to explain the liquidity management decision made by SMEs in Malaysia. The findings confirm the highly significant influence of profitability, asset tangibility, size and firm status on liquidity at the aggregate level using ordinary least squares (OLS) regression analysis. However, the influence of the variables vary at different quantile levels, especially at very low levels (θ = 0.10) and very high levels (θ = 0.90) of liquidity when quantile regression analysis is used. Differences in the effects of these variables on public and private SMEs are also found. The results provide a deeper understanding of the liquidity decisions of SMEs in general. The results also highlight the importance of quantile regression analysis in understanding financial decision making in both large and small firms.

This study is structured as follows: the next section discusses prior literature on liquidity and the possible relationships between the independent variables identified from the prior literature and the dependent variable. This is followed by a section on the data and the sampling method, where the empirical model and measurement of variables are also discussed. The data are then described through descriptive statistics and analyses on the mean differences in the variables between different types of firms. The section that follows then presents the results of the statistical analysis and discussions of the results. The final section concludes this study.

2. Literature review

Past studies on liquidity have measured it via traditional measures such as the current ratio and quick ratio or via alternative measures such as the net liquid balance, working capital requirement, comprehensive liquidity index and the cash conversion cycle (Appuhami, 2008; Banos-Caballero et al., 2009; Chiou et al., 2006; D’Mello et al., 2008; Hill et al., 2010; Nazir and Afza, 2009). Liquidity is also measured by the amount of cash holdings in a firm.
These studies have mostly referred to the trade-off, pecking order (Myers, 1984) and the free cash flow (Jensen, 1986) theories which emphasize on the role of information asymmetry and agency costs, respectively, in influencing the liquidity situation in a firm (Wasiuzzaman and Arumugam, 2013). Through these theories, the studies have identified various firm-specific and corporate governance variables that may be influential in explaining the liquidity decision of a firm. Studies on cash holdings have also identified various motives for holding cash (Wasiuzzaman, 2014), such as the transactionary and precautionary motives by Keynes (1936), the agency motive by Jensen (1986) and the tax motive by Foley et al. (2007). However, most of the studies on liquidity management, cash management or working capital management have focused on large firms (e.g. Afrifa and Gyapong, 2017; Al-Najjar, 2013; Kim et al., 1998; Nunn, 1981; Wasiuzzaman, 2014), mainly because data for large firms are easier to obtain and more information is available to enable a more comprehensive analysis, especially for firms that are public-listed. Studies on small firms are relatively few and mostly focus on the relationship between liquidity and profitability (Banos-Caballero et al., 2010, 2012; Ekanem, 2010; Garcia-Teruel and Martinez-Solano, 2007; Tran et al., 2017). When small firms are studied, in most cases, these firms are public-listed firms that have easy access to the capital market. Much less has been done on privately-held small firms, although it is the privately-held small firms which face the most challenges in their financial decisions. This study therefore considers all small firms – public and private – for which data are available to understand the liquidity management decision and its determinants.

Due to limitations in data as private firm data in Malaysia are very difficult to obtain, this study only focuses on some of the firm-specific variables identified in previous studies for which data are available, i.e. profitability, leverage, asset tangibility, sales growth, firm age and firm size. Additionally, the influence of the SME’s status (private or public) and the economic condition on SME liquidity are also considered. Past studies have found the financial characteristics of small and large firms to be different; therefore, the relationships of these variables with liquidity may be different for small firms.

2.1 Profitability

The relationship between liquidity and profitability is so far the most studied with most finding a negative impact of liquidity on profitability (Banos-Caballero et al., 2010; Fiador, 2016; Ghosh and Maji, 2004; Garcia-Teruel and Martinez-Solano, 2007; Talha et al., 2010; Wasiuzzaman, 2015; Tran et al., 2017). A similar result is found by Chiou et al. (2006) and Banos-Caballero et al. (2010) when investigating the impact of profitability on liquidity because it is argued that it is easier for profitable firms to obtain funding hence liquidity (cash and working capital) will be kept at a minimum. Firms with high profitability tend to adopt a more aggressive working capital policy and hence need to invest less in liquidity (Singh and Kumar, 2017). Also, according to Haron and Norman (2015), firms with low profitability will take advantage of the credit period extended to them by suppliers and therefore would take a longer time to pay their suppliers resulting in higher working capital and hence liquidity.

However, from the agency perspective, it can be argued that the excess cash available to highly profitable firms for their investment activities will result in inefficient management of liquidity hence a positive relationship can be expected between profitability and liquidity (Wasiuzzaman and Arumugam, 2013). Al-Najjar (2013) argues that more profitable firms will have enough cash flow to pay dividends, service debt payments and stockpile cash, and hence will hold higher levels of liquidity, particularly cash. Less profitable firms will not be able to hold as much cash and would issue debt to fund their projects (Al-Najjar, 2013). Asian firms are often characterized by weak corporate governance practices (Tan and Tan,
leading to severe agency problems and therefore inefficient management of liquidity by profitable firms is expected. Additionally, with the difficulty faced by small firms in accessing external financing, it is expected that these firms will invest in liquidity when the business is doing well to take advantage of investment opportunities and to avoid any shortfalls in the future. Therefore, these two characteristics of Asian firms would mean that firms with higher profitability tend to have higher levels of liquidity, indicating a positive association between profitability and liquidity.

2.2 Leverage
Robb (2002) emphasizes the importance of debt financing for small firms, especially the young ones but Upneja et al. (2000) find that small firms tend to have higher short-term debt ratios. Similarly, Walker and Petty (1978) find higher reliance on debt financing by small firms but there is a greater tendency for them to use short-term credit. Despite the wide range of financing alternatives, according to Holmes and Kent (1991), small firms tend to prefer bank financing but they tend to have less access to term loans compared to large firms. Higher leverage can be a proxy for a firm’s access to the external debt market (John, 1993). Since external financing is costlier compared to internal financing as a result of monitoring and issue costs, Chiou et al. (2006) find that, according to the pecking order theory, firms with high levels of debt would have less internal financing as this theory advocates the use of internal financing first followed by debt financing and then equity financing to reduce the associated costs of external financing. Therefore, a firm with a higher debt ratio will pay more attention to its liquidity to enable the available capital to be used in productive investment opportunities rather than have it stuck in the operating cycle (Nazir and Afza, 2009; Wasiuzzaman and Arumugam, 2013). However, since SMEs rely mainly on short-term credit, they will keep higher levels of liquidity when their debt ratios increase so that they can meet the short-term obligations when they fall due (Ozkan, 2001). Banks are the main source of financing for SMEs (Beck et al., 2008), especially in Asian economies, and SMEs tend to finance a large share of their working capital and investment via informal sources of financing as they face higher credit constraints (Beck et al., 2006). This is also true for Malaysian SMEs which have been found to rely more on internal financing and family and friends for their financing needs; hence, a negative association between leverage and liquidity is expected for Malaysian SMEs.

2.3 Tangibility
Large manufacturing firms tend to be more capital-intensive compared to smaller ones with the ratio of fixed capital costs to total operating cost of large firms being double that of small firms (Mercer-Blackman, 2008). This may be because of the larger equity base of large firms which is used to finance investment in risky fixed assets (Ishengoma, 2004). A firm’s investment in fixed assets is expected to have a negative relationship with its investment in current assets as both types of assets compete for the limited funds available (Banos-Caballero et al., 2010). However, another form of long-term asset is the intangible assets a firm invests in. Intangible assets are more difficult to value therefore higher investment in intangible assets means higher information asymmetry and agency problem. In such a situation, firms are likely to increase investment in liquidity (Banos-Caballero et al., 2010; Wasiuzzaman and Arumugam, 2013).

2.4 Growth
A firm’s sales growth depends on its business life cycle. However, an SME’s limited access to the capital markets may inhibit its growth compared to that of a large firm. The impact of sales growth on liquidity is more complicated. Moussawi et al. (2006) argue that firms
anticipating high sales growth would need to stock up on their inventory (current asset), thus indicating a positive impact of sales growth on liquidity. According to the pecking order theory, these firms will also have higher levels of cash to avoid cash shortfalls (Opler et al., 1999). Although this relationship is found by Chiou et al. (2006), Nazir and Afza (2009), Hill et al. (2010), Barros-Caballero et al. (2010) and Wasiuzzaman and Arumugam (2013), it is argued by Chiou et al. (2006) that the high sales growth will result in the need for the firm to keep operations-related working capital and liquidity at low levels. Since SMEs have limited access to capital, it is expected that those experiencing high growth will keep liquidity at low levels and use funds available to sustain the growth.

2.5 Age
A firm’s age is expected to impact its liquidity in a positive way since young firms have better growth opportunities and therefore have less capital retained whereas older firms become stable with less growth opportunities and retain more capital resulting in higher investment in liquidity (Chiou et al., 2006). On the other hand, over time, a firm builds strong relationship with its customers and suppliers and is able to manage its inventory better resulting in the reduced need to invest in working capital (Wasiuzzaman and Arumugam, 2013).

2.6 Size
Large firms are monitored by analysts resulting in less information asymmetry; therefore, they are expected to have better access to external capital and can thus afford to have relaxed receivables and inventory policies (Hill et al., 2010). Additionally, large firms have higher sales and therefore they need larger investments in working capital (Moussawi et al., 2006). However, Moss and Stein (1993) find that larger firms manage their working capital better and keep cash at a minimum due to easier access to capital. Large firms also tend to be more diversified, fail less often and have greater bargaining power with their suppliers and customers, thus reducing the need to invest in liquidity. Since there is a higher probability for smaller firms to be financially distressed (Titman and Wessels, 1988) and they tend to suffer from higher levels of information asymmetry (Ozkan and Ozkan, 2004), they usually find it difficult and expensive to access external financing (Ferreira and Vilela, 2004). Smaller firms will therefore invest in liquidity as a precautionary motive resulting in a negative relationship between size and liquidity, as confirmed by Nazir and Afza (2009) and Wasiuzzaman and Arumugam (2013).

2.7 Firm status
A publicly-listed SME has additional avenues (the capital market) to source for external financing compared to a privately-held SME and consequently, the privately-held SME faces higher costs of external funds (Ang, 1992). Therefore, the need for liquidity is higher for privately-held SMEs to allow these firms to readily fund valuable new investment opportunities. On the other hand, since publicly-listed SMEs have easy access to external capital, inefficient management of liquidity may be possible due to the agency problems associated with free cash flow (Jensen, 1986). Since Asian firms are characterized by weak governance practices, the existence of severe agency problems mean that liquidity levels are expected to be higher for publicly-listed SMEs in Malaysia.

2.8 Economic condition
The liquidity management of small firms tends to be different from those of large firms in response to the economic condition (Lamberson, 1995). Specifically, the working capital investment of small firms tends to be relatively stable across time compared to
that of large firms. Ekanem (2010) finds that small firms have difficulty in raising finance for their liquidity needs during economic crisis due to the higher cost of borrowing and inability to provide collateral. Investment in liquidity is believed to be affected by the economic condition. Usually, financing is abundant during boom periods in the economy and it is observed by trade magazines such as CFO.com that firms are less concerned about the level of working capital and the cash stuck in the working capital cycle during boom periods. In contrast, firms try to reduce working capital and squeeze out as much cash as possible during recessionary periods (Wasiuzzaman and Arumugam, 2013). However, Chiou et al. (2006) argue that higher volumes of working capital can be expected during recession periods as a precautionary measure to ensure smooth operations due to the difficulties firms face in selling off inventory and collecting receivables.

3. Data and methodology

3.1 Data and sampling

The sample for this study consists of small, public and private firms in Malaysia. Firms must satisfy the Malaysian SME definition to be classified an SME – for the manufacturing sector: sales turnover of less than RM50m or total number of employees less than 200; and for the services and other sector: sales turnover of less than RM20m or total number of employees less than 75. Since information on the number of employees is not available, the classification is based on sales turnover. SMEs from all industries except for the finance and regulated industries are considered. The sample consists of 986 SMEs with available data for the years 2011–2014. However, since the calculation of sales growth requires lag data, the analysis is based on only the three-year period (2012–2014) resulting in a total number of 2,683 observations. While most of the firms can be classified as a small firm throughout the three-year period, there are some which are large firms in some years and small in others, and vice versa as the definition of SME is based on the sales turnover; hence, a firm which crosses the required levels in any of the years is classified as a large firm and the data for that particular year are excluded from the sample. Therefore, the total number of observations varies across the years. The firm-level data are collected from the EMIS database and information on the economic condition for the three years (proxied by change in gross domestic product (GDP)) is obtained from the Department of Statistics Malaysia website. At the time of collection, only data for years 2011–2014 were available since private company data are limited.

3.2 Empirical model and measurement of variables

The effect of the independent variables on liquidity is analyzed through the following model:

\[ NWC_{\text{Sales},it} = \beta_0 + \beta_1 \text{PROFIT}_{it} + \beta_2 \text{GROWTH}_{it} + \beta_3 \text{TANG}_{it} + \beta_4 \text{LEV}_{it} \\
+ \beta_5 \text{SIZE}_{it} + \beta_6 \text{AGE}_{it} + \beta_7 \text{PUBLIC}_{it} + \beta_8 \text{ECONCOND}_{it} + \epsilon_{it}, \]

(1)

where the dependent variable \( NWC_{\text{Sales},i} \) represents firm \( i \)'s liquidity at time \( t \). It should be noted that due to limitations of data, net working capital (NWC) is measured as the difference between the current assets and the current liabilities of the SME to represent liquidity, instead of the more popular measurement, the cash conversion cycle. The value of NWC is divided by total revenue. Information on the components of current assets and current liabilities, i.e. cash, inventory, accounts receivables and accounts payables, is not available for the private firms; therefore, cash conversion cycle cannot be used in this study to measure liquidity. Additionally, the measure of working capital used in this study, i.e. current assets minus current liabilities, is recommended by Altman (1968) as an appropriate measure of liquidity and dividing it by total revenue takes into consideration the size characteristic of the SME (Zainuddin, 2006). The independent variables are \text{PROFIT}, which
represents firm $i$’s profitability at time $t$, $GROWTH_i$ which is the sales growth of firm $i$ at time $t$, $TANG_i$ which is firm $i$’s asset tangibility or its investment in fixed assets, $LEV_i$ which measures the level of leverage in firm $i$ at time $t$, $SIZE_i$ which represents the size of the assets of firm $i$ at time $t$, $AGE_i$ which is the firm $i$’s age at time $t$, $PUBLIC_i$ which is the status of the SME (either public or private) and $ECONCOND_i$ which represents the condition of the economy as measured by the growth in the country’s GDP. $e_{it}$ is the error term and $\beta_j$ ($j = 1, 2, \ldots, 7$) are the coefficients of the independent variables to be estimated. The independent variables are chosen as per the study of Wasiuzzaman and Arumugam (2013), but some variables in their model are excluded due to the unavailability of the data needed to calculate those variables. The variables’ measurements are provided in Table I.

4. Description of data

The variables are described in Table II based on their mean, standard deviation, maximum and minimum values. The average NWC_Sales for all SMEs is 0.4054 with the minimum value being −4.1140 and the maximum being 6.2181. The average return on assets (PROFIT) is only about 3.055 percent, which is quite low but the minimum value is −35.71 percent while the maximum value is 49.52 percent. The average sales growth rate (GROWTH) is 3.61 percent with some SMEs even having negative growth rates. On average, fixed assets (TANG) make up 33.82 percent of the total assets of the SMEs in the sample while only 8.33 percent of financing is through the use of long-term debt (LEV). The average asset size (SIZE – in natural logarithm) is 1.9739, which is around RM 7 million, while the smallest SME in terms of asset size is approximately RM20,000 and the largest is RM845m. The average SME is around 23 years old (AGE) with the youngest being just established and the oldest being 107 years old. A total of 22.5 percent of the SMEs in the sample are publicly-listed firms while the rest are privately-owned ones. The firm-specific variables, except for size and age,
are winsorized at 1 and 99 percent levels to control for extreme values which may distort the results.

Further preliminary analysis of the data is carried out by dividing the sample based on age, i.e. young (300 observations) and mature (2,383 observations) SMEs, and also based on firm status, i.e. private (2,086 observations) and public (597 observations) SMEs. Following Bhattacharya et al. (2004), a firm is classified as a young firm if its age is less than ten years old, otherwise it is classified as a mature firm. Mean values for all the variables and significance in the differences of the mean values tested via the non-parametric Mann–Whitney U test are provided in Table III.

It is observed that there are significant differences in the means values of young and mature SMEs for all variables, except for NWC_Sales and GROWTH. Although the difference is insignificant, mature SMEs are found to have higher investments in liquidity. This may indicate inefficient management of liquidity especially since the mature SME is found to be significantly more profitable but its sales is growing at a slower rate on average compared to the young firm. The higher profitability and slower sales growth means that the mature SME does not need liquidity for precautionary purposes. Accordingly, it is found that the mature SME has significantly lower investment in fixed assets (TANG) and long-term debt (LEV), possibly due to the lower sales growth. The mature firm is also significantly bigger in size. Interestingly, 49 percent of the young SMEs are public-listed while only around 81 percent of the mature SMEs are privately-owned.

Significant differences in the mean values of private and public SMEs are observed for all variables, except for GROWTH. It is found that the publicly-listed SME has higher investment in liquidity but has significantly lower profitability and sales growth. This again confirms the possibility of inefficient management of working capital since there is a wider range of financing choices available to the publicly-listed SME. The publicly-listed SME is also found to have significantly higher investment in fixed assets and also long-term borrowing. The publicly-listed SME is significantly larger in size but significantly younger than the privately-held firm.

5. Statistical analysis and discussion

Correlations between the variables are calculated to check for cases of multicollinearity before regression analysis is carried out. The correlation coefficients, the p-values and the variance inflation factors (VIFs) are reported in Table IV. Based on the low correlation and VIF values, there is no serious case of multicollinearity.

Equation (1) is estimated next using the OLS method, but due to the problems of heteroscedasticity and autocorrelation, and the short time frame (three years), the estimates are calculated with Rogers (1993) standard errors adjusted for firm-level clustering as recommended by Petersen (2009). Additionally, quantile regression is carried out as it is

### Table III

<table>
<thead>
<tr>
<th>Variable</th>
<th>Young</th>
<th>Mature</th>
<th>Mean difference (p-value)</th>
<th>Private</th>
<th>Public</th>
<th>Mean difference (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWC_Sales</td>
<td>0.3506</td>
<td>0.4123</td>
<td>0.731</td>
<td>0.3137</td>
<td>0.7259</td>
<td>0.000***</td>
</tr>
<tr>
<td>PROFIT</td>
<td>0.3708</td>
<td>3.3936</td>
<td>0.003***</td>
<td>3.8808</td>
<td>0.1732</td>
<td>0.000***</td>
</tr>
<tr>
<td>GROWTH</td>
<td>0.0688</td>
<td>0.0319</td>
<td>0.472</td>
<td>0.0363</td>
<td>0.0354</td>
<td>0.721</td>
</tr>
<tr>
<td>TANG</td>
<td>0.3653</td>
<td>0.3347</td>
<td>0.038***</td>
<td>0.2943</td>
<td>0.4912</td>
<td>0.000***</td>
</tr>
<tr>
<td>LEV</td>
<td>0.0983</td>
<td>0.0814</td>
<td>0.044***</td>
<td>0.0831</td>
<td>0.0843</td>
<td>0.000***</td>
</tr>
<tr>
<td>SIZE</td>
<td>1.7717</td>
<td>1.9993</td>
<td>0.044***</td>
<td>1.5853</td>
<td>3.3318</td>
<td>0.000***</td>
</tr>
<tr>
<td>AGE</td>
<td>2.1717</td>
<td>3.2331</td>
<td>0.000***</td>
<td>3.1512</td>
<td>2.9857</td>
<td>0.000***</td>
</tr>
<tr>
<td>PUBLIC</td>
<td>0.4900</td>
<td>0.1888</td>
<td>0.000***</td>
<td>2.086</td>
<td>597</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** **,** ***Significant at the 5 and 1 percent levels, respectively
Table IV.
Pearson correlation coefficients between variables

<table>
<thead>
<tr>
<th>NWC Sales</th>
<th>PROFIT</th>
<th>GROWTH</th>
<th>TANG</th>
<th>LEV</th>
<th>SIZE</th>
<th>AGE</th>
<th>GDP</th>
<th>PUBLIC</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1161***</td>
<td>−0.0435**</td>
<td>−0.1600***</td>
<td>−0.0440**</td>
<td>0.2189***</td>
<td>0.0868***</td>
<td>0.1132***</td>
<td>−0.1264***</td>
<td>1.07</td>
</tr>
<tr>
<td>0.1161***</td>
<td>1</td>
<td>−0.0611**</td>
<td>−0.1240***</td>
<td>−0.0504***</td>
<td>0.2286***</td>
<td>0.0789***</td>
<td>−0.1284***</td>
<td>−0.1302***</td>
<td>1.02</td>
</tr>
<tr>
<td>−0.0435**</td>
<td>−0.0611**</td>
<td>1</td>
<td>0.2825***</td>
<td>0.0930***</td>
<td>−0.0511**</td>
<td>−0.0410**</td>
<td>0.0205</td>
<td>0.3000***</td>
<td>1.24</td>
</tr>
<tr>
<td>−0.1600***</td>
<td>−0.1240***</td>
<td>0.2825***</td>
<td>1</td>
<td>0.0211</td>
<td>0.0258</td>
<td>0.0814***</td>
<td>0.2158***</td>
<td>0.0040</td>
<td>1.13</td>
</tr>
<tr>
<td>−0.0440**</td>
<td>−0.0504***</td>
<td>0.0930***</td>
<td>0.0258</td>
<td>1</td>
<td>0.0054</td>
<td>0.0008</td>
<td>0.0017</td>
<td>0.4036***</td>
<td>1.13</td>
</tr>
<tr>
<td>0.2189***</td>
<td>0.2286***</td>
<td>0.0211</td>
<td>0.0814***</td>
<td>0.0054</td>
<td>1</td>
<td>0.0008</td>
<td>0.0017</td>
<td>0.4036***</td>
<td>1.36</td>
</tr>
<tr>
<td>0.0868***</td>
<td>0.0789***</td>
<td>0.0205</td>
<td>0.0814***</td>
<td>0.0008</td>
<td>0.0017</td>
<td>1</td>
<td>0.0005</td>
<td>0.4036***</td>
<td>1.36</td>
</tr>
</tbody>
</table>

Note: **, *** Significant at the 5 and 1 percent levels, respectively

1574
purported to be more powerful than the OLS regression since it works better than the OLS regression in cases where the data are skewed, variance is unequal and outliers are present. Quantile regression “produces separate estimates for all conditional quantiles of a response variable’s distribution” (Wellalage and Locke, 2013, p. 859). To illustrate the need to carry out Quantile Regression analysis, Figure 1 shows the cumulative distribution function of NWC_Sales.

Since quantile regression has the equivariance property, the marginal effects on SME liquidity in terms of the underlying level variables is provided in Table V. It is found that all the independent variables, except for GROWTH and TANG, have positive marginal effects on liquidity. For example, a 0.0100 increase in PROFIT (measured by ROA) increases NWC_Sales by one unit.

The results of the OLS regression and the quantile regression analyses for \( \theta = 0.10, 0.25, 0.50, 0.75 \) and 0.90 are reported in Table VI. It should be noted that due to missing data, the final number of observations in all these analyses is 2,681. A visual representation of the quantile regression results is provided in Figure 2.

The OLS regression analysis shows highly significant positive influence of PROFIT, SIZE and PUBLIC on NWC_Sales. TANG has a highly significant negative influence on NWC_Sales. GROWTH and AGE significantly influence NWC_Sales but only at the 10 percent level, with the influence of GROWTH being positive and AGE being negative. GDP is insignificant in influencing NWC_Sales. Since OLS regression analysis provides the aggregate effect of the independent variables on NWC_Sales, the quantile regression analyses provide a clearer picture, and, at times, opposite results, at different levels of NWC_Sales.

Further, the sample is divided into public and private SMEs to assess whether the relationship between the variables is the same for private and public SMEs. OLS regression analysis with Rogers’(1993) standard errors adjusted for firm-level clustering is carried out for this purpose. The results of the regression analysis are presented in Table VII.

<table>
<thead>
<tr>
<th>PROFIT</th>
<th>GROWTH</th>
<th>TANG</th>
<th>LEV</th>
<th>SIZE</th>
<th>AGE</th>
<th>GDP</th>
<th>PUBLIC</th>
<th>CONSTANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWC_Sales</td>
<td>0.0100</td>
<td>-0.0849</td>
<td>-0.7803</td>
<td>0.1786</td>
<td>0.0906</td>
<td>0.0717</td>
<td>0.5291</td>
<td>0.4110</td>
</tr>
</tbody>
</table>

Figure 1. Cumulative distribution function (CDF) of NWC_sales

Table V. Marginal effects on liquidity
<table>
<thead>
<tr>
<th>OLS</th>
<th>Q(0.10)</th>
<th>Q(0.25)</th>
<th>Q(0.50)</th>
<th>Q(0.75)</th>
<th>Q(0.90)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROFIT</td>
<td>0.0130*** (-3.67)</td>
<td>0.0156*** (-3.57)</td>
<td>0.00978*** (-13.52)</td>
<td>0.00767*** (-13.4)</td>
<td>0.00495*** (-4.38)</td>
</tr>
<tr>
<td>GROWTH</td>
<td>-0.134* (-2.06)</td>
<td>-0.0545 (-0.78)</td>
<td>-0.0388*** (-3.05)</td>
<td>-0.0652*** (-5.33)</td>
<td>-0.0758** (-2.72)</td>
</tr>
<tr>
<td>TANG</td>
<td>-1.406*** (-8.18)</td>
<td>-1.078*** (-6.36)</td>
<td>-0.681*** (-21.12)</td>
<td>-0.690*** (-21.79)</td>
<td>-0.928*** (-16.25)</td>
</tr>
<tr>
<td>LEV</td>
<td>0.561 (-1.69)</td>
<td>-0.193 (-0.62)</td>
<td>0.0653 (-1.10)</td>
<td>0.137** (-2.56)</td>
<td>0.339*** (3.03)</td>
</tr>
<tr>
<td>SIZE</td>
<td>0.175*** (-5.66)</td>
<td>0.00568 (-0.2)</td>
<td>0.0433*** (-8.64)</td>
<td>0.0696*** (-15.93)</td>
<td>0.122*** (-12.91)</td>
</tr>
<tr>
<td>AGE</td>
<td>0.159* (-2.08)</td>
<td>-0.0424 (-0.57)</td>
<td>0.00212 (-0.14)</td>
<td>0.0551*** (-4.07)</td>
<td>0.108*** (-4.11)</td>
</tr>
<tr>
<td>GDP</td>
<td>-2.601 (-0.69)</td>
<td>-3.436 (-0.50)</td>
<td>0.075 (-0.49)</td>
<td>0.407 (-0.28)</td>
<td>1.896 (-0.69)</td>
</tr>
<tr>
<td>PUBLIC</td>
<td>0.455*** (-3.95)</td>
<td>0.515*** (-4.82)</td>
<td>0.291*** (-13.54)</td>
<td>0.316*** (-16.54)</td>
<td>0.505*** (-13.95)</td>
</tr>
</tbody>
</table>

Notes: t-statistics in parentheses, **,**,*Significant at the 10, 5 and 1 percent levels, respectively
From the quantile regression results in Table VI, it is found that PROF significantly influences liquidity for SMEs with liquidity at the quantile levels $\theta = 0.10, 0.25, 0.50$ and $0.75$ but not at $\theta = 0.90$. Profitability is found to have a significant positive influence on liquidity, implying that SMEs with low liquidity tend to rely on their profitability to provide the cash needed to invest in liquidity to be used to encounter any sudden needs for liquidity and for other purposes such as dividend payment and debt servicing (Al-Najjar, 2013). In terms of coefficient value, the impact of profitability on liquidity is highest for SMEs in the $\theta = 0.10$ quantile region, after which the impact continues to reduce to the point that profitability becomes negative in influencing liquidity at $\theta = 0.90$. The reducing impact of profitability on liquidity is observed in the quantile regression plots in Figure 1. It is expected that SMEs which already have high levels of liquidity have no reason to increase liquidity levels therefore any increase in profitability leads to a reduction in liquidity. Hence, the profitable
firm will keep liquidity at a minimum (Chiou et al., 2006; Banos-Caballero et al., 2010). On the other hand, the less profitable firm will increase its liquidity. In fact, when a firm is already holding high levels of liquidity, any further increase could be an indication of agency problems within the firm and is a waste of the cash flow which could have been used for more profitable investment opportunities. Holding high levels of liquidity could then correspond to the agency motive (Bates et al., 2009).

GROWTH has a significant influence on NWC_Sales only for the cases of $\theta = 0.25, 0.50$ and 0.75. At the extremes, i.e. $\theta = 0.10$ and 0.90, the influence is not significant. Additionally, in terms of the coefficient value, the impact of growth on liquidity is increasing, i.e. the impact of sales growth on liquidity is lowest for $\theta = 0.10$ and highest for $\theta = 0.90$. The negative effect of growth on liquidity is argued by Chiou et al. (2006) to be due to the need for the growing SME to invest in operations resulting in the firm keeping operations-related liquidity at low levels in order to be able to manage the high sales growth. The insignificant impact of growth on liquidity at $\theta = 0.10$ could be due to the lack of liquidity, which disables the SME from using the liquidity to manage its sales growth. On the other hand, the SME that has very high levels of liquidity ($\theta = 0.90$) does not need to sacrifice its liquidity to manage its growth.

TANG is found to have a significant negative effect on NWC_Sales and the effect is found to be in the shape of an inverted U (refer to Figure 1). When a firm invests more in its fixed assets (TANG), then there will be less available to be invested in liquidity. This negative impact of tangibility and liquidity confirms the findings of Banos-Caballero et al. (2010). Based on the absolute values of the coefficients, the negative impact of tangibility on liquidity is highest for SMEs with liquidity at the 0.10 and 0.90 quartile levels, and the effect decreases from $\theta = 0.10$ to 0.50, after which it increases. Asset tangibility of an SME, or its investment in fixed assets, tends to have a higher impact on liquidity either when the SME’s liquidity is very low or when it is very high.

LEV has an insignificant influence on NWC_Sales in most cases, except for the case where $\theta = 0.75$. The impact of leverage on liquidity is found to be negative at low levels of $\theta$ and is increasing as $\theta$ increases to that point that for $\theta = 0.50$ and above, the impact is positive. Since SMEs with low liquidity levels have limited capital and cash flow and so according to the pecking order theory, they would need to go for external financing to fund any productive investment opportunities. Unfortunately, since raising external financing is difficult for these SMEs and the financing is usually costly, coupled with the SMEs having very low levels of liquidity, a higher debt ratio means they will need to pay more attention to liquidity to enable the available capital to be used in productive investment opportunities rather than have it stuck in the operating cycle (Nazir and Afza, 2009; Wasiuzzaman and Arumugam, 2013). Liquidity will therefore be kept at a minimum, enough to meet the needs of the SME. However, at values of $\theta$ greater than 0.50, liquidity levels increase as the SME takes on more debt, possibly to ensure that they meet the debt obligations when due (Ozkan, 2001).

SIZE is also found to have a consistent significant influence on NWC_Sales. The impact of size on liquidity is found to be positive implying that as firm size increases, liquidity also increases significantly. Additionally, based on the size of the coefficients, the impact of firm size on liquidity is increasing as $\theta$ increases from 0.10 to 0.90. This result is contrary to expectations as it was expected that since SMEs tend to have difficulties in access to finance (Ferreira and Vilela, 2004) due to their higher probability of financial distress (Titman and Wessels, 1988) and higher levels of information asymmetry (Ozkan and Ozkan, 2004), the smaller the firm, the higher its investment in liquidity. As an SME increases in size, it needs to be able to manage its higher sales level and would therefore need larger investments in liquidity (Moussawi et al., 2006). Additionally, since in this study, both privately-held and publicly-listed SMEs are considered, most of the large SMEs are publicly-listed firms; hence, they have easier access to external capital and can therefore afford to have higher liquidity (Hill et al., 2010). Consistent with this, PUBLIC is found to have a positive influence on
Publicly-listed SMEs are found to have significantly higher liquidity values compared to privately-held SMEs as the publicly-listed ones have greater access to external capital. The significant impact is consistent for all quartiles of liquidity.

AGE is also found to have an increasing effect on NWC_Sales but the effect is only significant at the 5 percent level for values of $\theta = 0.50$ and 0.75. Age has a negative but insignificant impact on liquidity at $\theta = 0.10$, implying that the management of liquidity is the main reason for low liquidity at this quartile level. SMEs with very low levels of liquidity and those that are young are not able to manage their inventory and are yet to build strong relationships with customers and suppliers resulting in the need to invest in liquidity (Wasiuzzaman and Arumugam, 2013). However, at quartile levels greater than 0.10, a positive relationship is found between age and liquidity, confirming the expectations that younger SMEs have better growth opportunities therefore have less liquidity but older SMEs retain more liquidity due to the lack of growth opportunities.

GDP is found to have no significant influence on liquidity for all values of $\theta$. Although the effect is insignificant, the signs of the coefficients are worth noting. The relationship is found to be negative for the cases of $\theta = 0.10$ and 0.25 and positive for the cases of $\theta = 0.50$ and 0.90. The negative relationship of economic condition on liquidity means that during recessionary period, liquidity levels are increased as a precautionary measure to ensure smooth operations due to difficulties faced by these firms in selling off their inventory and collecting their receivables (Chiou et al., 2006) and in raising funds during the recessionary periods due to the higher cost of borrowing and inability to provide collateral (Ekanem, 2010). On the other hand, SMEs with high liquidity levels will be less concerned about their liquidity levels during boom periods but try to use their liquidity to support them through the recessionary period (Wasiuzzaman and Arumugam, 2013).

Comparison of the effect of the variables on NWC_Sales for public and private SMEs shows that profitability (positive), growth (negative), tangibility (negative), leverage (positive), size (positive) and age (positive) are significantly associated with liquidity for the private SME, although age is only significant at the 10 percent level. Only economic condition is found to be insignificant. The publicly-owned SMEs however behave a little differently. Only tangibility (negative) and size (positive) are significant in influencing liquidity. Aside from the significance of the variables, the main difference between the public and private SME is that leverage has a negative association with the liquidity of public SMEs but a positive one with that of the private SME. In fact, the impact of leverage is only significant for the private SME, not for the public one. Private SMEs that have higher leverage tend to have significantly higher liquidity. This is consistent with reports by the Malaysian Central Bank (Bank Negara Malaysia) and the Census of Establishment and Enterprise 2005 by the Department of Statistics Malaysia that working capital requirement is the main purpose of Malaysian SMEs to apply for financing. On the other hand, since the public SME has more avenues for funding, its level of leverage may not be an important consideration for its liquidity decision, hence the insignificant association. Additionally, since private SMEs lack funding opportunities, they need to rely on the firm's profitability to provide the cash needed to hold higher levels of liquidity. However, investment in the growth and tangible assets means that the liquidity decision has to compete with these two investment decisions for the limited funds available; hence, a negative relationship is observed between these two variables and liquidity, consistent with the arguments by Banos-Caballero et al. (2010).

6. Conclusion
The aim of this study is to investigate the determinants of liquidity of SMEs in Malaysia. The data consists of 986 small firms with available data for the years 2011–2014, resulting in a total of 2,683 observations after accounting for the lag data requirement. Firm specific variables
such as a firm's profitability, its sales growth, asset tangibility, leverage, size and age are considered in this study. To examine the effect of the economy on the liquidity decision of a firm, the economic condition as measured by the changes in the GDP is also considered.

Analysis is carried out on the entire sample using the quantile regression technique. OLS regression analysis technique with standard errors adjusted for firm-level clustering is also used. The OLS regression results show that profitability, tangibility, size and the SMEs firm status have highly significant influences on liquidity of an SME. The influence of growth and age is only significant at the 10 percent level, while leverage and economic condition are not significant in influencing liquidity levels. However, the quantile regression results provide a clearer picture of the relationships at different quantile levels of liquidity. The quantile regression results show that SMEs with liquidity at the low levels, specifically at $\theta = 0.10$, keep liquidity levels low mainly because of the lack of cash flow in the firm. The SMEs with low levels of liquidity mainly rely on their profitability to provide the cash flows to improve their liquidity. SMEs with high levels of liquidity do not rely on their profitability to improve their liquidity, instead increased profitability reduces liquidity. Additionally, the investment in liquidity has to compete with the investment in long-term assets. Sales growth has a negative influence on liquidity. Although insignificant, leverage has a negative influence on liquidity for SMEs with low levels of liquidity ($\theta = 0.10$ and $0.25$) but the influence becomes positive at higher levels of liquidity. This corresponds to the need for the SME with low liquidity to manage its liquidity as well as possible in order to not waste its cash flows in low-return liquid assets. In most cases, age and size have significant and positive influences on liquidity. Additionally, the publicly-listed SME has higher levels of liquidity at all quantile levels. Comparison of the determinants of liquidity of public and private SMEs shows differences in the effects of the variables on the liquidity of these SMEs. Only tangibility and size are found to influence the liquidity decision of public SMEs as these SMEs have adequate access to capital. On the other hand, profitability, growth, leverage, tangibility and size have highly significant influences on the liquidity of the private SME.

The results show that the need for liquidity based on the different motives such as the transactionary, the speculative and the precautionary motives influences the investment in liquidity of SMEs. This study contributes to the further understanding of the liquidity decision from both the theoretical and practical perspective. The lack of funding opportunities faced by SMEs in developing countries results in competition between the liquidity decision and the investment decision; hence, this study shows that the liquidity decision of small firms may be different from those of large firms due to their different characteristics. Additionally, factors influencing the liquidity decision may be different even for private and public SMEs. Finally, the use of quantile regression analysis technique allows analysis of the liquidity decision at various levels, and this study shows the importance of doing this rather than at the aggregate level to gain a clearer picture.

To the author's knowledge, this is the first study which provides a deeper understanding of the liquidity decisions of the SMEs using the quantile regression analysis technique. The results are applicable to not just the Malaysian but also the global context.

7. Limitations of the study and suggestions for future research

The main limitation of this study is that it considers only some variables while others which may be more important in influencing firm liquidity are ignored. The main reason for this limitation is the availability of data. Therefore, it is recommended that future studies consider additional variables to better understand the liquidity situation and its determinants.

Another limitation is the number of years considered in this study. Since only four years’ data were available, this limitation could not be overcome in this study. It is recommended that the study be conducted over more years to enable a more in depth analysis of liquidity,
particularly during times of crisis and other major economic events. Collection of data manually rather than through databases will result in availability of more data.

Finally, prior studies have discussed endogeneity issues between liquidity and profitability since it has been proven in past studies that liquidity does influence profitability. This study does not take into consideration endogeneity when analyzing the data therefore the OLS regression and quantile regression analysis were used. The reason for this is the limitation on the number of years analyzed. Therefore, it is recommended that future studies on the liquidity decision consider analysis which control for endogeneity although the issue of endogeneity needs further thought, since liquidity management is a short-term decision while profitability is a long-term measure.

References


About the author
Shaista Wasiuzzaman is Senior Lecturer at the Faculty of Management, Multimedia University, Cyberjaya. Her research interest is in the areas of corporate finance, corporate governance and Islamic finance. Shaista Wasiuzzaman can be contacted at: shaistaw76@gmail.com

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
How authentic leadership promotes individual performance
Mediating role of organizational citizenship behavior and creativity

Neuza Ribeiro
CARMÉ – Centre of Applied Research in Management and Economics, Polytechnic Institute of Leiria, Leiria, Portugal
Ana Patricia Duarte
Business Research Unit (BRU-IUL), ISCTE-Instituto Universitario de Lisboa, Lisbon, Portugal, and
Rita Filipe
Instituto Politecnico de Leiria Escola Superior de Tecnologia e Gestao, Leiria, Portugal

Abstract

Purpose – The purpose of this paper is to provide a more comprehensive understanding of how authentic leadership (AL) can affect individual performance through creativity and organizational citizenship behavior (OCB)’s mediating roles.

Design/methodology/approach – The sample included 177 leader-follower dyads from 26 private and small and medium-sized organizations. Followers reported their perceptions of AL, and leaders assessed each follower’s level of creativity, individual performance and OCB.

Findings – The findings show that AL has a positive impact on OCB (i.e. altruism, sportsmanship, civic virtue, conscientiousness and courtesy), employee creativity, and individual performance. Creativity partially mediates the relationship between AL and individual performance. Some dimensions of OCB, namely, altruism, civic virtue and courtesy, also play a mediating role in this relationship.

Research limitations/implications – Additional studies with larger samples are needed to determine more clearly not only AL’s influence on individual performance but also other psychosocial variables affecting that relationship.

Practical implications – Organizations can increase employees’ creativity, OCB and individual performance by encouraging managers to adopt more AL styles.

Originality/value – This study is the first to integrate AL, creativity, OCB and individual performance into a single research model, thereby extending previous research. The study also used a double-source method to collect data (i.e. leader-follower dyads) to minimize the risk of introducing common-method variance.

Keywords Creativity, Individual performance, Organizational citizenship behaviour, Authentic leadership

1. Introduction

The recent economic crisis and austerity measures implemented in various countries have resulted in more negative working conditions and put a strain on the relationships between organizational leaders and followers. In addition, the latest ethical and financial scandals have fostered a high level of mistrust among many organizational members and intensified the need to consider new positive forms of leadership such as authentic leadership (AL). This style of leadership is based on leaders’ moral character and concern for others, as well as congruency between ethical values and actions (Shahid, 2010).

AL has emerged as a potentially effective approach that fosters positive employee behaviors with benefits for specific workgroups and entire organizations. For instance, positive extra-role behaviors, such as organizational citizenship behavior (OCB), go beyond formal requirements but promote the effective functioning of organizations (Organ, 1988).
Employee creativity (i.e. the generation of novel and useful ideas) (Amabile, 1988) also is considered indispensable to organizations' ability to adapt in the current fast changing and competitive global business environment.

Leaders thus must promote their employees' OCB, which is a multidimensional concept including several facets of voluntary behavior (i.e. desirable comportment not reflected in formal job requirements). These behaviors are advantageous to organizations because of this conduct's association with improved employee performance and more effective organizations (Podsakoff et al., 2009). Similarly, leaders need to promote their followers' creativity through interactions and the creation of work environments that foster ingenuity. Creativity is of value to organizations to the extent that it impacts employees' performance (Gilson, 2008). Overall, "transparent, reliable, trustworthy, ethical and veritable" leaders (Luthans and Avolio, 2003, p. 4), namely, authentic leaders, are more likely to increase their followers' OCB and creativity, which, in turn, improve employees' individual performance.

AL is a new explanatory theory of organizational leadership that has attracted widespread attention from researchers because they believe that AL can reinforce employees' positive attitudes and behaviors (Avolio and Gardner, 2005). Ilies et al. (2005) suggest that, because leaders with relational authenticity strive for open, truthful relationships with their followers, this orientation has a number of positive consequences. The present study sought to develop a deeper understanding of the effects of AL on employees' outcomes – more specifically, OCB, creativity and individual performance. In particular, this research aimed to extend the existing knowledge in this field by exploring the mediating role of OCB and creativity in the relationship between AL and employee performance.

A few researchers have previously conducted empirical studies of AL's relationships with creativity (e.g. Banks et al., 2016; Li et al., 2014; Mubarak and Noor, 2018; Rego et al., 2014), OCB (e.g. Joo and Jo, 2017; Zaabi et al., 2016), and individual performance (e.g. Clapp-Smith et al., 2009; Walumbwa et al., 2011). However, none of these studies has integrated the four concepts into a single research model. This critical gap thus needs to be filled through positive organizational research that provides a model depicting the chain of effects from AL to employees' individual performance.

The recent literature on AL has increasingly called for studies considering different types of intermediary variables to explain the relationships between AL and various outcomes (Arda et al., 2016). In particular, the mediating role of creativity and OCB in the relationship between AL and employees' individual performance is still underexplored in this literature. Therefore, the present study set out to address the aforementioned calls for further research by testing these mediating effects. In addition, by examining the mediating mechanisms of creativity and OCB on the relationship between AL and individual performance, this study's results provide valuable insights into AL theory. Although the proposed research model included some relationships that had been previously examined, the further connections explored among these relationships represent an important extension of prior work that helps explain the process of how these elements operate in unison (see Figure 1).

Figure 1.
Research model
This study’s findings thus provide a more comprehensive understanding of how AL can relate to individual performance through creativity and OCB’s mediating roles. Compared with previous research, the present study conducted a more rigorous empirical examination by collecting data from two sources (i.e., dyads of leaders and followers). By making use of a double-source method for collecting data, the risk of introducing common method variance was avoided. The results thus contribute significantly to a better understanding of what is currently clearly a central challenge for many organizations: to develop authentic leaders capable of motivating employees to engage in positive behaviors.

2. Theoretical framework and hypotheses development

2.1 Authentic leadership

The notion of authenticity has been around for centuries. According to Kapasi et al. (2016, p. 340), authenticity is to “act in accordance with one’s true self.” Gardiner (2013) suggests that individuals ultimately define themselves through interactions with others, which raises the question of how a “true self” can exist. Researchers generally agree that people can show different degrees of authenticity and that no one is completely authentic or inauthentic (Gardner et al., 2011). Nonetheless, the apparent degeneration in the moral quality of contemporary leadership (Avolio and Mhatre, 2012) has created an even stronger need for more authentic leaders.

AL theory suggests that individuals whose internal and external selves are congruent express authenticity. The concept of AL emerged in the academic literature with the birth of the positive psychology movement (Luthans and Avolio, 2003). More specifically, AL was conceived of as a root construct of the positive practices of leadership needed to develop positive work contexts (Avolio and Gardner, 2005; Avolio et al., 2004). Walumbwa et al. (2008, p. 94) thus define AL as:

A pattern of leader behavior that draws upon and promotes both positive psychological capacities and a positive ethical climate, to foster greater self-awareness, an internalized moral perspective, balanced processing of information, and relational transparency on the part of leaders working with followers, fostering positive self-development.

In contrast, no single AL definition is accepted by all. Instead, a set of four components has been used in a growing number of empirical studies (Walumbwa et al., 2008). The first of these components is self-awareness or the degree to which leaders understand their strengths, weaknesses, values, beliefs and emotions, as well as how leaders impact others. The second is relational transparency, namely, presenting leaders’ authentic selves and sincere feelings, emotions and thoughts to followers. The third is a balanced processing of information in which leaders objectively analyze all relevant information before making decisions and solicit views that question their own positions.

The last component is an internalized moral perspective. That is, authentic leaders are guided by moral values and standards, and their behaviors and decisions are consistent with internalized values, as opposed to group, organization, or societal pressures (Avolio, 2005; Avolio and Gardner, 2005; Avolio et al., 2004; Gardner et al., 2005; Shamir and Eilam, 2005; Walumbwa et al., 2008). Leaders are considered authentic to the extent that they embody these four elements. The empirical and theoretical evidence from previous research also suggests that a core AL factor can be the relationships among the four components (Gardner et al., 2005; Kernis and Goldman, 2005; Rego et al., 2012; Walumbwa et al., 2008).

Despite researchers’ expectations of correlations between AL theory and transformational (Joo and Nimon, 2014), ethical (Brown and Treviño, 2006), charismatic, servant and spiritual leadership theories (Avolio and Gardner, 2005), studies have identified several differences. For example, Walumbwa et al. (2008) found that correlations between AL and transformational and ethical leadership are not high enough to indicate that the
concepts duplicate each other (Baron, 2016). Therefore, while some overlap can exist between AL and other leadership theories, AL has unique components.

Authentic leaders treat their employees with respect, look after their needs and protect their dignity. These leaders thus inspire employees to engage in behaviors that have positive outcomes for the workplace.

2.2 AL and creativity

Creativity is the generation of novel and useful ideas (Amabile, 1988; George and Zhou, 2007; Oldham and Cummings, 1996; Shalley et al., 2004; Zhou and Shalley, 2003). To be considered creative, innovative ideas must be unique relative to other ideas available in teams and/or organizations (Shalley et al., 2004), having the potential to create value in the short or long run (George, 2007). Thus, creativity represents an important aspect of organizational change and organizations’ efficiency and survival (Woodman et al., 1993). Although creativity can be considered both an outcome and a process (Shalley and Zhou, 2008), the present study treated individual creativity as an outcome, that is, the extent to which new ideas and useful solutions are produced by employees.

Leaders can promote their followers’ creativity in two main ways: encouraging interactions and developing a work climate supportive of ingenuity. More specifically, leaders can stimulate employees’ creativity by fostering positive, fair and transparent interactions (Madjar et al., 2002; Muceldili et al., 2013; Peterson et al., 2012); demonstrating the skills needed to solve technical and creative problems (Mumford et al., 2002); or providing support (Madjar et al., 2002; Oldham and Cummings, 1996). When leadership behaviors are a significant contextual factor, they can support or inhibit creativity at work (e.g. Liu et al., 2012; Tung and Yu, 2016).

Empirical research has linked AL with individual followers’ creativity (e.g. Ilies et al., 2005; Li et al., 2014; Rego et al., 2014; Walumbwa et al., 2008; Zhou et al., 2014). Since authentic leaders promote positive interactions between leaders and employees (George, 2003) and increase employees’ positive emotions, researchers have had good reason to posit that AL has a positive relationship with individual creativity (Banks et al., 2016; Černe et al., 2013; Leroy et al., 2012). Positive emotions create the urge to play, adopt unconventional thinking and be creative. These emotions result in wellbeing at work, increasing employees’ desire to explore and assimilate new ideas, discover fresh information and develop their individual potential, as well as inducing them to be more creative (Wright and Cropanzano, 2004).

Authentic leaders’ actions are guided by their words, values and beliefs, thereby contributing to open and truthful relationships with their followers and promoting a work environment in which creativity is fostered and encouraged (Ilies et al., 2005; Rego et al., 2013). Authentic leaders also focus on followers’ talents and strengths and seek to liberate their potential (Luthans and Avolio, 2003). As a result, followers feel more confident, flexible and original, that is, prone toward a more creative mindset. Authentic leaders feel less threatened by the changes that the creative ideas of subordinates might imply and more likely to accept such suggestions (Michie and Goody, 2005), thus welcoming and promoting employee creativity.

In summary, authentic leaders may encourage employees to be creative because these managers inspire trust and positive achievements, which stimulates “thinking outside the box” and improves emotional safety (Avalio et al., 2004). In view of these considerations, the following hypothesis was proposed for the present study:

**H1.** Employees’ perceptions of AL are positively related with their creativity.

2.3 Authentic leadership and organizational citizenship behavior

OCB refers to individuals’ autonomous behaviors that are not directly or explicitly recognized by formal reward systems but that help their organizations to function
effectively (Organ, 1988). The cited author suggests five OCB dimensions have been the most frequently examined by researchers (e.g. Podsakoff et al., 2000; Schnake and Dumler, 2003). These are, first, altruism or voluntarily helping others with organizationally relevant tasks or problems. The second is sportsmanship, which refers to tolerating the inevitable inconveniences and annoyances of organizational life without complaining and filing grievances. The third dimension is courtesy or behaviors that seek to prevent work-related problems with others. The fourth is conscientiousness, which involves employee behaviors that go beyond minimal requirements when carrying out work-related tasks. The last dimension is civic virtue or responsive, constructive involvement in the workplace.

Williams and Anderson (1991) further divide OCBs into two categories. The first is OCB toward individuals, namely, behaviors that are directly helpful to specific individuals and that indirectly contribute to organizations. The second is OCB toward organizations, that is, behaviors directly contributing to organizations. Overall, OCB improves social capital and enhances organizational functions (Bolino et al., 2002) and effectiveness (Podsakoff and MacKenzie, 1997; Rego and Cunha, 2008). Furthermore, researchers have found that OCB reduces adverse employees’ behaviors and attitudes such as turnover intention and absenteeism (Podsakoff et al., 2009; Shahin et al., 2014).

The assumption is that, if leaders make an effort to inspire positive types of behavior, employees are more likely to demonstrate these. Therefore, authentic leaders offer role models that can have a positive effect on their employees’ behavior, which, in turn, can translate into effective OCB. Hackett et al. (2003), for example, found a strong correlation between the quality of leader-member exchanges and overall OCB.

Researchers have also found that the more employees perceive their leaders to be authentic, the higher their level of satisfaction is with those leaders, which affects their organizational commitment and willingness “to walk the extra mile” (i.e. exhibit OCB) (Clapp-Smith et al., 2009; Moriano et al., 2011). The relevant literature has shown that AL creates a positive, transparent and fair environment that influences employees’ attitude and willingness to engage in OCB. For example, Walumbwa et al. (2007) found that the composite factor of AL dimensions was a significant positive predictor of self-reported OCB, as well as supervisor-rated job performance – after controlling for organization climate. According to Zaabi et al. (2016), AL significantly increases OCB among workers in the petroleum sector. Joo and Jo (2017) also found that perceived AL has a significant impact on employees’ OCB.

It can also be explained by social exchange theory (Blau, 1964), that is, when employees perceive their leaders as authentic, they develop a feeling of obligation to reciprocate by additional efforts in activities that extend beyond their normal duties. In summary, previous studies’ findings have suggested that authentic leaders are primarily responsible for the OCB of their followers (Avolio et al., 2004; Kernis, 2003; MacColl-Kennedy and Anderson, 2002). Therefore, the following hypothesis was proposed for the present study:

H2. Employees’ perceptions of AL are positively related with their OCB.

2.4 AL and individual performance
According to Campbell (1990), employees’ individual performance refers to a set of individual actions and behaviors that are relevant to their organizations’ goals. One of the simplest definitions of employee performance is “the extent to which a job is well done” (Campbell et al., 1993). The assessment of individual employees’ performance is important not only to ensure better management of companies but also to facilitate empirical analyses. Good individual performance thus means employees have accomplished job-related responsibilities to a satisfactory extent or to the degree expected by their employer.
Researchers have found AL to be related positively to employee performance (Clapp-Smith et al., 2009). That is, AL drives employee performance and OCB through higher trust and identification with the leaders in question (Walumbwa et al., 2011). Authentic leaders are true to their moral and ethical values and strive to achieve open and genuine relationships (Gardner et al., 2005; Ilies et al., 2005; Kernis, 2003), thereby promoting their followers’ development and, consequently, improving their performance. Furthermore, authentic leaders emphasize employees’ talents and strengths, liberate their potential (Gardner and Schermerhorn, 2004; Luthans and Avolio, 2003), and emphasize their growth. As a result, employees’ self-efficacy potentially increases through the effects of self-fulfilling prophecies (Avolio et al., 2004; Gardner et al., 2005; Luthans et al., 2007).

According to belongingness theory (Baumeister and Leary, 1995), one of the primary human drives is the need to belong and form strong positive interpersonal relationships. Thus, AL likely stimulates employees to perform better because of their need to belong. That is, workers want to feel a sense of belonging to their organization, and they need strong positive interpersonal relationships, thus workers’ perceptions of AL encourage them to improve their performance, thus providing benefits to their organization.

Overall, previous studies have confirmed that followers reciprocate their leaders’ authenticity with better individual performance (Ribeiro et al., 2018). Social exchange theory (Blau, 1964), therefore, can be used to explain why individuals who perceive AL develop a feeling of obligation to reciprocate by improving their performance. Accordingly, the following hypothesis was proposed for the present research:

**H3.** Employees’ perceptions of AL are positively related with their individual performance.

### 2.5 Creativity and individual performance

Creative employees have unique characteristics that help them differentiate themselves from co-workers, specifically through the exercise of independence, confidence and a willingness to take risks (Sternberg and Lubart, 1999). These features promote greater flexibility, thus making these individuals more adaptable and open to new experiences and, consequently, better able to achieve higher individual performance.

Luthans et al. (2007) report that subordinates with higher levels of creativity have better individual performance due to increased cognition and motivation and more positive behaviors. According to the cited authors, the most creative employees are continually looking for challenges and striving to meet set targets, thereby producing better performance. Lim and Gilson (2013) examined the effects of individual and team creativity on employee performance, and their results show that individual creativity has an inverted U-shaped relationship with individual performance. In addition, team creativity moderates the association between individual creativity and individual performance. According to Zhou (2003), if conditions are favorable, non-creative employees can be encouraged to develop their creativity to increase their performance.

In view of these findings, the following hypothesis was formulated for the present study:

**H4.** Employee creativity is positively related with employees’ individual performance.

### 2.6 OCB and individual performance

In general, OCB tends to have positive relationships with individual outcomes (Podsakoff et al., 2009). Previous studies have connected OCB with productivity (Podsakoff et al., 1997) and sales performance (Podsakoff and MacKenzie, 1994). Podsakoff and MacKenzie (1994) studied this relationship in 116 branches of a large insurance company, seeking to understand the impacts of three dimensions (i.e. helping behavior, sportsmanship and civic virtue)
on employees' performance. The cited authors' results confirm that all dimensions help explain employee performance.

Dunlop and Lee's (2004) findings also indicate a positive association between OCB and job performance. According to Farooqui (2012), many managers assume OCB is part of their employees' responsibilities, including helping co-workers, actively taking on work beyond their own specific job tasks, identifying problems, or creating new suggestions and innovative ideas for their organization. Prior research has shown that OCB is positively related to favorable evaluations of individuals' overall performance (Podsakoff et al., 2009; Whiting et al., 2008).

OCB is an important behavioral criterion used to measure individual performance. In other words, OCB is thought to improve performance by promoting social and psychological attitudes that contribute to employee performance (Allen et al., 2000). Anvari et al. (2015) confirmed that the OCB dimensions of altruism, conscientiousness, courtesy, sportsmanship and civic virtues improve employees' performance. Hence, the following hypothesis was proposed for the present research:

H5. OCB is positively related with employees' individual performance.

2.7 Creativity's mediation of the relationship between AL and individual performance

Given the assumption that employee creativity enhances work outcomes (e.g. individual performance), considerable attention needs to be devoted to identifying creativity's antecedents (e.g. AL). For instance, authentic leaders incorporate values, beliefs and moral principles into their behavior, which influence positively these leaders' relationship with their followers and organizations' climate (Avolio and Gardner, 2005) and, in turn, awakens these employees' creative spirit (Rego et al., 2014) and improves their performance (Lim and Gilson, 2013). The current study's research model thus suggested that the effect of AL in employees' individual performance can be both direct and indirect (e.g. through creativity).

Authentic leaders have characteristics such as persistence, intrinsic motivation (Rego et al., 2012), self-consciousness and an innovative cognitive style. These can increase followers' willingness to overcome obstacles and look for resources to achieve work goals (Licata et al., 2003). AL can further help employees to present new and useful ideas (Organ and Ryan, 1995), thereby influencing employees' level of creativity (Yan et al., 2011) and enhancing their individual performance (Lim and Gilson, 2013).

In summary, employees need to perceive that their leaders are guided by moral values and standards, as well as being self-aware of their attitudes, openly sharing information and considering all relevant information before coming to a decision. That is, when followers see their leaders as authentic, the former feel freer and more confident to be creative, and, as a result, they increase their performance. Thus, the following hypothesis was developed for the present research:

H6. Creativity mediates the relationship between AL and employees' individual performance.

2.8 OCB's mediation of the relationship between AL and individual performance

AL has a positive impact on OCB, and its dimensions have a positive influence on employees' individual performance. AL promotes OCB because employees are influenced by their leader and they are more likely to both adopt behaviors that help their organization and, consequently, perform better (Avolio et al., 2004; Ilies et al., 2005; MacColl-Kennedy and Anderson, 2002). The current study's conceptual model thus posited that AL can affect employees' individual performance both directly and indirectly (e.g. through OCB).
According to Podsakoff et al. (2000), a factor that can increase OCB is job satisfaction. Employees who are more satisfied by their work tend to deal with their tasks in more responsible and dedicated ways. Job satisfaction is also influenced by AL, that is, authentic leaders create ties with followers, making them feel more responsible and satisfied and thus more likely to demonstrate OCB and better performance. Karambayya (1990) also showed that employees who adopt more OCBs are more satisfied with their job and have better performance levels.

Avolio et al. (2004) argue that, if leaders express feelings of confidence, positive emotions, and optimism, followers respond with greater commitment and satisfaction with their work, as well as attributing greater value to their company. Employees also show less turnover intentions and make an extra effort, consequently producing better performance. Thus, a positive relationship may exist between AL and employees’ performance through the mediating effect of OCB. The following hypothesis was therefore proposed for the present study:

\[ H7. \text{ OCB mediates the relationship between AL and employees' individual performance.} \]

3. Method
3.1 Sample and procedures
The present study gathered a convenience sample that included 177 leader-follower dyads from 26 private and small and medium-sized organizations operating in Portugal. The organizations were from different sectors, including office supplies, food, construction, and automobiles.

The respondents were between 20 and 57 years old (mean = 31.44; standard deviation [SD] = 7.30), and 55.0 percent were females. The surveyed individuals’ level of education was as follows: 6.8 percent had completed 6 years of schooling or less, 63.3 percent had between 7 and 12 years and 29.9 percent had a higher education degree. Regarding job tenure, followers had been employed in their organizations for an average of 6.12 years (SD = 4.36 years; minimum = 1 year; maximum = 25 years). The average contact time of each follower with his or her leader was 5.56 years (SD = 3.96 years; minimum = 1 year; maximum = 23 years).

Data were collected using questionnaires. Several organizations were asked to collaborate in the data collection process. In the case of positive replies, the individuals in charge of each organization’s work teams and/or departments were identified and invited to participate in the study voluntarily. The requirement of collecting data from leader-follower dyads was made clear to each supervisor. If consent was granted, questionnaires were delivered to leaders and followers by a member of the research team.

A letter accompanied each questionnaire, explaining the research goals and guaranteeing maximum confidentiality of the data collected and anonymity of respondents. The instructions explicitly stated that there were no right or wrong answers, and respondents were asked to answer the questions as honestly as possible. Some instructions were also provided regarding how to complete the questionnaire and, thus, reduce possible errors. To avoid any embarrassment, leaders and followers were asked to complete questionnaires in separate locations and, upon completion, to place the questionnaires in the envelope provided and seal it. According to Podsakoff et al. (2003), protecting respondents’ anonymity and diminishing evaluation apprehension contribute to reducing lenient, acquiescent and socially desirable answers.

Followers were asked to report their perceptions of AL. Leaders were asked to assess their followers’ individual level of creativity, performance, and OCB (i.e. a maximum of six employees per supervisor). Both questionnaires were pre-tested. Each contained questions regarding the respondents’ age, gender, education, tenure in their organization and tenure in their dyads (i.e. length of time in leader-follower dyads).
3.2 Measures

3.2.1 AL (predictive variable). AL was measured using the 16 items of the AL Questionnaire (ALQ) developed by Walumbwa et al. (2008). The Portuguese version used had previously been translated and back-translated by Rego et al. (2012). The ALQ measures four dimensions:

1. self-awareness (e.g. “[…] seeks feedback to improve interactions with others”);
2. relational transparency (e.g. “[…] is willing to admit mistakes when they are made”);
3. internalized moral perspective (e.g. “[…] makes decisions based on his/her core beliefs”); and
4. balanced processing of information (e.g. “[…] listens carefully to different points of view before coming to conclusions”).

Followers reported the frequency with which their leaders adopted each behavior on a five-point Likert response scale (1 = “Never” to 5 = “Often, if not always”).

Based on relevant studies in the literature on AL (e.g. Rego et al., 2010, 2013; Walumbwa et al., 2008), this variable was considered an overall construct. To obtain a composite AL score, the procedure suggested by Luthans et al. (2008) was followed. First, the items for each of the four dimensions were calculated to obtain a composite average for each dimension. Then, the averages for each of the four dimensions were calculated to arrive at a composite AL score for each follower (α = 0.94). Higher scores represent perceptions of stronger AL.

3.2.2 Individual creativity (mediator variable). Employees’ creativity was measured using an adapted version of the 13-item scale developed and validated by Zhou and George (2001). The translation of these items into Portuguese followed the standard procedures for translations of research instruments (Brislin and Berry, 1986). A sample item is “[…] is not afraid to take risks” (α = 0.98). Leaders indicated how characteristic the 13 behaviors were of each follower on a five-point Likert response scale (1 = “Not at all characteristic” to 5 = “Very characteristic”).

3.2.3 OCB (mediator variable). Followers’ OCB was measured using 17 items adapted from Konovsky and Organ’s (1996) research and later validated by Rego et al. (2010) in Portugal. These items measure five dimensions:

1. altruism (four items, e.g. “[…] helps make others more productive”; α = 0.94);
2. sportsmanship (four items, e.g. “[…] is able to tolerate occasional inconveniences when they arise”; α = 0.83);
3. civic virtue (three items, e.g. “[…] stays informed about developments in the company”; α = 0.89);
4. conscientiousness (three items, e.g. “[…] gives advance notice when unable to come to work”; α = 0.88); and
5. courtesy (three items, e.g. “[…] respects the rights and privileges of others”; α = 0.72).

Leaders indicated the degree to which each statement applied to each follower on a seven-point Likert response scale (1 = “Does not apply to this employee at all” to 7 = “Applies to this employee completely”).

In order to obtain a composite OCB score, the items for each of the five dimensions were calculated to obtain a composite average for each dimension. Then, the averages for each of the five dimensions were calculated to arrive at a composite OCB score for each employee (α = 0.94). Higher scores represent perceptions of higher OCB.
3.2.4 Individual performance (criterion variable). Followers’ individual performance levels were measured using three items adapted from Mackenzie et al. (1993). These items were also translated into Portuguese by following the standard procedures for research instrument translation (Brislin and Berry, 1986). A sample item is: “All things considered, this employee is outstanding” \( (\alpha = 0.97) \). Leaders indicated the degree to which statements applied to each follower on a seven-point Likert response scale (1 = “Does not apply to this subordinate at all” to 7 = “Applies to this subordinate completely”).

4. Data analysis and results

The data were analyzed using IBM’s SPSS software 22.0 version. Table I lists the means, SDs and correlations between variables. Gender correlates positively with education and conscientiousness and negatively with overall AL, with females reporting lower perceptions of their leaders’ AL. Age correlates negatively with education and positively with tenure in dyad. Education correlates positively with tenure in dyad, all OCB dimensions, overall OCB, individual performance, creativity and overall AL. Tenure in dyad also correlates positively with all OCB dimensions, overall OCB, individual performance, creativity and overall AL. Moreover, all the main variables in the study intercorrelate positively.

4.1 Perceived AL as a predictor of OCB, creativity and individual performance

Table II presents the results of regression analyses performed for the five OCB dimensions, overall OCB, creativity and individual performance – with AL as the predictor variable. These analyses facilitated the testing of the first three research hypotheses.

In the first step of the analyses, gender, age, education and tenure in dyad were set as control variables given that they were correlated with some variables. As can be seen in Table II above, age has no statistically significant effects on the variables to be explained. Gender has statistically significant negative effects on sportsmanship \( (\beta = -0.21; p < 0.05) \) and creativity \( (\beta = -0.26; p < 0.05) \). Both follower education and tenure in dyad show significant positive effects on all variable, with \( \beta \) coefficients ranging from 0.25 to 0.46 in the case of education and 0.03 and 0.09 in the case of tenure in dyad. Taken together, the control variables explain a sizable proportion of the variance of OCB, creativity and individual performance, with adjusted \( R^2 \) values ranging from 13 to 28 percent.

In the second step, AL was added to observe its effects on variables. The results indicate that AL has a significant positive effect on all variables, increasing the ability to explain their unique variance from 10 to 25 percent. More specifically, AL has a positive effect on each OCB dimension – altruism \( (\beta = 0.81; p < 0.000) \), sportsmanship \( (\beta = 0.52; p < 0.000) \), civic virtue \( (\beta = 0.70; p < 0.000) \), conscientiousness \( (\beta = 0.44; p < 0.000) \) and courtesy \( (\beta = 0.56; p < 0.000) \) – as well as on overall OCB \( (\beta = 0.61; p < 0.000) \). AL also positively impacts follower creativity \( (\beta = 0.59; p < 0.000) \) and individual performance \( (\beta = 0.82; p < 0.000) \). Overall, these findings suggest that followers’ perceptions of AL lead them to engage more frequently in OCB, be more creative, and perform better at work, which provides empirical support for \( H1–H3 \).

4.2 Creativity as a predictor of individual performance

Regression analyses were also performed to test \( H4 \). As can be seen in Table III, after controlling for gender, age, education and tenure in dyad, employees’ creativity explains 42 percent of the unique variance of individual performance \( (\beta = 0.88; p < 0.000) \). \( H4 \) was thus supported.

4.3 OCB as a predictor of individual performance

Similar analyses were conducted to examine OCB’s effect on individual performance. The five OCB dimensions were entered into the analyses as predictors, and
<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gender(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Age</td>
<td>31.44</td>
<td>7.30</td>
<td>–0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Education(^b)</td>
<td>2.46</td>
<td>0.56</td>
<td>0.19*</td>
<td>0.92**</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Tenure in dyad</td>
<td>5.56</td>
<td>3.98</td>
<td>–0.03</td>
<td>0.41***</td>
<td>0.18*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Altruism</td>
<td>5.71</td>
<td>1.10</td>
<td>–0.03</td>
<td>0.05</td>
<td>0.32***</td>
<td>0.38***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Sportsmanship</td>
<td>5.93</td>
<td>0.89</td>
<td>–0.05</td>
<td>0.01</td>
<td>0.42***</td>
<td>0.36***</td>
<td>0.76***</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Civic virtue</td>
<td>5.66</td>
<td>1.10</td>
<td>–0.00</td>
<td>0.16</td>
<td>0.41***</td>
<td>0.40***</td>
<td>0.88***</td>
<td>0.77***</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Conscientiousness</td>
<td>6.01</td>
<td>0.89</td>
<td>0.14*</td>
<td>–0.01</td>
<td>0.31***</td>
<td>0.20**</td>
<td>0.68***</td>
<td>0.57***</td>
<td>0.70***</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Courtesy</td>
<td>5.92</td>
<td>0.88</td>
<td>–0.03</td>
<td>0.06</td>
<td>0.47***</td>
<td>0.31***</td>
<td>0.82***</td>
<td>0.72***</td>
<td>0.82***</td>
<td>0.73***</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Individual performance</td>
<td>5.65</td>
<td>1.14</td>
<td>0.07</td>
<td>0.07</td>
<td>0.31***</td>
<td>0.37***</td>
<td>0.90***</td>
<td>0.70***</td>
<td>0.87***</td>
<td>0.76***</td>
<td>0.80***</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Creativity</td>
<td>3.67</td>
<td>0.94</td>
<td>–0.10</td>
<td>0.01</td>
<td>0.30***</td>
<td>0.37***</td>
<td>0.79***</td>
<td>0.76***</td>
<td>0.81***</td>
<td>0.53***</td>
<td>0.71***</td>
<td>0.76***</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>12. Overall AL</td>
<td>4.02</td>
<td>0.71</td>
<td>–0.20**</td>
<td>0.12</td>
<td>0.13*</td>
<td>0.29***</td>
<td>0.60***</td>
<td>0.52***</td>
<td>0.55***</td>
<td>0.37***</td>
<td>0.53***</td>
<td>0.56***</td>
<td>0.54***</td>
<td>–</td>
</tr>
<tr>
<td>13. Overall OCB</td>
<td>5.84</td>
<td>0.87</td>
<td>0.02</td>
<td>0.03</td>
<td>0.45***</td>
<td>0.37***</td>
<td>0.90***</td>
<td>0.85***</td>
<td>0.91***</td>
<td>0.82***</td>
<td>0.91***</td>
<td>0.91***</td>
<td>0.81***</td>
<td>0.58***</td>
</tr>
</tbody>
</table>

**Notes:**
- 0 = male; 1 = female;
- \(^b\)1 = 6 or less years of education; 2 = 7 to 12 years of education; 3 = higher education.
- AL and creativity = five-point response scale; OCB and individual performance = seven-point response scale.
- *p < 0.05; **p < 0.01; ***p < 0.001
socio-professional variables were set as controlled variables. Table IV presents the results obtained, according to which only altruism ($\beta = 0.49$; $p < 0.000$), civic virtue ($\beta = 0.31$; $p < 0.000$) and conscientiousness ($\beta = 0.35$; $p < 0.000$) significantly affect individual performance. Sportsmanship ($\beta = -0.04$; nonsignificant (ns)) and courtesy ($\beta = 0.11$; ns) have no statistically significant effect on individual performance at work.

Taken together, OCB dimensions explain 68 percent of the unique variance of individual performance. These results indicate that those followers who show higher altruism, civic
virtue, and conscientiousness are also those who carry out their functions more competently. Since only some of the dimensions of OCB, but not all, explain individual performance, H5 received only partial support.

4.4 Creativity as a mediator of the relationship between AL and individual performance

To test H6, Baron and Kenny’s (1986) procedure for mediation analysis was followed and complemented with Sobel’s (1982) test. Table V presents the results obtained, which suggest that creativity ($\beta = 0.74; p < 0.000$) partially mediates the effect of AL ($\beta = 0.38; p < 0.000$) on individual performance. Thus, perceptions of AL appear to encourage employees to be more creative at work and, subsequently, to achieve better job performance levels.

Table IV. Hierarchical regression analyses: how five dimensions of OCB predict individual performance (standardized coefficients)

<table>
<thead>
<tr>
<th>Individual performance</th>
<th>Authentic leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td>First step</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.00</td>
</tr>
<tr>
<td>Gender$^a$</td>
<td>0.09</td>
</tr>
<tr>
<td>Education$^b$</td>
<td>0.30***</td>
</tr>
<tr>
<td>Tenure in dyad</td>
<td>0.09****</td>
</tr>
<tr>
<td>$F$</td>
<td>10.49****</td>
</tr>
<tr>
<td>$R^2$ adj.</td>
<td>0.18</td>
</tr>
<tr>
<td>Second step</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.00</td>
</tr>
<tr>
<td>Gender$^a$</td>
<td>0.19***</td>
</tr>
<tr>
<td>Education$^b$</td>
<td>−0.11*</td>
</tr>
<tr>
<td>Tenure in dyad</td>
<td>0.01</td>
</tr>
<tr>
<td>Altruism</td>
<td>0.49****</td>
</tr>
<tr>
<td>Sportsmanship</td>
<td>−0.04</td>
</tr>
<tr>
<td>Civic virtue</td>
<td>0.31****</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>0.35****</td>
</tr>
<tr>
<td>Courtesy</td>
<td>0.11</td>
</tr>
<tr>
<td>$F$</td>
<td>122.96****</td>
</tr>
<tr>
<td>$R^2$ adj.</td>
<td>0.86</td>
</tr>
<tr>
<td>$R^2$ change</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Notes: $^a0 = \text{male}; 1 = \text{female}; \ ^b1 = 6 \text{ or less years of education}; 2 = 7 \text{ to 12 years of education}; 3 = \text{higher education}$. $^*p < 0.05; \ ^{**}p < 0.01; \ ^{***}p < 0.001$

Table V. Hierarchical regression analyses: how creativity mediates effect of AL on individual performance (standardized coefficients)

<table>
<thead>
<tr>
<th>Individual performance</th>
<th>Authentic leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Third step</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.01</td>
</tr>
<tr>
<td>Gender$^a$</td>
<td>0.39***</td>
</tr>
<tr>
<td>Education$^b$</td>
<td>0.08</td>
</tr>
<tr>
<td>Tenure in dyad</td>
<td>0.01</td>
</tr>
<tr>
<td>Overall AL</td>
<td>0.38****</td>
</tr>
<tr>
<td>Creativity</td>
<td>0.74****</td>
</tr>
<tr>
<td>$F$</td>
<td>52.50****</td>
</tr>
<tr>
<td>$R^2$ adj.</td>
<td>0.64</td>
</tr>
<tr>
<td>$R^2$ change</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Notes: $^a0 = \text{male}; 1 = \text{female}; \ ^b1 = 6 \text{ or less years of education}; 2 = 7 \text{ to 12 years of education}; 3 = \text{higher education}$. $^*p < 0.05; \ ^{**}p < 0.01; \ ^{***}p < 0.001$
The results for Sobel’s (1982) test support the existence of a partial mediation effect ($z = 6.81; p < 0.05$). The full model explains 64 percent of the unique variance of individual performance. $H6$ was thus empirically supported.

### 4.5 OCB as a mediator of the relationship between AL and individual performance

The same analytical procedure was followed to test $H7$. As can be seen in Table VI, followers’ altruism ($\beta = 0.48; p < 0.000; z = 8.95; p < 0.05$), civic virtue ($\beta = 0.31; p < 0.000; z = 7.79; p < 0.05$) and conscientiousness ($\beta = 0.26; p < 0.000; z = 4.90; p < 0.05$) fully mediate the effect of AL ($\beta = 0.08; ns$) on individual performance. Sportsmanship ($\beta = -0.05; ns$) and courtesy ($\beta = 0.19; ns$) again have no statistically significant effect on individual performance.

Thus, perceptions of AL appear to motivate followers to show more altruism, civic virtue and conscientiousness at work, and this, subsequently, leads them to achieve better levels of job performance. The full model explains 86 percent of the unique variance of individual performance. However, since only some OCB dimensions mediate the relationship between AL and individual performance, $H7$ received only partial support.

### 5. Discussion and conclusions

#### 5.1 Main findings

##### 5.1.1 AL perceptions as a predictor of OCB, creativity and individual performance

The findings show that AL has a positive impact on OCB (i.e. altruism, sportsmanship, civic virtue, conscientiousness and courtesy), employee creativity and individual performance. The positive effect of AL on OCB confirmed by the present study is consistent with the results reported by previous researchers (e.g. Avolio et al., 2004; Joo and Jo, 2017; Kernis, 2003; Walumbwa et al., 2007; Zaabi et al., 2016). These state that perceptions of employees regarding leaders’ authenticity influence followers’ willingness to engage in OCB. That is, when employees regard their leaders as being authentic and creating a positive, transparent, and fair work environment, workers tend to reciprocate with higher levels of OCB.

The current study’s results indicate that AL has a positive influence on employees’ creativity, which is also in line with past research (e.g. Banks et al., 2016; Černe et al., 2013; Li et al., 2014; Rego et al., 2014; Zhou et al., 2014). Authentic leaders contribute to developing open, truthful relationships with their followers, thereby stimulating positive psychological

<table>
<thead>
<tr>
<th>Individual performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$ adj.</td>
</tr>
<tr>
<td>$R^2$ change</td>
</tr>
</tbody>
</table>

Table VI. Hierarchical regression analyses: how OCB mediates effect of AL on individual performance (standardized coefficients)

<table>
<thead>
<tr>
<th>Third step</th>
<th>Age</th>
<th>Gender*</th>
<th>Education b</th>
<th>Tenure in dyad</th>
<th>Overall AL</th>
<th>Altruism</th>
<th>Sportsmanship</th>
<th>Civic virtue</th>
<th>Conscientiousness</th>
<th>Courtesy</th>
<th>F</th>
<th>$R^2$ adj.</th>
<th>$R^2$ change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.00</td>
<td></td>
<td></td>
<td>0.20**</td>
<td>-0.11*</td>
<td>0.08</td>
<td>0.48***</td>
<td>0.31****</td>
<td>0.26***</td>
<td>0.09</td>
<td>0.64</td>
<td>0.24</td>
<td>0.46</td>
</tr>
<tr>
<td>Gender a</td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.20**</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Education b</td>
<td></td>
<td>0.20**</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Tenure in dyad</td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Overall AL</td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Altruism</td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Sportsmanship</td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Civic virtue</td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Courtesy</td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td>0.64</td>
<td></td>
<td>0.24</td>
<td>0.46</td>
<td>0.64</td>
<td>0.24</td>
<td>0.46</td>
<td>0.64</td>
<td>0.24</td>
<td>0.46</td>
<td>0.24</td>
<td>0.46</td>
</tr>
<tr>
<td>$R^2$ change</td>
<td></td>
<td>0.46</td>
<td></td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Notes: $*0 = \text{male}; 1 = \text{female}; b1 = 6 \text{ or less years of education}; 2 = 7 \text{ to 12 years of education}; 3 = \text{higher education}$. $*p < 0.05; **p < 0.01; ****p < 0.001$
states that produce wellbeing at work and promoting work environments in which creativity is fostered and encouraged. Hence, when followers perceive AL, they feel more confident, flexible and original – in other words, more creative.

Finally, the present findings are congruent with prior studies indicating a positive association exists between AL and individual performance (Clapp-Smith et al., 2009; Walumbwa et al., 2011). Authentic leaders emphasize employees’ talents and strengths, promote followers’ development, and, consequently, improve their performance. The current study’s results thus suggest that followers’ perceptions of AL lead them to engage more frequently in OCB, become more creative, and perform better at work.

5.1.2 Creativity as a predictor of individual performance. The findings also indicate that creativity promotes employees’ individual performance, which is consistent with other studies (Gilson, 2008; Lim and Gilson, 2013; Luthans et al., 2007). A creative employee has the independence, confidence and awareness to take reasonable risks, producing greater flexibility and openness to new experiences and, consequently, higher individual performance. Employees with higher levels of creativity have better individual performance due to increased cognition, motivation and positive behavior (Luthans et al., 2007). Creative employees are continually looking for challenges, striving to meet set targets, and thereby increasing their performance.

5.1.3 OCB as a predictor of individual performance. According to the results obtained in the present study, OCB has a positive effect on employees’ individual performance, which confirms previous studies’ findings (Dunlop and Lee, 2004; Podsakoff et al., 1997, 2009; Whiting et al., 2008). However, the current findings reveal that only altruism, civic virtue and conscientiousness influence individual performance. Sportmanship and courtesy have no statistically significant effect on individual performance. This may indicate that, when employees are mindful of how their actions affect co-workers (i.e. courtesy) and tolerate the inconveniences and annoyances of organizational life without complaining and filing grievances (i.e. sportmanship), these OCBs do not influence these individuals’ performance. Altruism, civic virtue and conscientiousness have a comparatively greater effect.

The latter may thus imply more active behaviors. Those employees who are more altruistic help co-workers with organizationally relevant tasks or problems. Individuals exhibiting civic virtue participate in their organizations’ political processes (e.g. contribute to discussions or get involved in organizational activities). Conscientious employees engage in behaviors that go beyond minimal requirements, such as punctuality, housekeeping, resource conservation and above average work attendance. These behaviors are also characteristic of those who carry out their job functions better.

5.1.4 Creativity and OCB as mediators of the relationship between AL and individual performance. The results obtained suggest that creativity mediates the effect of AL on individual performance. This mediation is only partial because other mechanisms (e.g. OCB) can also influence the relationship between perceptions of AL and individual performance. Authentic leaders have characteristics that influence positively relationships with followers and their work environment. Followers feel freer and more confident to be creative, and, as a result, they improve their performance. Thus, perceptions of AL appear to motivate employees to be more creative at work and, subsequently, to achieve better job performance.

This study’s results also suggest that OCB mediates the effect of AL on individual performance. Employees’ altruism, civic virtue and conscientiousness fully mediate the effect of AL on individual performance, but sportmanship and courtesy do not have a statistically significant effect on individual performance. Thus, perceptions of AL apparently inspire employees to show more altruism, civic virtue and conscientiousness at work, so this leads them to achieve better levels of job performance.
In summary, authentic leaders promote employees’ positive behaviors. These include helping co-workers, participating in organizational activities, and actively taking on tasks outside their specific job definition, as well as developing new suggestions and creative ideas— in other words, bettering their performance.

5.2 Theoretical contributions
From a theoretical perspective, the present study confirms the results of prior studies that found positive effects of perceived AL on employees’ work output, including creativity (Banks et al., 2016; Li et al., 2014; Rego et al., 2014), OCB (Joo and Jo, 2017; Zaabi et al., 2016) and individual performance (Clapp-Smith et al., 2009; Walumbwa et al., 2011). Therefore, the current results support theoretical findings that suggest the quality of work environments leaders create may play an important role in the extent to which employees develop positive behaviors and a creative spirit.

This study answers the call for further empirical research from those who have argued that more information is needed to understand fully the mechanisms through which AL influences employees’ performance and to expand the nomological network for AL (Avolio and Mhatre, 2012; Leroy et al., 2012). Cooper et al. (2005) emphasize the importance of emerging areas such as AL theory and stress that researchers need to pay careful attention to identifying significant construct outcomes. To date, few studies have examined AL’s relationships with creativity, OCB and individual performance, so the present research sought to integrate these four key constructs into a single conceptual model. In particular, the mediating role of creativity and OCB in the relationship between AL and employees’ individual performance remains underexplored in the literature. Therefore, this study addressed significant research gaps by testing the relevant variables’ mediating effects.

As a result, this study’s findings provide a more comprehensive understanding of how AL can relate to individual performance through creativity, as well as of OCB’s mediating role. Furthermore, as compared with previous research, the current results were obtained using a more rigorously empirical methodology that included collecting data from two sources (i.e. dyads of leaders and followers).

5.3 Managerial implications
From a practical perspective, this study’s findings show the importance of how leaders influence their followers. Practitioners looking to increase employee creativity, OCB and performance can do so by enhancing their AL, while organizations can encourage managers to become more authentic in their leadership styles.

These research outcomes have important practical implications for leaders, which can be summarized, in part, as follows. First, leaders should give their followers the freedom to choose what actions they take because this is vital to developing their creativity and job performance. Second, by seeking out different opinions and visions, managers can be better informed in their decision making. Third, leaders need to admit mistakes that they have made and encourage followers to do the same. Fourth, managers should develop channels of multisource feedback that offer anonymous advice from superiors, peers and followers. Fifth, leaders need to take the time to recognize their followers’ values and beliefs and guarantee that job requirements are congruent with these principles. Sixth, managers should seek to become more aware of their own moral weaknesses when pressured by external influences. At last, leaders need to motivate, guide and give their followers all the support they require to become more creative and productive, as well as better organizational citizens.

From the perspective of human resources management, managers should pay more attention to how they select leaders. The hiring process could include evaluating managers’ performance by paying attention to the authenticity of their behaviors. Organizations can
also prepare ways to promote authenticity in leaders through training, coaching and mentoring programs. Further, to improve employees’ performance, human resources management practices should be developed to stimulate creativity and OCB among employees. Specifically, it is important to implement mechanisms to monitor and recognize employee’s creativity and additional efforts in activities that extend beyond normal expectations. Therefore, organizational systems such as compensation and promotion may be aligned with that. All these interventions should take into account the context in which leaders and followers interact (Avolio, 2007), including the organizational culture of the company. This should value the authenticity in the relationships between all organizational members and the development of AL processes.

5.4 Limitations and future studies
The present study was subject to some limitations that indicate opportunities for future research. The first limitation was convenience sampling, which limits the generalizability of the results. Future research along these lines needs to obtain larger, more diversified samples to achieve better generalizability. In addition, this study was based on a cross-sectional design, so any causal influences suggested by the findings cannot be treated as conclusive. To examine causal relationships, a longitudinal or experimental design would be needed to ensure a deeper understanding, which suggests the need to conduct further research in this field.

The respondents in this study are from a single country. Future studies could investigate if the current conclusions can be replicated for other nationalities, taking a more cross-cultural approach to the issues in question and considering the broader cultural context in which leaders and followers interact (Avolio, 2007). The studies on AL developed so far in different national cultures suggest that the AL construct has a cultural equivalence (e.g. Caza et al., 2010; Walumbwa et al., 2008), but more research using multi-culture data is needed to fully demonstrate this (Caza et al., 2010).

The research also included two mediating variables, but others are also plausible. For example, authentic leaders may stimulate followers’ hope, optimism, resilience and happiness, which, in turn, increase employee performance.

Further research may be needed to analyze other relevant dependent variables with negative effects on employees’ performance, such as stress, burnout, or turnover. Future studies could also test the incremental validity of AL theory for other positive forms of leadership such as shared, transformational, ethical, empowering and leader-member exchange leadership. Despite the above limitations, this study’s findings contribute significantly to expanding the field of positive organizational research, thereby deepening the understanding of the influence of AL on employees’ behavior and performance and opening up new avenues of research.

5.5 Concluding remarks
The study responds to a call by researchers who have argued that more empirical research is needed to understand fully the mechanisms through which AL influences employees’ performance and other outcomes and to expand the nomological network for AL (Arda et al., 2016; Avolio and Mhatre, 2012; Gardner et al., 2011; Leroy et al., 2012; Mubarak and Noor, 2018). This research also represents an important opportunity for legitimizing the authenticity in leadership, in both the scholarly and practitioner context. Although AL is seen as a core value in itself, without empirically demonstrated pragmatic outcomes, it is less likely to obtain attention in both contexts. Considering that creativity and OCB promote employees’ performance, our study suggests that acting authentically, leaders are more able to stimulate important outcomes, such as, creativity, OCB and performance.
References


Avolio, B. (2005), Leadership Development in Balance: Made/Born, Lawrence Erlbaum, Mahwah, NJ.


About the authors

Neuza Ribeiro is Professor of Human Resources Management and Organizational Behavior at the School of Technology and Management, Polytechnic Institute of Leiria, Portugal. She received a Post-doc in Management and a Doctorate in Industrial Management with a specialization in organizational behavior from Aveiro University, in 2010. Her research focuses broadly on human resources management and organizational behavior. Neuza Ribeiro is the corresponding author and can be contacted at: neuza.ribeiro@ipleiria.pt

Ana Patrícia Duarte earned a Doctorate in Social and Organizational Psychology from the Instituto Universitário de Lisboa (ISCTE-IUL) in 2011. She is currently a postdoctoral research fellow at the same university, and her research focuses broadly on management, organizational behavior and corporate social responsibility.

Rita Filipe received a Master of Management Control from the School of Technology and Management, Polytechnic Institute of Leiria, in 2013. She currently works with investment projects in mould industry. Her primary research interests include leadership, commitment, creativity and performance.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Four facets of learning in performance measurement

Jason Martin and Mattias Elg
Department of Management and Engineering, Linköping University, Linköping, Sweden, and
Andreas Wallo and Henrik Kock
Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

Abstract
Purpose – The purpose of this paper is to contribute to a better understanding of the role of learning in performance measurement.
Design/methodology/approach – The authors develop a theoretical framework combining workplace learning theory with purposes of performance measurement. The authors elaborate this framework empirically by identifying critical incidents from a case set within a context containing a broad range of different performance measurement activities. Finally, the authors discuss the results and the possible implications for using the theoretical framework in order to better understand facets of learning regarding the design of performance measurement.
Findings – Workplace learning theory provides a deeper understanding of how the mechanisms of performance measurements support control or improvement purposes. The authors propose a tentative framework for learning as a driver for performance measurement and four facets of learning are identified: reproductive, rule-oriented, goal-oriented and creative learning.
Research limitations/implications – The empirical material is limited to the healthcare context and further studies are needed in order to validate the findings in other settings.
Practical implications – The authors argue that all managers must consider what kind of learning environment and what kind of learning outcomes best serve the interests of their organisation. Purposeful and carefully designed organisational arrangements and learning environments are more likely to induce intended learning outcomes.
Originality/value – Previous connections between the fields of “performance measurement” and “workplace learning” often lack any deeper conceptualisations or problematisations of the concept of learning. In this paper, the authors provide a more nuanced discussion about the process of learning in performance measurement, which may provide a basis for further research and scholarly attention.

Keywords Performance measurement, Control, Improvement, Workplace learning, Adaptive learning, Developmental learning

Paper type Research paper

Introduction
A number of challenges caused by globalisation, digitalisation and the changing nature of work have influenced how most organisations operate and improve their performance. There is widespread agreement that performance measurement is one way to handle these challenges, and is therefore a core element of modern management in any business (Bititci et al., 2012; Neely, 1999; Pollitt, 2013). This broad interest emanates from the central management problem of increasing efficiency as well as effectiveness (Neely et al., 2005).

The versatility and ability to use performance measurement for different purposes is striking. It plays a key role in developing strategic plans, prioritising strategies, evaluating achievements of objectives and setting the stage for organisational improvements. An important feature of performance measurement, cutting across any context or purpose of use, is its potential to create different conditions for learning within organisations (cf. Bohner, 2009; Chenhall, 2005; Lucas, 2015; Wang, 2018). The notion of learning in the context of performance measurement has been proposed in relation to a broad range of topics, such as manager development (cf. Appuhami, 2017; Lebas, 1995; Sun and Henderson, 2016),
as a means for employees to understand their business (cf. Nudurupati et al., 2011; Taylor and Taylor, 2014), stimulation of problem-solving (Mausolff, 2004), and critical reflection on strategy (cf. Santos-Vijande et al., 2012; Neely and Al Najjar, 2006). Although there is consensus about the importance of learning in performance measurement (cf. Aranda et al., 2017; Franco-Santos et al., 2012; Melnyk et al., 2014), a shift towards an increased focus on learning can be seen as a key challenge for actual performance measurement practices (cf. Aragon et al., 2014; Bititci et al., 2012).

There are numerous examples of theoretical frameworks encompassing both learning and performance measurement (cf. Aragon et al., 2014; Brudan, 2010; Zhou et al., 2015). Although these studies provide valuable insights, they do not further problematise and elaborate on the concept of learning (cf. Santos-Vijande et al., 2012; Fang et al., 2016). Empirical, qualitative studies extending a deeper understanding on learning processes in relation to performance measurement practices are scarce. By drawing on theories from workplace learning (cf. Argyris and Schön, 1978; Billet, 2001; Engeström, 1987; Ellström, 2011; Tynjälä, 2013), we want to explore the diverse and complex ways in which learning is embedded in performance measurement practices. The complexity of learning is manifested in its duality as constituting both a process and an outcome. Therefore, it could be claimed that the content of performance measurement is defined by learning and vice versa.

The purpose of this study is to understand this duality and explore the role of learning in performance measurement. We aim to develop a framework in which two main contrasting purposes of performance measurement – control and improvement – are connected to four different facets of learning processes in organisations: reproductive, rule-oriented, goal-oriented and creative learning. A key issue is to understand what characterises learning in situations when performance measurement facilitates control (i.e. where there is a need for efficient, routinised work), and what characterises learning when performance measurement drives creative and developmental actions.

We extend our investigation in four main steps. First, we present previous research and a theoretical framework. Second, we propose a tentative framework for learning processes in performance measurement. Third, we present data from a case study set within a performance measurement rich healthcare context. In the final step, we discuss the framework and present implications in using performance measurement for learning in organisations.

**Previous research**

A recurring theme within performance measurement literature is to define and understand the different purposes and roles of performance measurement in organisations (cf. Bond, 1999; Elg, 2007; Neely, 1999; Franco-Santos et al., 2007; Pidd, 2012; Serrat, 2017). Behn (2003) described eight main purposes of performance measurement: evaluate, control, budget, motivate, promote, celebrate, learn and improve. Although the studied literature offers an array of different performance measurement purposes, two overarching and governing “meta-purposes” that can be readily identified and control and improvement. A central problem in the use of performance measurement is that it “just keeps the score” (Bourne, 2008, p. 68), instead of facilitating change. Therefore, we stress a need to explore learning purposes and improvement in performance measurement practice.

Early examples of literature incorporating learning and performance measurement often adopt a strategic and systems-oriented perspective (cf. Bourne et al., 2002; Dervitsiotis, 2004; Neely et al., 2005; Neely, 1999; Chenhall, 2005; Franco-Santos et al., 2007). Recent studies have continued the strategic and systems-oriented approach (cf. Appuhami, 2017; Fang et al., 2016; Zhou et al., 2015). Some studies have challenged the traditional control perspective on the use of measurements and offered an alternative view on performance measurement systems as “a means of challenging assumptions” (Neely and Al Najjar, 2006, p. 102) and also as a learning tool (Canonico et al., 2015; Dossi and Patelli, 2010).
Most studies on learning and performance measurement are quantitative, with few qualitative or conceptual studies more elaborately integrating the processes and practices of learning and performance measurement. Learning is often addressed only briefly and most studies lack a deeper understanding and analysis of how learning theory and learning processes affect performance measurement and vice versa. Also, several studies have adopted a somewhat instrumental view, presuming a causal chain with learning supporting or facilitating the use of performance measurement systems and rarely, if at all, the other way around (cf. Aragon et al., 2014; Fink et al., 2017; Hussain et al., 2016). Few empirical studies have described the particular learning process or elaborated how learning processes and performance measurement practices interact. Even fewer have extended a deeper understanding of what actually constitutes learning in performance measurement systems, or indeed the very nature of the particular learning processes contained therein. According to Behn (2003), learning and performance measurement are about determining what is working, or not working, and then figuring out how to make improvements based on these new insights. The present study is an attempt to increase knowledge about the relation between “determining what is working, or not” and the nature and quality of the “figuring out”.

Based on the examined literature, two main conclusions can be discerned. First, the often piecemeal handling of the concept of learning, in relation to performance measurement, implies that learning is perceived as something of a “black box” that is often left without any further problematisation or extended discussion. Questions about what is learnt and by whom are rarely addressed and previous research has made few attempts to distinguish between different meanings of the concept. Second, there appears to be a gap in the performance measurement literature on the needs of managers to better understand learning processes in order to utilise performance measurement to its full potential. We argue that managers need to have an equal understanding of both performance measurement and learning processes in order to optimise them both in successful organisations.

Theoretical framework
To open up the “black box” of learning, we will employ a workplace learning perspective (Billet, 2001; Engeström and Kerosuo, 2007; Ellström, 2011; Tynjälä, 2013) to discuss learning in relation to performance measurement. The field of workplace learning research has grown extensively and a large number of interdisciplinary studies have been carried out in this area (Manuti et al., 2015). Rapid economic, social and cultural changes have led organisations to realise the importance of issues concerning workplace learning and employee development (Tynjälä, 2013). Therefore, the workplace is no longer only seen as a site for production, but also as an environment for facilitating learning (Ellström, 2001).

From a workplace learning perspective, learning implies both formal training activities (such as courses) and informal learning. The latter refers to learning as a continuous activity that is inherently integrated into all work activities (Zuboff, 1988). Theoretically, this distinction between formal and informal learning is parallel to Sfard’s (1998) notion of learning as acquisition and learning as participation. As learning is truly at the heart of any productive activity (Zuboff, 1988) and productive activity incorporates performance measurement, learning and performance measurement appear to be truly inseparable. Thus, it could be argued that learning forms a point of departure for just about any notion of performance measurement and is not, as Behn (2003) stated, a separate end exclusive performance measurement purpose. Therefore, learning can be said to be not just an outcome but perhaps also be a prerequisite for an activity, or a part of the activity itself (Antonacopoulou, 2006; Tynjälä, 2013).
There are several similar ways of conceptualising different types of learning. Argyris and Schön (1978) elaborated the influential concepts of single-loop and double-loop learning. Engeström (1987) developed expansive and restrictive learning, and March (1991) proposed a difference between exploration and exploitation in organisational learning. In this paper, we draw mainly on Ellström’s (2010) distinction between adaptive and developmental learning, which builds on, and, to some extent, also elaborates these previous conceptualisations of learning in the workplace.

Adaptive learning is obtained through routine action, where rules and instructions are followed for known problems or situations. This type of learning is primarily instrumental and is valued when it promotes effective action, leading to a learning environment that is characterised by security, standardisation, exploitation and consensus. Thus, adaptive learning happens on a more routine basis, often aiming to improve something that has already been mastered. Developmental learning, by contrast, is characterised by reflective and alternative thinking, risk-taking, critical reflection and the desire to experiment (Ellström, 2010, 2011). Uncertainty and divergence are seen as potential generators of exploration rather than as threats or inconveniences. According to Ellström (2011), these two learning types should be perceived as complementary.

The question of whether learning is adaptive or developmental requires a discussion on the conditions in the workplace that may enable or constrain learning; that is, the learning environment (Coetzer, 2007; Fuller and Unwin, 2004; Kock and Ellström, 2011). This learning environment comprises conditions provided by an employer, such as the learning potential of the work tasks, the work organisation and the available learning resources. Subjective factors that also seem to affect the individual’s engagement in learning activities include self-confidence, motivation, values, previous learning experiences and learning readiness (Billet, 2001). Engaging in a learning process is not automatic (Tidd, 1997) and there may also be specific problems that affect engagement in learning activities. In their study on learning, learning networks and shared learning in organisations, Bessant and Tsekouras (2001) listed seven main learning blocking factors together with their underlying problems: lack of entry to learning cycle, incomplete learning cycle, weak links in the learning cycle, lack of learning skills or structure, knowledge remains in tacit form, repeated learning (without capturing or codifying the learning content) and infrequent, sporadic and unsustained learning.

When it comes to learning in performance measurement, it is particularly interesting to explore the learning potential of the task, and especially the discretion that the task entails. Discretion refers to the degrees of freedom, or scope of action, that the learning subject has regarding how the task is defined, how the methods for solving the task are chosen and how the results should be evaluated (Ellström, 2001). Hackman (1969) listed understanding, acceptance, idiosyncratic needs and values, as well as previous experience with similar tasks, as factors that affect the motivation, performance and outcome (or results) of tasks. In combining these factors – task, method and results – Ellström distinguished four levels or, as we prefer to label them, facets of learning (see Table I).

<table>
<thead>
<tr>
<th>Aspects of the learning situation</th>
<th>Adaptive learning</th>
<th>Developmental learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tasks</td>
<td>Given</td>
<td>Given</td>
</tr>
<tr>
<td>Methods</td>
<td>Given</td>
<td>Not given</td>
</tr>
<tr>
<td>Results</td>
<td>Given</td>
<td>Not given</td>
</tr>
</tbody>
</table>

**Table I.** A taxonomy of facets of learning as a function of the discretion that exists in the learning environment.
By using facets rather than levels, we embrace the notion that learning is not a hierarchical phenomenon in which “higher-level” learning processes presuppose superiority over “lower-level” learning processes. We adopt the view that different contexts and settings induce different learning processes, sometimes even simultaneously.

In Table I, reproductive learning represents a facet of learning that occurs when the tasks, methods and results are pre-determined. This facet of learning is sufficient and necessary for many activities, but its primary function is in the formation of competencies for automatically solving routine assignments. Reproductive learning primarily supports the accumulation of tacit knowledge (Polyani, 1967) within individuals. An example of reproductive learning situations could be the use of routinised day-to-day measurements serving a documented purpose used to adapt or refine existing routines. The next facet of learning is called productive learning and it comes in two slightly different types: rule-oriented and goal-oriented. Rule-based productive learning is characterised by a higher degree of discretion when it comes to the evaluation of results. The learning subject must evaluate the outcomes and make minor changes to the methods used to solve the task. One example of rule-oriented learning is the use of routinised day-to-day measurements that serve a documented purpose but are used to develop new ways of working. As well as having a higher degree of discretion regarding results, goal-oriented productive learning has a higher degree of choice and use of methods. So, even though the task may be given, the learning subject can experiment with ways of solving the problem at hand. One example of goal-oriented learning is adhering to documented and previously designed purposes but allowing for new approaches in the design of measurements and utilisation of results.

In the last facet, creative learning, the learner must use his or her own authority not only to evaluate results or chose methods, but also to define the task itself; that is, to diagnose the situation. In this facet of learning, the learner is free to question the definitions of tasks and problems posed by the environment (such as colleagues or management) and to act to transform standardised or institutionalised solutions. An example of creative learning could be an environment in which individuals or groups of individuals within an organisation are entrusted to freely reflect on and question established ideologies, devise own measurements and develop new routines, structures and practices.

Towards a tentative framework for performance measurement as a driver for learning

Based on our theoretical perspective on learning, we propose a tentative framework as a first attempt to build a theory of various forms of learning in contexts of performance measurement (see Figure 1). The elements of the proposed framework are based on four components: the purpose of performance measurement, the learning type, the learning facet, and the level of agent discretion.
facets and the level of agent discretion. The framework is intended to be used in associating learning with performance measurement.

The purpose of performance measurement plays a fundamental role in management. A key assumption in performance measurement is that organisational activities are best controlled and improved through a carefully set of measures that drives performance (Melnyk et al., 2014). Efficient and effective operations management rely on this mechanism (Neely et al., 2005). However, purposes of performance measurement are rarely clear-cut and rarely serve purely either control or improvement purposes. Therefore, with our framework, we suggest that measurement for control and measurements for improvement can include both learning types and the different learning facets as described below.

The second component is learning type (Ellström, 2010, 2011). Within adaptive learning type contexts, learning from performance measurement is primarily valued if it can contribute to operational efficiency. There is a strong imperative for reliability and learning is strongly connected to the task of reducing variation in operations (Ashton, 2004; Ellström, 2011). With a predominantly adaptive learning type, learning from performance measurement can be said to focus on eliminating errors and aiming for organisational compliance with operational routines and rules. Within adaptive learning, the context and culture supports learning conditions aimed at enhancing and making current operations more efficient (Ellström, 2001; Barley and Kunda, 2001). By contrast, within developmental learning type contexts, the context supports exploration of variation. Any deviation or divergence from routine is viewed as an opportunity for knowledge development and as an incentive for incremental improvements and, in some cases, even radical improvement. In such situations, failures are accepted and a certain level of risk-taking is encouraged. The most critical component in a developmental context is organisational acceptance and encouragement for critical thinking and critical reflection (Argyris and Schön, 1978; Ellström, 2001; March and Olsen, 1976).

The third component in the framework consists of the four learning facets (Ellström, 2001). The continuum of reproductive, rule-oriented, goal-oriented and creative learning is situated along the same axis as the adaptive and development learning type axis. Thus, contexts within a predominantly adaptive learning type would support reproductive and rule-oriented learning to a higher degree, while contexts within a predominantly developmental learning type would be better suited to support goal-oriented and possibly also creative learning.

The fourth and final component in the framework – the level of agent discretion in relation to tasks, methods and results (Ellström, 2001) – is outlined vis-à-vis the other components. As used here, “agent” is an intentionally acting subject or collective that occupies a structural position (Danermark et al., 2002). Agent discretion, in the form of bounded conditions for participation in operational and organisational development, has been widely addressed in the literature (cf. Ellström, 2001; March and Olsen, 1976; Norros, 1995). Discretion not only requires the individual agent to be able and willing to participate but also to have the right to do so, in the form of organisational acceptance and encouragement (Kock, 2007). Agent discretion will differ for the agents involved in performance measurement activities. In more adaptive learning type contexts, the level of discretion can be characterised as low to medium, and more developmental learning type contexts as medium to high.

Tasks refer to directives or demands that express needs, desired outcomes and intended purposes of the performance measurement (cf. Ellström, 2001; Hackman, 1969). Examples include targets, rules, regulations, orders, professional needs and/or insights. Methods specify the performance measurement methods or practices to be utilised. These can be routine-based, “one-off” designs and/or experimentally developed. Results are performance measurement outputs that can be specified with clear intentions, or developed according to un-foreseen contingencies or arising needs.
Research method
The data reported in this paper were collected in a research project based on case study methodology (Eisenhardt, 1989; Voss et al., 2002). Our case was an orthopaedic and rheumatologic department located in southeast Sweden, a context in which performance measurement is widely used. The department has about 150 employees in nine sections, treating both planned and acute patients. Each year, approximately 30,000 patients visit the department and approximately 3,000 surgeries are performed. The department is considered to be best-in-class when it comes to utilisation of performance measurement in both production and developmental work.

In line with Voss et al. (2002), the starting point of the present work is that it is necessary to have a rich amount of primary data in order to make sense of the key dimensions and the types of environments that frame learning in performance measurement. We consider this case as instrumental (Stake, 1995). Thus, we are not primarily interested in its intrinsic qualities, but rather in how it enables us to understand and develop our theoretical ideas through analytical generalisation (Firestone, 1993). In previous studies, we have reported the studied organisation’s wide variety of activities linked to performance measurement (reference anonymised for blind review). These include higher management’s strategic initiatives, strategic goal deployment, continuous follow-up in management meetings, improvement work and individual initiatives for decision making. By and large, this provides us with a single case with great potential for studying many variations in which learning maybe linked to performance measurement. Thus, we use the case for theory building in the sense that the case supports our work in establishing links between various forms of learning and performance measurement (Voss et al., 2002).

When conducting empirical field work, one of the central problems is grasping real-world activities. This is especially difficult in situations where primary data are collected through interviews. Silverman (2004) argued that it is important for the researcher to acknowledge that leading actors in a studied organisation are interested in representing themselves as goal-oriented, trustworthy, change-oriented and rational. This needs to be taken into account when studying performance measurement. Therefore, general statements by top leaders are of limited value, especially if an organisation is best-in-class. One strategy that we used to deal with this is to collect information from a large range of respondents in the organisations (Alvesson and Sköldberg, 2009). Another strategy is to allow the respondents to recall specific situations from which the dialogue can evolve. This was done through the critical incident (CI) technique (Flanagan, 1954), which helped us identify critical events related to the application of performance measurement among interviewees. In the interviews, respondents were asked to elaborate on incidents with respect to: what was going on in this situation; the types of performance measurements being used; who was involved; and short-term results and the perceived long-term effects. All interviews were recorded and transcribed. A total of 73 CIs were extracted from the empirical material. The CIs were first reported in a previous study (reference anonymised for blind review).

As discussed above, the respondents for the study were selected in order to maintain a high degree of variation in CIs. The following respondent groups within the orthopaedic unit were included: medical doctors (MD) (n = 3), nurses (n = 3), assistant nurses (n = 2), care unit managers (n = 2, also with background as nurses), financial manager (n = 1), administrative unit manager (n = 1), administrators (n = 2), the clinical department manager (n = 1, also MD), organisational developer (n = 1, physiotherapist) and physiotherapist (n = 1). In total, 17 respondents were interviewed.

We viewed each CI as a specific type of situation (i.e. learning environment) that creates various conditions for learning in performance measurement. The analytical procedure followed a structured approach in which our theoretically proposed framework guided the coding of each incident. This helped us validate and make further refinements of our key dimensions and definition of types of learning environments that determined facets of learning.
Results

Critical incidents
The analysis of the total number of CIs ($n = 73$) revealed 55 incidents in which learning environments could be adequately categorised. Table II lists the four facets of learning, each accompanied by two representative CIs describing the particular learning environment. The analysis of each CI concerning discretion (the CI nature of task, methods and results) is outlined briefly. We have also extracted a selection of quotes from the data in order to further illustrate how the particular facet of learning is related to the specific performance measurement.

Of the 73 CIs, 55 comprising the empirical material could be plotted in the framework. These 55 CIs could further be elaborated and categorised in accordance with the particular learning facets each of the supported CIs. Ten CIs (18.2 per cent of the total) could be categorised as performance measurement supporting creative learning. Eight CIs (14.5 per cent) were categorised as performance measurement supporting goal-oriented learning. In total, 17 CIs (30.9 per cent) were categorised as performance measurements supporting rule-oriented learning and 20 (36.4 per cent) as supporting reproductive learning.

Reproductive learning
The facet of reproductive learning is reflected in the prescribed nature of the tasks, methods and results within the performance measurement context. The CIs reveal examples with tasks specified according to routines and/or management directives (e.g. CIs no. 10 and 11). Performance measurement is initiated as an organisational management action in order to follow standards and regulations, which are either locally or centrally governed. The methods are managed, designed and implemented on a top-down basis and the desired results are set, often in accordance with the task. The staff have little, if any, influence or choice regarding what, how and when the performance measurements should be used.

Rule-oriented learning
For rule-oriented learning, the CIs reveal a higher agent discretion level concerning results and outcomes. The effects and outcomes of performance measurements are not specified or defined in any detailed sense. CI examples reveal dynamic team discussions where alternatives for solutions are discussed before decisions are made, and there is also room for individual influence within a team context (e.g. CIs no. 8 and 18). However, the alignment between task and method remains, and the examples reveal adherence to routines and pre-designed measurement methods such as complying with requirements and using pre-set screening methods.

Goal-oriented learning
CIs with goal-oriented learning show performance measurements that are usually decided “top-down”, specifying particular tasks, such as a specific target or goal to achieve. However, the choice of performance measurement method and the responsibility for results are delegated to the staff. One featured example (CI no. 9) describes direct and defined problems, as defined by a pre-determined quota; in this case, the number of faulty telephone directions made by the automatic system. The choice of measurement design and execution and how to handle the results is left to the staff. In CI no. 19, the task is pre-defined but the method and results are conjointly devised within the teams, with high individual and team discretion. Both of these CIs have little alignment among task, method and result.
<table>
<thead>
<tr>
<th>Facet of learning</th>
<th>Performance measurement content</th>
<th>Analysis of discretion level (task, method, results)</th>
<th>Illustrative quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reproductive (Task, method and results given)</td>
<td>Measures of waiting times (CI no. 10)</td>
<td>T: how to comply with documented routine and regulations M: routine referrals compared to MD schedule R: matching procedure</td>
<td>“[... ] then we measured the timings [waiting times] and then I had the schedules to compare and developed a way to comply with the target goal of three months [... ]”</td>
</tr>
<tr>
<td></td>
<td>Staff hygiene measures (CI no. 11)</td>
<td>T: higher-level orders M: pre-determined self-assessment routine R: measures used to reach pre-defined target goals</td>
<td>“I don’t remember how many questions it is but you have to self-assess whether you disinfect your hands before putting on gloves /.../ to check our dress, how we look, if”</td>
</tr>
<tr>
<td></td>
<td>Rule-based hospital time measures (CI no. 8)</td>
<td>T: decrease post-surgery hospital time M: number of patients leaving within a pre-set time + follow-up call + return rate R: instead of anticipated results, un-foreseen results appears</td>
<td>“/.../ we were told to call the patient at home on day seven or a follow-up [... ] and this was good because at first there was some resistance /.../ but now when we measure, we see that not many at all return”</td>
</tr>
<tr>
<td></td>
<td>Observation of hygiene procedures (CI no. 18)</td>
<td>T: how to comply with documented target (levels of hygiene) M: established measures R: team development of new routines and procedures</td>
<td>“I get a report on the routine observations every month/.../I always distribute the reports in our meetings and let the staff discuss observations and possible actions”</td>
</tr>
<tr>
<td>Goal-oriented (Task given. Method and results not given)</td>
<td>Measures of automatic telephone system (CI no. 9)</td>
<td>T: fix automatic telephone system directing problems M: task analysis led to development of measuring method: survey R: efforts made to change Imp-temp scheduling</td>
<td>“So, we created some simple measurements [... ] afterwards, we sat down in our little group of colleagues and went through the figures and found common issues and patient needs”</td>
</tr>
<tr>
<td></td>
<td>Team performance follow-up (CI no. 19)</td>
<td>T: reduce long hospital times M: benchmark and collective analysis of factors affecting hospital time R: contingent problem solution cutting across professional areas</td>
<td>“And then we sat around the table and analysed what affected hospital times [... ] we looked into what to change and then we returned to our different worlds and changed”</td>
</tr>
<tr>
<td>Creative (Neither task, method nor results given)</td>
<td>Measures of emergency patient re-visits (CI no. 12)</td>
<td>T: experienced problem triggers necessary contingency M: emergent analysis of adequate measurements (statistics) in collective meeting. Also scanning journals and time availability R: contingency-dependent prioritisation</td>
<td>“So I figured out issues with availability [... ] we [the team] ran through the available statistics and then we were allowed to do our priorities /.../ we created a very good booking system”</td>
</tr>
<tr>
<td></td>
<td>Deep infections (CI no. 46)</td>
<td>T: a realisation of an increase in deep infections M: task awareness initiates investigation. Hypotheses development R: contingency-dependent action; change of procedure</td>
<td>“Then I started to ponder how we really do the sterile dressing and did an evaluation of that. It was due to the figures in these discoveries that we understood we had to do something”</td>
</tr>
</tbody>
</table>
Creative learning

In creative learning, the agent has full discretion over the task, method and results. Our featured CI examples (CIs no. 12 and 46) reveal a high degree of integrity and responsibility in the definition of the task, the design of the method, and the handling of results. In CI no. 12, an emergent problem (booking re-visits) triggers the collective choice of a particular task (better planning) and the joint analysis of journals and schedules in order to handle the situation. The result ("prioritisation") is contingency-dependent and highly flexible. The same levels of discretion within task, method and result are also mirrored in CI no. 46 and assume a high degree of responsibility for the entire performance measurement process. From an agent perspective, our identified case CIs reflecting creative learning also show a dominance of highly qualified professionals (such as MDs) who can be assumed to have strong professional legitimacy and authority in the organisation, granting liberties but also high levels of responsibility and freedom. The structural context may be assumed to contribute to agent discretion, providing conditions for higher learning readiness and developmental initiatives.

Notes on the absence of learning conditions

Contextual and/or individual factors inhibited or obstructed observable conditions for learning in 20 of the CIs. Descriptive examples of CIs where very limited conditions for learning and/or improvement could be found include CI no. 38, in which the sheer numbers of quantitative performance measurements produced within the studied organisation creates an overload of information, which led to lost opportunities for learning. CI no. 54 describes how the visualisation of quantified performance measurements is not enough to create the conditions necessary for any kind of learning and/or improvement. In CI no. 68, MDs describe a particular performance measurement as a nuisance, which irritates the staff and discourages any kind of conditions for learning.

Summary of results

Two-thirds of the CIs containing identifiable learning processes were of an adaptive nature (i.e. either reproductive or rule-based) and the remaining one-third were developmental in nature (i.e. goal-oriented or creative). Table III shows the distribution of the four different learning facets found in the CIs within the different personnel categories. Our results, as displayed with our featured examples, show that performance measurement is achieved in many different forms and that the nature of the learning is highly reflected in the design, context and execution of performance measurement. We further show that the degrees of freedom induce certain types of learning and vice versa and that agent discretion plays an important role in determining the facet of learning. The continuum from control to improvement is conditioned by three aspects characterising the workplace: task, method and results.

Table III shows that many different groups of professionals are involved in the utilisation of performance measurement. Most of the CIs were from the nurses and other care personnel group. Also, the managers and administrator group, as well as medical doctors, were linked to many CI. These three groups were involved in all of the four different learning situations, albeit with a general tendency towards more adaptive processes. It is noteworthy that MDs score relatively high on creative learning.

<table>
<thead>
<tr>
<th></th>
<th>Reproductive</th>
<th>Rule-based</th>
<th>Goal-oriented</th>
<th>Creative</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD</td>
<td>5</td>
<td>11</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Nurses and other care personnel</td>
<td>12</td>
<td>15</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Managers and administrators</td>
<td>14</td>
<td>4</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Organisational developers</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>31</td>
<td>17</td>
<td>22</td>
</tr>
</tbody>
</table>

Table III. The distribution of learning facets within CIs amongst studied personnel categories.
Discussion

The present study sheds light on learning processes that occur over a wide range of workplace situations in which performance measurement are being used. In line with this, the purpose of this study was to contribute to a better understanding of the opportunities for learning found in the practice of performance measurement. In utilising the established workplace learning theory, we proposed a tentative framework for learning as a driver for performance measurement. Four facets of learning were identified: reproductive, rule-oriented, goal-oriented and creative learning.

Our empirical investigation showed that it is possible to map performance measurement activities according to the learning conditions they create or support. The findings presented herein suggest that research on management benefits from theory development involving theories and concepts from other, related research fields. The workplace learning literature pointed to key issues interconnecting needs for learning and improvement with performance measurement.

While problems with learning from performance measurements may occur, we observed that some kind of learning always takes place, although it may not be as intended or serving stated organisational purposes. As Ellström (2010) pointed out, the processes of learning, change and improvement do not always need to be positively related. If there is a lack of learning skills, unsupportive structures or if the motivation or learning skills are absent (Bessant and Tsekouras, 2001), learning can lead to unintended learning outcomes. The conscious design of performance measurements could be viewed as a way of overcoming possible problems in the creation of the environment (Gibson, 1950, 1966). Thus, the individual perception of the context as providing a means for learning and also creating motivation for learning forms a powerful tool in successful operations management aiming for continuous improvement.

Discretion or control (Dworkin, 1978; Ellström, 2001, 2011; Karasek, 1979; Weststar, 2009) provides a valuable concept that supports understanding of how certain facets of learning occurs in performance measurement practice. Task understanding (Hackman, 1969), as well as technical and social factors (Karasek, 1979; Weststar, 2009), affects the individual discretion. This influences job control; that is, the degree of freedom and roles for people in (re-)defining and deciding on the task, methods and results within their professional context (Dworkin, 1978). According to Taris et al. (2003), increasing job control has a positive effect on learning. This would indicate that empowering staff members and increasing discretion, with more individual control over tasks, methods and results, would enable more influence over the learning situation, resulting in more productive type II (creative) learning outcomes. On the other hand, lower discretion levels, with performance measurements having pre-determined tasks, methods and results, would constrain the possibilities to influence or change the aspects of learning, resulting in more productive type I (reproductive) learning situations.

Within these settings, an important distinction is whether performance measurement is used for control purposes, where the actor degree of freedom is low or whether performance measurement is used for improvement purposes, in situations with possibilities for actors to make decisions about how to do things and sometimes even what to achieve. One example of the former is when the studied healthcare organisation is exposed to performance measurements targeted at compliance with county council regulations (CI no. 11). This is a situation in which the healthcare personnel need to focus on doing the job; that is, following the pre-determined hygiene standards with very low levels of discretion. The latter purpose may be exemplified through the creative job of identifying the causes for a sudden increase in infections (CI no. 46). This situation required a lot of new thinking, team efforts and root cause analysis of the central problems. In this situation, tasks, methods and results were defined by the team that solved the problem – a situation with high discretion. In between these two activities, we
find CIs that enable some degrees of freedom and some discretion for the parties involved. We illustrate this in the results section using the example of an incident of compliance with time for new visits where the team had freedom in how to investigate this problem and what the solution could be (CI no. 9). However, the task was defined by the standard rules for compliance (in this particular CI, a maximum of 90-days waiting time). So, the continuum of learning can be expressed in terms of whether the task, method and results are determined in advance. That will give the healthcare personnel various opportunities and degrees of freedom in terms of how they solve problems and learn from situations.

In situations of goal-oriented and creative learning, participants are more involved in determining and interacting with the data. In such situations the aim is to make sense of and obtain a feeling for the data. Competence in using performance measurement in those situations is, arguably, very important. Participants need to be able to use existing measures in new ways or develop new measures for the specific task at hand. They also need to be aware of the various ways in which data may be represented and analysed. Thus, goal-oriented and creative learning as a driver for improvement seems to be reliant on active engagement and competence among participants.

We found similarities with Canonico et al. (2015) in that performance measurements used in contexts of control often entail a low interaction with data and primarily reflect management directives. In these situations, measurements are used in a top-down fashion, mainly to instruct participants about what is important, what the gap is and what is needed to close the gap. For instance, the instructions for compliance with pre-surgery hygiene routines are quite clear and there is not so much room for variation in terms of how to carry out this activity. The learning subjects only need to know that they should follow the instructions that are based on evidence from research.

A major difference between performance measurements used for control and for improvement is “contact” and discretion to utilise data. When the degrees of freedom for various forms of actions increase, performance measurement becomes one’s own learning and makes it possible to make sense of the data. In this context, the designer of the performance measurement is also the interpreter of the information. All this creates a robust process for creative knowledge creation. A special case resulting in creative learning occurred when the method was given but the task and results were not. This particular situation was not featured by Ellström (2001) but is exemplified in CI no. 12 (see Table II) and also found in CIs no. 35, 55, 56, 66, 72 and 73. The task is realised by the individual who utilises available data from earlier performance measurement activities and acts according to his or her own professional discretion in order to achieve desired results.

On the other hand, when the measurement is used for control purposes, the pre-conditions are sometimes the opposite. Few of the people involved in the control context have been involved in the design and development of knowledge. Our study also indicates that learning facets may interchange, back and forth, within a particular performance measurement activity, sometimes even allowing simultaneous learning levels. Therefore, the design of performance measurement requires continuous reflection and conscious balancing of learning facets (Ellström, 2011).

Another finding for future studies to explore is the importance of legitimacy. Legitimacy refers to the lawful gain or collective acceptance of power to exercise authority within organisations. Our results suggest that agent discretion and possibilities for creative learning seem to be affected by the organisational legitimacy of the involved actors. A majority of the respondents that describe performance measurement as supportive of creative learning were MDs, team leaders or organisational developers; that is, relatively senior members of organisational groups that typically enjoy superior access structures (March and Olsen, 1976) and professional or positional legitimacy (Dowling and Pfeffer, 1975). This finding raises the question of how organisations can empower
broader staff groups, thereby increasing legitimacy and the possibility for a wider range of actors to utilise performance measurement for developmental learning.

Performance measurement design has the potential for greater impact on improvement efforts in operations management. Mapping operations management practices (such as performance measurement) and the different conditions for learning give practitioners the opportunity to harness and manage inevitable learning participation (Sfard, 1998) and increase the potential for organisational improvement. Understanding the workplace learning theory in performance measurement also helps to detect dysfunctional learning environments, offering the possibility to break inhibiting hierarchies of learning and overcoming learning obstacles and constraints.

Limitations
We used a single case situated in a healthcare context, which arguably contains a positivist and evidence-focussed paradigm, governing performance measurement and learning. However, an underlying rationale in the sample choice was its richness and variety in performance measurement methods and practices, which make the case fill illustrative purposes. We have tried to follow the notions of Walker et al. (2015) in a modest effort to make a conceptual advance by incorporating and integrating behavioural science theory in in an operations management perspective. As such, our sampling strategy could be described as “critical case” and “typical case” (Miles and Huberman, 1994) in an effort to support analytical generalisation (Firestone, 1993). The present study can be seen as an evolved theoretical perspective on learning processes in relation to performance measurement and addressing a perceived gap in traditional production and operations management research. However, in order to conceptually validate and pave the way for further analytical generalisation, additional cases will have to be studied (Miles and Huberman, 1994; Firestone, 1993).

Managerial implications
We argue that all managers must consider what kind of learning environment and what kind of learning outcomes best serve the interests of their organisation. Purposeful and carefully designed organisational arrangements and learning environments are more likely to induce intended learning outcomes. This has a great impact on the design and implementation of necessary performance measurement. Managers who empower employees and entrust them with the capabilities and means that enable them to manage the tasks, methods and results of performance measurement are more likely to drive successful organisational improvement efforts. Balancing adaptive and developmental learning (Ellström, 2010) with exploitation and exploration (March, 1991) is necessary for modern organisations, which must include an analysis of the required facets of learning. Therefore, managers on every level need to consider what metrics are integrated into management meetings, professional meetings, workplace meetings, in development programmes and even in daily work. These arenas can be seen as learning resources that can potentially increase opportunities to critically question the prevailing conditions and to find new solutions to problems that arise. The arenas create opportunities to integrate measurements in the workplace activities and, ultimately, facilitate performance measurement as a driver for increasing knowledge in organisations.

References


Argyris, C. and Schön, D.A. (1978), Organizational Learning, Addison Wesley, Reading, MA.


**Further reading**


**About the authors**

Jason Martin is PhD Student at the Department of Management and Engineering (IEI), Linköping University in Sweden. His research topic is about learning processes within operations management practices. Jason Martin is the corresponding author and can be contacted at: jason.martin@liu.se

Mattias Elg is Professor at the Department of Management and Engineering (IEI), Linköping University in Sweden. He is Director of the HELIX VINN Excellence Centre and conducts research in the areas of healthcare service development, performance measurement and quality improvement.

Andreas Wallo is Senior Lecturer in Education at the Department for Behavioural Sciences and Learning (IBL) at Linköping University, Sweden. He is also affiliated to the HELIX VINN Excellence Centre in the field of working life research. His research interests include studies of leadership, managerial work, workplace learning, organisational development, human resource management, human resource development and interactive research.

Henrik Kock is Associated Professor in Education in the Department for Behavioural Sciences and Learning (IBL) at Linköping University, Sweden. He is also affiliated to HELIX VINN Centre of Excellence in the field of working life research, where he is Director of HELIX Graduate School. His research involves issues on learning and competence development, workplace learning, leadership and organisational renewal.

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: permissions@emeraldinsight.com
Income diversification and bank performance: evidence from BRICS nations

Shweta Sharma and Anand Anand
Indian Institute of Management Ranchi, Ranchi, India

Abstract

Purpose – The purpose of this paper is to examine the impact of income diversification on bank performance in BRICS countries as a structural response to concentration risk. The authors argue that effectiveness of this approach is conditional upon its extent and quality. To understand the role of firm-specific characteristics on effectiveness of diversification, the authors examine this relationship across asset sizes.

Design/methodology/approach – An unbalanced panel data set of 169 BRICS banks is sampled over the period 2001–2015. Fixed effect models and system generalized method of moments techniques are used to test the relationship between diversification and bank performance using alternate measures.

Findings – Results indicate a positive relationship between diversification and performance measured in terms of bank risk and returns for medium and large size banks. However, for small banks this relationship is negative suggesting a “diversification discount.”

Originality/value – The study indicates that diversification as a risk mitigating tool can be effective but the managers and regulators should not emphasize on the “one-size-fits-all” approach for all banks. Policy frameworks for controlling concentration risk should be developed keeping in mind factors like bank size, customer base and financial leverage which brings variations to the risk profile of banks.

Keywords Concentration risk, Risk management, Bank performance, BRICS, Income diversification

Paper type Research paper

1. Introduction

Most of the financial institutions pursue some form of diversification to reduce risk and improve performance. Elsas et al. (2010) documented that for world’s largest banks, level of average diversification had been increased around 33 percent from 1996 to 2003 and have only been marginally declined since then. Financial liberalization and reforms have further prompted banks to explore alternate income channels to maintain their earnings and capture market share in the emerging business areas (Meslier et al., 2014). As a result, diversification of sources of income to fee, trading, underwriting and cross-selling-based businesses has become the most prevalent form of diversification in banks reducing their dependence on the traditional interest-based income (Stiroh and Rumble, 2006). The underlying understanding in favor of the overwhelming emphasis on income diversification is that it results significantly higher reduction in concentration risk due to imperfectly correlated income streams when compared to other forms of diversification, namely, across economic sectors or geographical boundaries (Acharya et al., 2006). Moreover, in case of emerging economies the policy changes pertaining to liberalization and opening up of the economy persuaded banks to diversify their income resources as a part of their strategic decisions to capture the untapped opportunity (Beck et al., 2006).

Diversification has been primarily studied in corporate finance and strategy literature. These studies showcase two complimentary views on the effectiveness of diversification as a risk management technique. The first view originates from the finance literature which is based on the modern portfolio theory[1]. The second view originates from strategy literature which revolves around the concepts of relatedness vs un-relatedness[2], core competencies[3], functional synergies[4] and corporate coherences[5].

Proponents of diversification as a risk management technique argue that revenue coming from different businesses of a diversified firm, which have imperfectly correlated
income streams, reduces the bankruptcy or the overall risk when bank engage in broader scope of activities (Lewellen, 1971; Berger and Ofek, 1995). Expansion in broader activities through diversification helps in leveraging managerial skills and creation of economies of scope. Operational synergy is created through these functional economies, which help in controlling operational risk (Amihud and Lev, 1981). Further, diversification helps in creating a financial supermarket, by catering to all financial needs (loans, investments, trading and underwriting) of a customer, which helps in retaining and improving the customer base, which is very crucial in the present competitive scenario (Laevan and Levine, 2007).

Schmid and Walter (2009) studied income diversification in US commercial banks and concluded that as diversified banks have propriety customer information acquired through underwriting, insurance or trading of securities, this can be exploited for sanctioning of new advances and credit risk management. Information synergies are generated which help in improving bank performance and market valuations. For examining banks’ profitability, Elsas et al. (2010) took a global sample of nine different countries and concluded that revenue diversification enhances bank profitability via higher margins from non-interest businesses and lower cost income ratios. Greater bank profitability renders greater market valuations; eluding to the fact that diversification has an indirect positive impact on bank valuation. Similarly, Sawada (2013) claimed that the benefit associated with revenue diversification outweighs the cost incurred and points toward the existence of “diversification premium” for Japanese banks. To understand the impact of ownership on income diversification, Pemathur et al. (2012) examined a sample of Indian banks. Their findings suggest that public sector banks are benefited more by reduction in insolvency risk and increase in profitability. In a similar line of argument, Trivedi (2015) also reported a positive impact of increasing share of non-interest income and fee-based income on the total operating income and banks’ profitability, for Indian Banks.

On the other hand, opponents of diversification as risk management technique argue that the cost attached with diversification offsets its benefit. Diversification results in inefficient allocation of capital across different divisions of diversified firm thus resulting in cross-subsidization of business (Denis et al., 1997), wherein increase information asymmetries and thus results in higher agency costs (Jensen, 1986; Rajan et al., 2000). Multiple supervisory processes may invoke the regulatory costs, and therefore the financial synergies that are created due to diversification are surpassed by the cost associated with it due to increase in operational risk (Jensen and Meckling, 1976; Palepu, 1985). Several empirical studies suggest that diversification discount exists, i.e. costs may outweigh benefits when banks opt to diversify their product mix (Mercieca et al., 2007; Li and Zhang, 2013). Thus, DeYoung and Roland (2001) documented that diversification across sources of income in banks may result in an enormous increase in profits, but as income from these new streams (fee based or trading) is highly volatile, it may bring in new business risk as per their study based on US markets.

DeYoung and Rice (2004) studied the effect of non-interest income on the financial performance of 4,712 US commercial banks (small vs big) over the period of 1989–2001. These banks adopted different business strategies based on market conditions and technological change, and have documented slow expansion in non-interest activities for well managed banks; the benefits of diversification were offset by the risk-return tradeoff. Similarly, Stiroh and Rumble (2006) studied US markets and concluded that due to the highly volatile alternate sources of income, the costs (risk) associated with diversification outweigh its benefits (increase in the return). Stiroh (2004) explored the link between the growing reliance on non-interest income and the volatility of bank profits and revenue. The results of the study from both aggregate and bank data do not provide any evidence that this shift offers large diversification benefits in the form of more stable
profits or revenue. Similar studies conducted on Chinese and Australian Markets by Berger et al. (2010) and Williams and Prather (2010), respectively, also focus on the risk-return tradeoff related to diversification. However, no significant impact on risk-adjusted performance and stability was found. Chen et al. (2014) reported a growing reliance on non-interest income in Chinese banks over the period of 1986–2008. They reported a negative correlation among interest and non-interest income at the individual bank level and thereby justified that diversification creates benefits. However, their overall findings suggest that growing reliance on non-traditional sources of income results in worsening the risk-return tradeoff. High cyclicality and volatility of non-interest income results in the decrease of marginal benefits of income diversification for Chinese banks as reported by Li and Zhang (2013). Alhassan (2015) using a stochastic frontier analysis also reported a non-linear relationship between income diversification and banks’ cost and profit efficiency for Ghanaian banks. He emphasized that asset size is also an important determinant, enabling banks to reap the benefits of income diversification.

Based on the theoretical justification in favor of diversification, regulators are recommending banks to diversify income sources for better risk management. However, due to the mixed results obtained from banks across different parts of the world empirically, it becomes important to reach a consensus on effects of income diversification on bank performance. This paper examines the relationship between income diversification and performance of banks in terms of risk and return across sizes. We employ fixed effect model (FEM) technique on a panel data set similar to previous studies. Further, as a test of robustness we employ system-generalized method of moments (system-GMM) estimator proposed by Arellano and Bover (1995). We sample banks from BRICS countries for the study. This will further help us to explore the impact of stage of economic development on the relationship that we are going to examine here.

Our findings suggest that for banks with medium and large asset size, diversification has a significant impact on improving bank returns and reducing risk. However, for small commercial banks there exists diversification discount. Therefore, this study rejects the “one-size-fits-all” approach of regulators and policy makers who recommend income diversification as a tool to improve bank performance. We conclude that an efficient diversification strategy should focus on the improved performance through business development along with minimizing the risk carrying out of those exposures. Therefore, while making diversification choices banks should emphasize on profitability along with risk mitigation.

The remainder of this paper is organized into following sections. Section 2 discusses the research methodology and the measurement of key variables used in the study. Section 3 analyzes the data and discusses the results of the study followed by robustness tests in Section 4. Finally, Section 5 concludes the study while discussing possible implications.

2. Research methodology

As discussed in the previous section, we use FEM technique following previous studies. As our data set forms an unbalanced panel, the standard OLS models may give inappropriate results due to the potential bias caused by omitted heterogeneity and problem of endogeneity (Stiroh and Rumble, 2006). To address this problem, Gujarati and Porter (2009) advocated FEM, as the basic assumptions of this model are strict exogeneity between the covariates, constant variance and no serial correlation. This model also controls for unobserved heterogeneity pertaining to firm-specific and time-specific features. The relationship under examination is as follows:

$$Y_{jt} = \alpha_0 + \beta_1 \text{DIV}_{jt-1} + \beta_2 X_{jt-1} + \epsilon_{jt},$$

(1)
where \( j \) refers to the bank and \( t \) indicates the time. \( \beta_1 \) is the estimated coefficient of bank diversification. Here, we test the null hypothesis that there is no significant impact of income diversification on bank performance. If \( \beta_1 \) is significant then income diversification effect holds.

The description of variables is as follows: to measure the bank performance \( (Y) \), we calculate returns with ROA and ROE and riskiness with RISK (Berger et al., 2010). ROA is calculated as the ratio of net income to total assets (Stiroh and Rumble, 2006; Acharya et al., 2006). ROE is calculated as the ratio of net income to total equity (Berger et al., 2010; Tabak et al., 2011). RISK is calculated as the ratio of net non-performing assets to net loans and advances similar to Berger et al. (2010).

DIV is used as a proxy for income diversification. Empirical studies on non-financial conglomerates usually measure diversification using SIC codes and segment-wise accounting data. We use Herfindahl–Hirschman index (HHI) as a measure of diversification in line with previous researchers (Stiroh, 2004; Acharya et al., 2006; Bebczuk and Galindo, 2008; Tabak et al., 2011). HHI index is calculated based on proportion of income coming from across \( n \) income sources. The lower limit of DIV is \( 1/n \) which represents perfectly diversified portfolio and the upper limit is equal to \( 1 \) denoting all income coming from the single source.

To capture effects of bank-specific exogenous variables, we use a set of control variables in our regression analysis denoted by vector \( X \). To capture the effect of bank size we use natural logarithm of year-end total assets (LTA) following the previous literature (Acharya et al., 2006). We use capital adequacy ratio (CAR) calculated as total capital to risk weighted assets, as an indicator of the financial health of bank, similar to Acharya et al. (2006). We use equity ratio (EQR) calculated as the ratio of total equity to total assets of bank, to capture the effect of capital structure on banks performance (Chen et al., 2014).

We also examine an explicit interaction term for the role of bank size on the relationship between income diversification and bank performance. We use additional dummy variables for small, medium and large banks based on their asset size. Banks with an average asset size of less than $1m were categorized as small between $1m and $20m as medium and more than $20m as large banks. Thus, our extended model is as follows:

\[
Y_{jt} = \beta_0 + \beta_1 \text{DIV}_{jt-1} + \beta_2 (\text{DIV}_{jt-1} \times \text{SizeDummy}_{jt-1}) + \beta_3 X_{jt-1} + \beta_4 \text{Size Dummy}_{jt-1} + \epsilon_{jt}. \tag{2}
\]

The dummy coefficients should be interpreted as the difference between the individual bank size captured by the respective dummy variable and the omitted group of another bank size.

Some of our model specifications are based on dynamic panel data and at times FEM may give biased estimates for such panels. Therefore, we employ system-GMM estimators proposed by Blundell–Bond outline in Arellano and Bover (1995) as a robustness check and as an alternative estimator to address the perseverance of endogeneity bias in dynamic panels. The system-GMM estimator combines the standard set of equations in first-differences with lagged levels as instruments and an additional set of equation in levels with lagged first-differences as instruments (Tabak et al., 2011).

The study further tests the relationship using system-GMM estimators as robustness check:

\[
Y_{jt} = \beta_0 + \beta_1 Y_{jt-1} + \beta_2 \text{DIV}_{jt-1} + \beta_3 X_{jt-1} + \epsilon_{jt}, \tag{3}
\]

\[
Y_{jt} = \beta_0 + \beta_1 Y_{jt-1} + \beta_2 \text{DIV}_{jt-1} + \beta_3 (\text{DIV}_{jt-1} \times \text{SizeDummy}_{jt-1}) + \beta_4 X_{jt-1} + \beta_0 \text{Size Dummy}_{jt-1} + \epsilon_{jt}, \tag{4}
\]

where \( Y_j \) is one of three performance measures ROA, ROE and RISK of bank \( j \) at time \( t \). \( Y_{jt-1} \) are the same measures observed in the previous period \( (t–1) \). A positive and significant \( \beta_1 \) indicates that performance levels will be persistence from one year to the next. Further Hansen’s over identifying test is used to validate these instruments, i.e. to test for serial correlation with the estimators. A rejection of null hypothesis recommends the model be well specified.
We also report the Arellano–Bond serial correlation test in the residuals AR (1) and AR (2). The above methodologies have been applied to the data, description of which has been given in the next section.

3. Data and descriptive statistics
The period of our study ranges from 2007 to 2015. The sample comprises financial data of banks from BRICS nations. The data form unbalanced panel of annual financial figures which are sourced from Bloomberg and World Bank database. Banks which have at least three consecutive years of time series observations and are consistently in operation during the study period are included in the sample. To control for the measurement errors, we have omitted extreme values in the bank year observations (3 percent highest and lowest values) for each computed variable. By comparing mean values of the primary sample (356 banks) and filtered sample (169 banks), it is verified that their statistical properties are similar. The final sample consists of 169 banks, dominated by Russian banks (70), followed by Indian banks (43), Brazilian and Chinese banks (21 each) and finally South African banks (14). Total bank year covered in the study is 1,415 (12 percent of data sets from Brazil; 13 percent of data sets from China; 25 percent of data sets from India; 40 percent of data sets from Russia and 8 percent of data sets from South African banks).

The shift from interest income to non-interest income in BRICS nations is evident as shown in Figure 1. The share of non-interest income in BRICS nations was quite negligible in 2001, which has been increased to 30 percent in 2015. This growth in alternate sources of income for BRICS banks is in line with developed countries where banks have a relatively high share of non-interest income.

Table I provides summary statistics for all banks. The mean total assets of all banks in our sample are $169bn and mean total operating income is approximately $11bn. The RISK measure, i.e. the ratio of non-performing loans to total assets, is 0.320. The mean diversification measure is 0.383, denoting that BRICS banks are moderately diversified.

Table II represents descriptive statistics based on the bank’s asset size under three different categories: small, medium and large. The small banks are comparatively less diversified as compared to medium and large banks evident from the mean diversification measures. The mean ROA of large banks is 1.634, which is comparatively higher to smaller banks pegged at 0.733. Similarly, the smaller banks are riskier as compared to

![Figure 1](source: The figure is compiled by authors using data from World Bank reports and Bloomberg data source)

Income diversification and bank performance

---

**Figure 1.** Interest income vs non-interest income of BRICS banks over the period 2001–2015
medium and large size banks in the sample. The equity to total asset ratio (EQR) is much higher for small banks as compared to medium and large size banks. This could be because of the fact that in these emerging economies the larger banks are usually under the government control as compared to smaller banks which are private players.

<table>
<thead>
<tr>
<th>Variable</th>
<th>DIV</th>
<th>ROA</th>
<th>ROE</th>
<th>RISK</th>
<th>LTA</th>
<th>EQR</th>
<th>CAR</th>
<th>TR</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of observations</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
</tr>
<tr>
<td>No. of banks</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
</tr>
<tr>
<td>Mean</td>
<td>0.383</td>
<td>1.682</td>
<td>19.840</td>
<td>0.320</td>
<td>10.164</td>
<td>−0.032</td>
<td>15.870</td>
<td>10.966</td>
<td>169.735</td>
</tr>
<tr>
<td>Median</td>
<td>0.380</td>
<td>1.905</td>
<td>23.540</td>
<td>0.280</td>
<td>10.240</td>
<td>−0.550</td>
<td>13.920</td>
<td>1.781</td>
<td>15.992</td>
</tr>
<tr>
<td>Mode</td>
<td>0.400</td>
<td>1.80</td>
<td>3.88</td>
<td>−0.701</td>
<td>10.540</td>
<td>−4.300</td>
<td>14.370</td>
<td>0.181</td>
<td>0.013</td>
</tr>
<tr>
<td>SD</td>
<td>0.096</td>
<td>1.918</td>
<td>16.011</td>
<td>1.026</td>
<td>4.294</td>
<td>7.397</td>
<td>2323.692</td>
<td>4510.753</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>0.070</td>
<td>−0.034</td>
<td>−0.650</td>
<td>−0.060</td>
<td>0.563</td>
<td>3.252</td>
<td>3.418</td>
<td>4.243</td>
<td></td>
</tr>
<tr>
<td>SE of skewness</td>
<td>0.101</td>
<td>0.101</td>
<td>0.101</td>
<td>0.101</td>
<td>0.101</td>
<td>0.101</td>
<td>0.101</td>
<td>0.101</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.477</td>
<td>0.535</td>
<td>0.282</td>
<td>−0.040</td>
<td>10.540</td>
<td>−4.300</td>
<td>14.370</td>
<td>0.181</td>
<td>0.013</td>
</tr>
<tr>
<td>SE of kurtosis</td>
<td>0.201</td>
<td>0.201</td>
<td>0.201</td>
<td>0.201</td>
<td>0.201</td>
<td>0.201</td>
<td>0.201</td>
<td>0.201</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>0.090</td>
<td>−4.020</td>
<td>−31.820</td>
<td>−4.330</td>
<td>6.620</td>
<td>−8.970</td>
<td>2.300</td>
<td>0.013</td>
<td>0.013</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.490</td>
<td>8.670</td>
<td>62.290</td>
<td>4.210</td>
<td>13.270</td>
<td>12.850</td>
<td>18.600</td>
<td>162.970</td>
<td>3,320.223</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>DIV</th>
<th>ROA</th>
<th>ROE</th>
<th>RISK</th>
<th>LTA</th>
<th>EQR</th>
<th>CAR</th>
<th>TR</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.320</td>
<td>0.280</td>
<td>10.020</td>
<td>−0.530</td>
<td>9.550</td>
<td>−3.600</td>
<td>12.020</td>
<td>0.644</td>
<td>2.589</td>
</tr>
<tr>
<td>50</td>
<td>0.380</td>
<td>1.905</td>
<td>23.540</td>
<td>0.280</td>
<td>10.240</td>
<td>−0.550</td>
<td>13.920</td>
<td>1.781</td>
<td>15.992</td>
</tr>
<tr>
<td>75</td>
<td>0.430</td>
<td>2.863</td>
<td>30.600</td>
<td>1.240</td>
<td>10.850</td>
<td>2.760</td>
<td>17.370</td>
<td>8.070</td>
<td>19.368</td>
</tr>
</tbody>
</table>

**Notes:** DIV, diversification measure using Hirschman–Herfindahl index; ROA, return on assets; ROE, return on equity; RISK, ratio of net non-performing assets to net loans and advances; LTA, log of total assets; EQR, total equity to total assets; CAR, capital adequacy ratio; TR, total operating revenue; TA, total assets. TR and TA are denoted in billion US dollars.

<table>
<thead>
<tr>
<th>Variable</th>
<th>DIV</th>
<th>ROA</th>
<th>ROE</th>
<th>RISK</th>
<th>LTA</th>
<th>EQR</th>
<th>CAR</th>
<th>TR</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of observations</td>
<td>182</td>
<td>182</td>
<td>182</td>
<td>182</td>
<td>182</td>
<td>182</td>
<td>182</td>
<td>182</td>
<td>182</td>
</tr>
<tr>
<td>Mean</td>
<td>0.125</td>
<td>0.733</td>
<td>4.687</td>
<td>0.452</td>
<td>89.744</td>
<td>19.200</td>
<td>4.139</td>
<td>0.116</td>
<td>11.161</td>
</tr>
<tr>
<td>Median</td>
<td>0.050</td>
<td>0.460</td>
<td>0.410</td>
<td>0.600</td>
<td>111.936</td>
<td>24.453</td>
<td>1.000</td>
<td>0.120</td>
<td>11.161</td>
</tr>
<tr>
<td>SD</td>
<td>0.155</td>
<td>1.835</td>
<td>12.496</td>
<td>0.972</td>
<td>5.633</td>
<td>6.476</td>
<td>6.226</td>
<td>123.239</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>0.030</td>
<td>−4.020</td>
<td>−23.820</td>
<td>−4.330</td>
<td>6.620</td>
<td>−7.040</td>
<td>2.300</td>
<td>0.013</td>
<td>0.013</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.650</td>
<td>7.760</td>
<td>56.780</td>
<td>4.210</td>
<td>176.063</td>
<td>28.831</td>
<td>9.000</td>
<td>0.974</td>
<td>123.239</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>DIV</th>
<th>ROA</th>
<th>ROE</th>
<th>RISK</th>
<th>LTA</th>
<th>EQR</th>
<th>CAR</th>
<th>TR</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.320</td>
<td>0.280</td>
<td>10.020</td>
<td>−0.530</td>
<td>9.550</td>
<td>−3.600</td>
<td>12.020</td>
<td>0.644</td>
<td>2.589</td>
</tr>
<tr>
<td>50</td>
<td>0.380</td>
<td>1.905</td>
<td>23.540</td>
<td>0.280</td>
<td>10.240</td>
<td>−0.550</td>
<td>13.920</td>
<td>1.781</td>
<td>15.992</td>
</tr>
<tr>
<td>75</td>
<td>0.430</td>
<td>2.863</td>
<td>30.600</td>
<td>1.240</td>
<td>10.850</td>
<td>2.760</td>
<td>17.370</td>
<td>8.070</td>
<td>19.368</td>
</tr>
</tbody>
</table>

**Notes:** DIV, diversification measure using Hirschman–Herfindahl index; ROA, return on assets; ROE, return on equity; RISK, ratio of net non-performing assets to net loans and advances; LTA, log of total assets; EQR, total equity to total assets; CAR, capital adequacy ratio; TR, total operating revenue; TA, total assets. TR and TA are denoted in billion US dollars.

<table>
<thead>
<tr>
<th>Variable</th>
<th>DIV</th>
<th>ROA</th>
<th>ROE</th>
<th>RISK</th>
<th>LTA</th>
<th>EQR</th>
<th>CAR</th>
<th>TR</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of observations</td>
<td>673</td>
<td>673</td>
<td>673</td>
<td>673</td>
<td>673</td>
<td>673</td>
<td>673</td>
<td>673</td>
<td>673</td>
</tr>
<tr>
<td>Mean</td>
<td>0.394</td>
<td>1.634</td>
<td>22.742</td>
<td>0.346</td>
<td>10.974</td>
<td>1.883</td>
<td>15.131</td>
<td>349.563</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>0.390</td>
<td>1.900</td>
<td>26.160</td>
<td>0.090</td>
<td>10.880</td>
<td>2.880</td>
<td>13.140</td>
<td>90.747</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>0.091</td>
<td>1.417</td>
<td>12.738</td>
<td>1.115</td>
<td>0.950</td>
<td>3.586</td>
<td>6.583</td>
<td>6,028.846</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>0.160</td>
<td>−2.740</td>
<td>−25.960</td>
<td>−2.28</td>
<td>8.960</td>
<td>−4.250</td>
<td>3.710</td>
<td>27.702</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>0.650</td>
<td>6.190</td>
<td>51.940</td>
<td>3.170</td>
<td>13.270</td>
<td>10.530</td>
<td>15.200</td>
<td>5,689.223</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** DIV, diversification measure using Hirschman–Herfindahl index; ROA, return on assets; ROE, return on equity; RISK, ratio of net non-performing assets to net loans and advances; LTA, log of total assets; EQR, total equity to total asset; CAR, capital adequacy ratio; TR, total revenue; TA, total assets. TR and TA are denoted in billion US dollars.
We also calculate the annual average value of diversification measures over the period to understand its tendency and behavioral changes as depicted in Figure 2. The diversification trends start declining in the aftermath of 2008 financial crisis, as banks start concentrating their sources of income. However, as banks of BRICS nations start recovering from the financial crisis, an upward trend in diversification measures can be seen in Figure 2.

Table III provides the correlation matrix of the explanatory variables. The two alternative measures of bank return ROA and ROE have a positive correlation of 0.713 which is statistically significant. Substantiating the risk-return tradeoff, measures of bank return and risk are negatively correlated. As some independent variables were correlated at a 5 percent level of significance, we calculated the variance inflation factor (VIF) statistics for regression model which is used to quantify the severity of multicollinearity in the model. The tests do not suggest that any variables should be dropped from the regression as the VIF statistics are within the specified range.

4. Result and discussion

The empirical results of our fixed effect regression models are presented in Table IV, which investigates the relationship between income diversification and various dependent

<table>
<thead>
<tr>
<th></th>
<th>ROA</th>
<th>ROE</th>
<th>RISK</th>
<th>DIV</th>
<th>LTA</th>
<th>EQR</th>
<th>CAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROA</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROE</td>
<td>0.713**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RISK</td>
<td>-0.100*</td>
<td>-0.333**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIV</td>
<td>0.112**</td>
<td>0.008</td>
<td>-0.197**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTA</td>
<td>0.002</td>
<td>0.244**</td>
<td>-0.029</td>
<td>-0.090*</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQR</td>
<td>0.286**</td>
<td>-0.076</td>
<td>0.214**</td>
<td>-0.093*</td>
<td>-0.459**</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>CAR</td>
<td>0.087**</td>
<td>0.086</td>
<td>-0.056</td>
<td>0.177**</td>
<td>0.103</td>
<td>0.048</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Notes: ROA, return on assets; ROE, return on equity; RISK, ratio of net non-performing assets to net loans and advances; DIV, diversification measure using Hirschman-Herfindahl index; LTA, log of total assets; EQR, total equity to total asset; CAR, capital adequacy ratio. *,**Statistically significant at the 90 and 95 percent levels, respectively.
Table IV. Diversification and bank performance (fixed effect model)

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIV&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>2.314***(0.862)</td>
<td>2.748***(1.519)</td>
<td>1.826**(0.924)</td>
<td>2.312***(1.247)</td>
<td>2.937***(1.423)</td>
<td>1610**(0.998)</td>
<td>-0.041*(0.566)</td>
<td>-1.405***(0.695)</td>
<td>-1.189***(0.915)</td>
</tr>
<tr>
<td>Small×DIV&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-0.122***(0.390)</td>
<td>-0.160*(0.910)</td>
<td>-2.878***(1.720)</td>
<td>-2.360*(1.824)</td>
<td>2.392**(1.283)</td>
<td>2.728*(1.441)</td>
<td>-1.384****(0.223)</td>
<td>-0.619* (1.139)</td>
<td></td>
</tr>
<tr>
<td>Medium×DIV&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>0.362***(0.090)</td>
<td>0.221***(0.090)</td>
<td>0.202***(0.090)</td>
<td>0.428 (0.865)</td>
<td>0.466 (0.863)</td>
<td>0.480 (0.863)</td>
<td>0.463****(0.070)</td>
<td>0.143****(0.065)</td>
<td>0.325***(0.088)</td>
</tr>
<tr>
<td>LTA&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>0.027***(0.017)</td>
<td>0.022* (0.017)</td>
<td>0.019***(0.017)</td>
<td>0.002***(0.171)</td>
<td>0.006***(0.184)</td>
<td>0.079* (0.164)</td>
<td>0.060****(0.141)</td>
<td>0.075****(0.013)</td>
<td>0.065* (0.013)</td>
</tr>
<tr>
<td>EQR&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>0.007***(0.016)</td>
<td>0.005* (0.016)</td>
<td>0.006***(0.016)</td>
<td>0.136* (0.156)</td>
<td>0.115* (0.156)</td>
<td>0.125* (0.156)</td>
<td>0.006***(0.013)</td>
<td>0.004****(0.012)</td>
<td>0.005* (0.013)</td>
</tr>
<tr>
<td>CAR&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>2.321** (1.232)</td>
<td>2.348** (1.532)</td>
<td>2.348** (1.532)</td>
<td>2.348** (1.532)</td>
<td>2.348** (1.532)</td>
<td>2.348** (1.532)</td>
<td>2.348** (1.532)</td>
<td>2.348** (1.532)</td>
<td>2.348** (1.532)</td>
</tr>
<tr>
<td>Small</td>
<td>1.014 (0.902)</td>
<td>1.207 (1.962)</td>
<td>1.207 (1.962)</td>
<td>1.207 (1.962)</td>
<td>1.207 (1.962)</td>
<td>1.207 (1.962)</td>
<td>1.207 (1.962)</td>
<td>1.207 (1.962)</td>
<td>1.207 (1.962)</td>
</tr>
<tr>
<td>Medium</td>
<td>0.262** (0.090)</td>
<td>0.221** (0.090)</td>
<td>0.202** (0.090)</td>
<td>0.428 (0.865)</td>
<td>0.466 (0.863)</td>
<td>0.480 (0.863)</td>
<td>0.463*** (0.070)</td>
<td>0.143*** (0.065)</td>
<td>0.325** (0.088)</td>
</tr>
<tr>
<td>No. of observations</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
</tr>
<tr>
<td>No. of cross-sections</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.634</td>
<td>0.637</td>
<td>0.639</td>
<td>0.554</td>
<td>0.556</td>
<td>0.539</td>
<td>0.549</td>
<td>0.569</td>
<td>0.570</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.624</td>
<td>0.627</td>
<td>0.629</td>
<td>0.511</td>
<td>0.513</td>
<td>0.515</td>
<td>0.505</td>
<td>0.532</td>
<td>0.531</td>
</tr>
<tr>
<td>Durbin–Watson stat</td>
<td>1.842</td>
<td>1.892</td>
<td>1.991</td>
<td>1.950</td>
<td>1.980</td>
<td>2.101</td>
<td>1.922</td>
<td>1.976</td>
<td>1.965</td>
</tr>
<tr>
<td>Akaike information criterion</td>
<td>3.400</td>
<td>3.392</td>
<td>3.390</td>
<td>7.908</td>
<td>7.905</td>
<td>7.903</td>
<td>2.876</td>
<td>2.826</td>
<td>2.821</td>
</tr>
<tr>
<td>Hausman test ($\chi^2$ statistic)</td>
<td>38.315***</td>
<td>40.190***</td>
<td>66.278***</td>
<td>29.171***</td>
<td>27.566***</td>
<td>35.964***</td>
<td>30.942***</td>
<td>34.377***</td>
<td>34.214***</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are reported in parentheses. DIV, diversification measure using Hirschman–Herfindahl index; ROA, return on assets; ROE, return on equity; RISK, ratio of net non-performing assets to net loans and advances; LTA, log of total assets; EQR, total equity to total asset; CAR, capital adequacy ratio. We do not present the coefficients of time dummies for the sake of brevity. ***,***Statistically significant at the 90, 95 and 99 percent levels, respectively.
variables such as ROA, ROE and RISK in alternate models. The study reports mixed results for bank returns and risk across different bank categories. Column (1) denotes the regression results for Equation (1) with ROA as the dependent variable. The estimated coefficients of diversification measure are positive and significant indicating income diversification to be beneficial for improving ROA. To examine the role of size, intersection terms for size dummies (small and medium) are used in Columns (2) and (3). In Column (2), the intersection term corresponding to small banks is negative and significant indicating an adverse effect of diversification on small banks ROA. However, a significant and positive effect of diversification is reported for medium banks. In Column (3), the dummy coefficient of small banks indicates that these banks are significantly different from the other banks (medium and large). Similar results are found for medium banks.

Columns (4)–(6) of Table IV represent regression results with ROE as the dependent variable. The results are similar to ROA. Income diversification plays a positive and significant impact on ROE as denoted by the coefficients of DIV. As for the other control variables, the coefficient of equity ratio and CAR shows a direct relationship with alternate performance measures and indicates sound financial health of banks. These results are in line with Pennathur et al. (2012) who studied Indian banks and reported a positive impact of diversification on bank performance. On the contrary, these results are in contrast to results from developed countries (Stiroh and Rumble, 2006) where benefits of diversification are offset by the highly volatile nature of non-interest income.

One possible explanation for the contrasting findings for developing economies, similar to Lee et al. (2014), can be if fee-based activities are considered as less capital intensive with a low startup cost. Therefore, it acts as a cost effective way for improving banks’ income. All the regressions reject the null hypothesis of Hausman test and thus make the FEM well specified. To check for the problem of autocorrelation, we calculated the Durbin–Watson statistics as reported in Table IV. The value of statistics is close to 2 indicating no problem of autocorrelation in our data set. We also report the Akaike information criterion (AIC) which is used for model selection. Smaller values of AIC are preferred indicating the intersection-dummy model to be the best fit.

Bank risk, measured using “RISK,” has a negative and significant relationship with DIV as reported in Columns (7) to (9) of Table IV. This indicates that higher diversification leads to lower risk. Therefore, BRICS banks can mitigate their risk by diversifying across non-interest income sources. These results are in line with the traditional intermediation theory and corporate finance literature, which argue that due to the uncorrelated income streams across different sources of income, diversification does help to control for concentration risk by creating a premium. However, in the case of small banks, higher diversification leads to higher risk, indicating adverse effects of diversification on the performance of small banks. The possible reason of our findings could be that small banks usually work on the principal of relationship banking and have expertise in the traditional line of business that concentrates on interest income (Rajan et al., 2000). For small banks, our findings are aligned with the studies on financial conglomerates which indicates the existence of “diversification discount” based on corporate finance literature (Jensen and Meckling, 1976; Laeven and Levine, 2007; Rajan et al., 2000). The agency conflict among the managers and shareholder and inefficient allocation of resources between different segments within a firm can explain the finding that diversification has negative consequences for corporate performance. Our results are aligned with the findings of Mercieca et al. (2007) who reported that small European banks are not able to reap the benefits of diversification due to lack of expertise and experience in the area.

For banks with medium and large asset size, diversification has a positive impact on bank performance, i.e. higher diversification increases bank returns and reduces risk. These results indicate that as the size of bank increases, efficient diversification strategies
are being adopted by the banks to control for concentration risk. Our findings are aligned with the results of Amidu and Kuipo (2015) who reported that diversification helps in leveraging managerial skills and abilities across products and geographic regions in financial conglomerates. Skillful implementation of diversification policy using well-defined roles of managers and regular follow up through supervisory review process can lead banks to mitigate risk and thus enhance bank performance. These findings are in line with the findings of Stiroh (2004) and Elsas et al. (2010) who emphasized that larger banks can significantly improve their profitability by relying on non-traditional income sources and thus reducing the dependency on the traditional banking activities. The results for larger banks especially in the context of emerging economies could be due to high correlations between the development of financial intermediary market and stock market indicators as emphasized by Demirguc-Kunt and Levine (1996). They reported that large stock markets having a well-diversified portfolio are comparatively more internationally integrated as compared to smaller markets having concentrated portfolios. Similarly, well-diversified larger banks follow strong information disclosures and internationally accepted standards which complement them to reap the benefits of diversification.

Robustness tests results
As some of our model specifications are based on dynamic panel data and considering that at times FEM may give biased estimates for such panels, we employed system-GMM technique as a robustness check. The empirical results for GMM regression models are presented in Table V using ROA, ROE and RISK as the dependent variable in alternate models.

In the GMM specification, the study reports Hansen $J$-statistic and corresponding $p$-values, which test the null hypothesis on whether the instruments are uncorrelated with the error term. In all the cases, we reject the null hypothesis and thus conclude that the GMM approach is well specified. For all the models, the value of AR (2) is insignificant, indicating we cannot reject the null hypothesis of the second-order correlation, a finding which is likely in the first difference where the basic assumption is no serial correlation among the original disturbance terms. The results support our earlier findings of all diversification coefficients being significant for dependent variables ROA and ROE. RISK also shows an inverse relationship with diversification measure. For small banks, GMM estimators also report the adverse impact of diversification on their performance similar to FEMs.

Although the debate on income diversification as advantageous or disadvantageous in banks continues, banks across the globe are adopting diversification and the average diversification level in banks has increased threefold from 1996 to 2003 and has not declined significantly since then (Elsas et al., 2010). The given study incorporates several tests and robustness checks to verify the relationship between income diversification and bank performance. The findings suggest that in general, income diversification positively affects the bank profitability by improving risk-adjusted returns, which is in line with the earlier studies of Stiroh and Rumble, 2006 (US markets), Elsas et al., 2010 (global) and Williams and Prather, 2010 (Australia). In the case of large and medium size banks, our results are aligned with traditional intermediation theory, which suggests that sources of income diversification are likely to result in improved bank performance. In particular, our analysis suggests a positive link between diversification and bank returns and an inverse relationship between diversification and bank risk. For small commercial banks, we find that diversified revenue streams have an adverse impact on bank returns and risk. A limited expertise of small banks managers could be one possible explanation of our results. Thus, diversification does not contribute to improved safety and soundness of small banks. These results are in line with the results of Mercieca et al. (2007) who investigated how the performance of small European banks was affected due to the
<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROA_{t-1}</td>
<td>0.357*** (0.005)</td>
<td>0.337*** (0.006)</td>
<td>0.339*** (0.010)</td>
<td>0.342*** (0.002)</td>
<td>0.354*** (0.003)</td>
<td>0.272*** (0.028)</td>
<td>0.275*** (0.011)</td>
<td>0.365*** (0.007)</td>
<td>0.355*** (0.044)</td>
</tr>
<tr>
<td>ROE_{t-1}</td>
<td>1.902*** (0.002)</td>
<td>1.528** (0.008)</td>
<td>1.764** (0.006)</td>
<td>1.921*** (0.006)</td>
<td>1.596** (0.069)</td>
<td>1.764** (0.065)</td>
<td>-0.361 (0.630)</td>
<td>-1.747** (0.242)</td>
<td>-1.654** (0.063)</td>
</tr>
<tr>
<td>RISK_{t-1}</td>
<td>-0.615*** (0.004)</td>
<td>-0.525*** (0.036)</td>
<td>-0.562** (0.893)</td>
<td>-0.675*** (0.001)</td>
<td>-0.675*** (0.001)</td>
<td>-0.675*** (0.001)</td>
<td>-0.675*** (0.001)</td>
<td>-0.675*** (0.001)</td>
<td>-0.675*** (0.001)</td>
</tr>
<tr>
<td>DIV_{t-1}</td>
<td>1.363*** (0.003)</td>
<td>1.500*** (0.416)</td>
<td>1.500*** (0.416)</td>
<td>1.500*** (0.416)</td>
<td>1.500*** (0.416)</td>
<td>1.500*** (0.416)</td>
<td>1.500*** (0.416)</td>
<td>1.500*** (0.416)</td>
<td>1.500*** (0.416)</td>
</tr>
<tr>
<td>Small × DIV_{t-1}</td>
<td>-0.311*** (0.041)</td>
<td>-0.227*** (0.020)</td>
<td>-0.227*** (0.020)</td>
<td>-0.227*** (0.020)</td>
<td>-0.227*** (0.020)</td>
<td>-0.227*** (0.020)</td>
<td>-0.227*** (0.020)</td>
<td>-0.227*** (0.020)</td>
<td>-0.227*** (0.020)</td>
</tr>
<tr>
<td>Medium × DIV_{t-1}</td>
<td>0.007*** (0.001)</td>
<td>0.000*** (0.002)</td>
<td>0.000*** (0.002)</td>
<td>0.000*** (0.002)</td>
<td>0.000*** (0.002)</td>
<td>0.000*** (0.002)</td>
<td>0.000*** (0.002)</td>
<td>0.000*** (0.002)</td>
<td>0.000*** (0.002)</td>
</tr>
<tr>
<td>EQR_{t-1}</td>
<td>0.013*** (0.001)</td>
<td>0.011*** (0.003)</td>
<td>0.011*** (0.003)</td>
<td>0.011*** (0.003)</td>
<td>0.011*** (0.003)</td>
<td>0.011*** (0.003)</td>
<td>0.011*** (0.003)</td>
<td>0.011*** (0.003)</td>
<td>0.011*** (0.003)</td>
</tr>
<tr>
<td>CAR_{t-1}</td>
<td>-0.151*** (0.240)</td>
<td>-0.151*** (0.240)</td>
<td>-0.151*** (0.240)</td>
<td>-0.151*** (0.240)</td>
<td>-0.151*** (0.240)</td>
<td>-0.151*** (0.240)</td>
<td>-0.151*** (0.240)</td>
<td>-0.151*** (0.240)</td>
<td>-0.151*** (0.240)</td>
</tr>
<tr>
<td>Year dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>No. of observations</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
<td>1,415</td>
</tr>
<tr>
<td>No. of cross-sections</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
<td>169</td>
</tr>
<tr>
<td>Hansen statistic</td>
<td>51.537</td>
<td>48.199</td>
<td>0.668</td>
<td>47.928</td>
<td>45.072</td>
<td>48.421</td>
<td>41.288</td>
<td>43.433</td>
<td>39.304</td>
</tr>
<tr>
<td>Hansen p-value</td>
<td>0.337</td>
<td>0.456</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>AR (1)</td>
<td>-1.345**</td>
<td>-1.765**</td>
<td>-1.765**</td>
<td>-1.765**</td>
<td>-1.765**</td>
<td>-1.765**</td>
<td>-1.765**</td>
<td>-1.765**</td>
<td>-1.765**</td>
</tr>
<tr>
<td>AR (2)</td>
<td>0.056</td>
<td>0.056</td>
<td>0.056</td>
<td>0.056</td>
<td>0.056</td>
<td>0.056</td>
<td>0.056</td>
<td>0.056</td>
<td>0.056</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are reported in parentheses. ROA, return on assets; ROE, return on equity; RISK, ratio of net non-performing assets to net loans and advances; DIV, diversification measure using Hirschman–Herfindahl index; ROA, return on assets; LTA, log of total assets; EQR, total equity to total asset; CAR, capital adequacy ratio. *,**,***Statistically significant at the 90, 95 and 99 percent levels, respectively.
observed shift into non-interest income activities. They suggested that banks should concentrate on those lines of business where management has the greatest expertise and experience. They highlighted the importance of specialization which is aligned with the recent studies which reports diversification discount in financial conglomerates.

5. Conclusion

Financial liberalization in developed and developing nations over the past three decades has resulted in deregulation of interest rates, liberalization of credit control, privatization of financial institutions and removal of entry barriers in capital markets. These changes have prompted banks to explore alternate channels of income to maintain their income and capture market share in emerging business areas (Beju and Ciupac-Ulici, 2012). The compelling force behind this is the opportunity of business growth in the new business segments. This shift in orientation of banks presents with a dual opportunity where they should be able to enhance their returns by entering into high growth business segments, while also reducing their dependence onto traditional sources of income (mostly interest income). The idea comes from the modern portfolio theory that suggests that efficient diversification should reduce unique risk and therefore is a structural response to the concentration risk.

Our findings suggest that in most of the cases sources of income diversification as a by-product have shown increase in returns along with reduction in risk for banks. However, in few cases this diversification has resulted in decrease in returns with an increase in risk indicating toward inefficient diversification. This seems to be a result of particular focus of banks to enter into the new business segments while ignoring the perspective of efficiency of diversification achieved in the process (Alhassan, 2015). This reminds us that all diversifications are not efficient and therefore, reinforce that while making diversification choices banks should focus on profitability as well as risk mitigation. An efficient diversification strategy should focus on returns improvement yet minimizing the risk while entering into the new business areas. This implies that while entering into new business segments, banks must recognize incremental unique risks in their portfolio and should identify specific dimensions to achieve efficient diversification of these risks. A rampant diversification along all possible dimensions is not necessarily efficient and may turn counterproductive resulting in increased concentration risk and/or reduced returns. Our results suggest that a ceiling-based approach (Bhide et al., 2002) to diversification prescribed by the regulators needs to be reexamined and revised in favor of a unique risk identification and efficient diversification. A lapse to this end will increase high severity–low-frequency risk increasing the probability of a bank failure.

Further, our results contest the notion that a joint adoption of traditional and non-traditional banking activities creates diseconomies of scope. Our finding partially supports the diversification benefits as postulated by the “Shadow Banking” literature suggesting that banking system has evolved over time and the importance of non-interest activities has been identified by the banks to improve profitability. Regulators and managers should realize the scope for reducing risk through diversification along with business development.

Our findings suggest distinct impacts of diversification on different banks. Therefore, policy frameworks for improving bank performance should be developed keeping in mind bank-specific factors like bank size, customer base and financial leverage. The distinct institutional framework, macroeconomic conditions and political environment of emerging economies, the results of diversification strategies adopted by developed nations, might result in dissimilar impacts of this diversification. Some other factors like fragility of banking system, structure of industries and regulatory environment as suggested by Beck et al. (2006) also impact the benefits of diversification and could be explored as an extension of the current study. The share of non-interest income in banks has increased considerably
over the past few years which perhaps offer the right incentives to the managers and the regulators to identify and resolve the inefficient diversifications. As this study examines only few emerging economies over a limited period, the results have to be taken with a note of caution. However, the results supplement a growing body of empirical research in financial intermediation literature that indicates difficulties in reaping benefits of diversification due to the "one-size-fits-all" regulatory approach. Moreover, it beckons the need for future empirical studies from the perspective of bank age and ownership for robustness of results.

Notes

1. Markowitz (1952) laid the foundation for the modern portfolio theory which states that an investor could choose a diversified portfolio of uncorrelated investment opportunities that will minimize his/her risk for a given level of return.

2. The resource-based view suggests that related diversification should lead to superior firm performance compared to unrelated diversification and single business strategy (Rumelt, 1982; Wernerfelt, 1984; Barney, 1991). Proponents of related diversification believed that (i.e. firms with moderate levels of product diversification) a firm can exploit economies of scope that derive from the ability to share resources and capabilities among business units that are mutually reinforcing (Barney, 1991; Lubatkin and Chatterjee, 1994; Palich et al., 2000).

3. The proponents of related diversification advocated that firms should concentrate on their core competencies defined as “a harmonized combination of multiple resources and skills that distinguish a firm in the marketplace” by Schilling.

4. Palich et al. (2000) suggested that firms which are diversifying may obtain potential gains by enhancing functional synergies derived from intra-firm product technology propagations, limited access to factors of production and effective use of learning curve efficiencies.

5. Firms are coherent to the extent that their constituent businesses are related to each other (Teece, 1980).

6. Shadow banking typically describes financial activities occurring outside the regulated banking sector (Gennaioli et al., 2013). The traditional intermediation activities remain the core of banking business, but now they are increasingly being complemented by non-interest income to increase banks’ efficiency.

References


About the authors
Shweta Sharma is Research Scholar at the Indian Institute of Management, Ranchi, India. She has more than seven years of experience in banking with State Bank Group. She received Master’s degree in Financial Management from the Indian Institute of Management, Ranchi. Her areas of interest include risk management, financial institutions appraisal, investment management and market microstructure. Shweta Sharma is the corresponding author and can be contacted at: shweta.sharma13fpm@iimranchi.ac.in

Anand Anand is Assistant Professor in the Finance and Accounts Area at the Indian Institute of Management, Ranchi, India. He has more than seven years of academic experience which includes research, teaching and training. His areas of interest include investment management, market microstructure, financial markets, risk management, corporate finance, project finance, financial institutions appraisal and MIS technologies. He has several refereed publications to his credit which include national and international journal and case study publications.

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Flexible budgeting influence on organizational inertia and flexibility

José Carlos Tiomatsu Oyadomari
Department of Accounting,
Mackenzie Presbiterian University, São Paulo, Brazil and
Insper, São Paulo, Brazil

Paulo Sérgio Lima Pereira Afonso
Department of Production and Systems,
University of Minho – Campus of Azurem, Guimaraes, Portugal

Ronaldo Gomes Dultra-de-Lima
Department of Accounting, Mackenzie Presbiterian University,
São Paulo, Brazil

Octavio Ribeiro Ribeiro Mendonça Neto
Department of Accounting, Mackenzie Presbiterian University,
São Paulo, Brazil and
Department of Accounting, Methodist University of São Paulo,
São Bernardo do Campo, Brazil, and

Maria Carolina Gazso Righetti
Mackenzie Presbiterian University, São Paulo, Brazil

Abstract
Purpose – The purpose of this paper is to investigate how the use of flexible budgets may influence different institutional logics (organizational inertia and flexibility).
Design/methodology/approach – A qualitative research based on a single case study in a multinational subsidiary company was carried out. The data were mainly collected using the dialog technique through open-ended and semi-structured interviews and complemented with direct observation in informal and formal meetings and the analysis of internal documents. Content analysis was used for the analysis of the findings.
Findings – The use of flexible budgets, which isolates the negative variations due to the decrease in sales volume, may contribute to organizational inertia. However, this can be counterbalanced if the managers try to minimize the decline in performance through initiatives that promote organizational flexibility. In this case study, it was found that the alignment between the production director and the controller, who frequently work under different institutional logics, was important to stimulate organizational flexibility particularly in continuous improvement projects.
Research limitations/implications – The findings of this paper are based on only one in-depth case study. Hence, the results cannot be generalized, but a theoretical contribution can be made. Furthermore, the findings are constrained by the constructs used and the specific managerial and theoretical perspectives that have supported the analysis.
Practical implications – These results can be useful particularly for companies that are dealing with the abrupt drop in the sales volume and use the flexible budget as a performance assessment technique. These firms must pay attention because this combination can stimulate organizational inertia. To counteract this problem, it is necessary that controllers and the managers work by understanding the initiatives that promote organizational flexibility, mainly by Kaizen projects, which can minimize performance decline.
Social implications – The main contribution may be how to deal with the different managers’ behaviors, given the decrease in sales volume, and it can help an organization survives in times of economic recession and fierce competition environments.
Originality/value – This paper contributes to both practical and academic dimensions. Indeed, despite being widely used, flexible budgeting is not a widely researched topic.
Keywords Institutional logics, Management accounting, Kaizen, Organizational flexibility, Flexible budgeting, Organizational inertia
Paper type Research paper
1. Introduction
Companies use a series of activities and tools to support planning, control, and decision making in order to enhance organizational performance and competitive advantages over their competitors. In this context, management accounting systems and practices identify, collect, measure, classify and report information for these purposes (Hansen and Moven, 2003). Particularly, since the 1920s that budgeting has been widely used for planning and controlling in organizations, projecting revenues and expenses for the upcoming period based on both past information and forecasts (Garrison et al., 2007; Horngren et al., 1999).

Despite criticisms and difficulties to deal with uncertainties in dynamic environments and the advent of beyond budgeting (Hope and Fraser, 2003), Libby and Lindsay (2010) point out that budgets continue to be widely used in the USA and Canada. Furthermore, and apparently contradictorily, the use of budgets has been found positively associated with hostile and turbulent external environments (Chenhall, 2003). Uncertainty is not a new issue in management control (Otley and Soin, 2014); however, the permanent and higher unpredictability that characterizes more and more industries and markets means that the changes in sales volume budgeted due to strategic uncertainties have been contributing to the obsolescence of budgeting as performance evaluation tool.

In this context, the use of flexible budgeting has been recommended to deal with the strategic uncertainty that commonly affects the planned sales volume (Garrison et al., 2007; Horngren et al., 1999). In practice, flexible budgeting is operationalized by graphical tools known as waterfall charts or bridge analysis, which are considered diagnostic tools according to Simons’ (2000) framework.

On one hand, we may consider that this tool can help managers to deal with strategic uncertainties, making the company more flexible (Shuabi et al., 2005), by planning through more manageable budget dimensions. By contrast, while isolating the variations between budgeted and actual results due to volume variation, the use of the flexible budgeting may generate an excuse that limits the potential performance of the organization as its use may induce behaviors by which the performance variation can be attributed to external causes (Merchant and Van der Stede, 2012). Therefore, flexible budgeting can contribute to organizational inertia (Näslund and Pemer, 2012), especially when the managerial actions did not impact on exogenous independent variables such as the sales volume and when variations are due to events that are not controllable by managers.

The literature explains volume variation considering two factors: the market volume variation and the market share variation (Horngren et al., 2012; Shank and Govindarajan, 2008). The first is associated with the determinants of market volume, which is less controllable by the organization, but the second can be associated with managerial initiatives.

If a significant drop in sales occurs and the organization cannot revert it through its commercial efforts, organizational inertia can be generated. This problem occurs due to the slow response of the organization to changes in the environment (Hannan and Freeman, 1984; van der Steen, 2009). In these situations, managers may justify performance losses to external causes adopting a uniformity of discourse (Näslund and Pemer, 2012), generating inertia regarding behavior, initiatives and insights (Godkin and Allcorn, 2008). Kelly and Amburgey (1991) point out that organizational inertia is positively associated with organizational size. Thus, management control can be associated to inertia particularly in large organizations.

By contrast, the literature recognizes the role of management control instruments as influencers of flexibility. For example, Joiner et al. (2009) found that financial performance measures are essential in the relationship between the strategic focus of flexible manufacturing and organizational (financial) performance. Thus, organizational flexibility is the counterpoint to organizational inertia, being conceptualized as the capacity of an organization to adjust quickly to changes in conditions such as sales volume, product mix,
cost structure, direct labor, equipment use and others (Karuppan, 2004; Koste et al., 2004; Patel et al., 2012). In this case, a flexible budget would encourage managers to seek to minimize the effects of changes through managerial initiatives enabling the organization to deal with environmental uncertainties, such as when a drop in sales volume occurs.

These contradictions and complexity of the influence of flexible budgeting in the organization ask for empirical findings and theoretical conceptualization, which are lacking in the literature. Furthermore, in an organization, different managers may have different perspectives and act differently in these situations. Thus, the different roles and interaction of managers and collaborators must be taken into consideration.

In a company, the different views and actions of managers and collaborators influence and also are influenced by different institutional logics. Indeed, institutional logics shape people behavior, but also individual and organizational actors can change institutional logics (Thornton, 2004). Institutions have a central logic that constrains both the means and the ends of individual behavior. However, while institutions constrain action they also provide sources of change. The contradictions inherent in the differentiated set of institutional logics provide individuals and organizations with resources for transforming individual identities, organizations and society (Friedland and Alford, 1991; Thornton and Ocasio, 2008).

In this context, the groups with sufficient power to prevail their views can define dominant institutional logics, but this is partially determined by the positioning of the managers, in general, and managerial accountants in particular (Lambert and Sponem, 2011). The academic studies in management accounting that used institutional logics have focused on the adoption of new management accounting practices or performance measurement systems (PMS) (Carlsson-Wall et al., 2016; Dambrin et al., 2007; Goretzki et al., 2013; Lambert and Sponem, 2011), and on the budgetary process (Amans et al., 2015; Ezzamel et al., 2012). However, the extant literature and particularly in the management accounting field is still providing incipient studies regarding flexible budgeting – this subject appears only in a few surveys (Ekholm and Wallin, 2011), accounting textbooks and normative papers (Yahya-Zadeh, 2012).

In the context of this research, it is important to highlight that management accounting research points out that controllers are critical elements to change the dominant institutional logic in the organizations (Lambert and Sponem, 2011). Furthermore, research that has used institutional logics has usually shown that controllers and operational managers have acted in opposing the current status quo and with different institutional logics (Dambrin et al., 2007; Goretzki et al., 2013). On the other hand, despite the apparent coexistence and interdependence between organizational inertia and flexibility, no study has inquired how these two different dimensions coexist and how they can be influenced by the adoption of managerial control practices, particularly flexible budgeting.

To address this research gap, the following question was examined in this study considering the theoretical lens of institutional logics: how does the use of a flexible budget may contribute and explain to organizational flexibility and inertia? This study follows a proposition recommended by Franco-Santos et al. (2012) who investigated how the contemporary management accounting practices and performance systems can support and enhance the companies’ flexibility and dynamism, particularly when emerging organizational inertia.

In this research, the empirical data were collected from a Brazilian company acquired about ten years ago by a first-tier USA multinational corporation that develops and produces automotive components for automakers. In the last years, the decreasing in sales volume has affected the level of profitability in the industry. The high degree of strategic uncertainties required significant flexibility in order to achieve and retain competitive advantages, but the use of flexible budgets may have contributed to organizational inertia. Thus, the research findings are particularly relevant and rich for discussing the aforementioned issues.
The data were mainly collected using the dialogue technique through open-ended and semi-structured interviews and complemented with direct observation and the analysis of internal documents. Content analysis was used for the analysis of the findings.

This paper is organized as follows. Section 2 provides a literature review on the constructs that were used and on institutional logics. In Section 3, the case study and the research methodology are explained. Sections 4 and 5 are devoted to present and discuss the main findings of this research. Finally, Section 6 is dedicated to present the final remarks, the main contributions of this research study, its limitations, and the opportunities for further research.

2. Constructs and theoretical framework

There is a need for more dynamic management accounting and control, in general, and, particularly in multinationals that face high business and market uncertainties. Unanticipated external events may invalidate companies' strategies facing strategic uncertainties. In this context, flexible budgeting has been proposed as a way of dealing with such uncertainties. Notwithstanding, the apparent need for greater flexibility, organizational inertia is recognized as prevalent in many of these situations. Therefore, different institutional logics may emerge and evolve in a process in which the discourse of organizations' key elements may simultaneously converge and diverge in more complex and dynamic ways supporting, enhancing, but also constraining management accounting and control in practice.

2.1 Flexible budget

To deal with strategic uncertainty and be able to use the budget as a diagnostic tool, the literature in management accounting and control (Garrison et al., 2007; Horngren et al., 1999) recommends the use of the flexible budgeting technique. This technique consists of comparing the budget vs actual results, mainly in the income statement. The main advantage of this process is to isolate the budget variations due to the changes in market demand, market share, prices, consumption, and raw material prices, among others. In a scenario where the sales volume is decreasing, the most significant variation in results is likely to be attributed to the environment such as external causes, primarily when the company cannot influence the market. In this regard, managers can visualize the leading causes of variations and make the best decisions aiming to guarantee the budgeted bottom line.

Our definition of flexible budget is consistent with the definition of Ekholm and Wallin (2011) for whom it is a general term used for the budget that can be changed in shorter intervals. But the scope of our construct does not include the four types of budget under this general concept, namely, variable, flexible, revised and rolling forecast (Ekholm and Wallin, 2011, p. 147). However, one of the weak points of flexible budgeting is that it may be used to invoke external causes to justify the failure in accomplishing organizational objectives, as explained by the literature (Merchant and Van der Stede, 2012), primarily when large volume variations are observed in the firm. Therefore, it can induce managers in adopting different behaviors contributing to the organizational inertia or flexibility.

Even though the literature has pointed out the relevance of flexible budgeting, few researches have discussed this subject in-depth. An exception is a work of Ekholm and Wallin (2011) in Scandinavian companies who identified that traditional and flexible budgets are not considered mutually exclusive. The two types of budget can be useful for organizations who have a strong commitment to their strategy. The authors also found a negative relationship between environmental uncertainty and the perceived usefulness of traditional annual budgets, but no significant association with flexible budgeting (Ekholm and Wallin, 2011).
In the light of contingency theory, whether a company has a high volatility regarding sales, it should be flexible enough to adjust to the changes in the external environment.

2.2 Organizational flexibility and inertia

Organizational flexibility and inertia are opposite conceptualizations of a much more diffuse reality where organizations are between these two. Inertia can be measured by the cost of change, while organizational flexibility can be measured by efficiency improvement of change. There is no absolute flexibility, otherwise organizational change cannot last and impact on the organization and there is no absolute inertia either. Otherwise, the organization will be totally obsolete. During organizational change, inertia and flexibility are balanced through a continuous and dynamic process of change and stabilization (Hannan and Freeman (1984).

Flexibility is the ability to deal quickly with changes (Karuppan, 2004; Shuiabi et al., 2005), and it is considered especially important in an environment characterized by unanticipated external events that may modify the current strategies (Simons, 2000). Additionally, flexibility has been recognized as an important factor for competitiveness (Koste et al., 2004; Combe et al., 2012; Scherrer-Rathje et al., 2014).

The degree of flexibility depends on a set of external and internal aspects of the organization. For example, Pekkola et al. (2016) argued that in small companies, the lack of bureaucracy contributes to have a higher level of flexibility. These authors developed their construct (i.e. flexibility) as the ability to respond quickly to environmental changes. In operations management literature, flexibility constructs are associated to changes in terms of resources such as equipment, workforce, material handling, product mix, new products, among others. However, in operations, the most common approach is to have this construct (i.e. flexibility) related more to skills than to resources (Pagell and Krause, 2004). But flexibility can also be discussed through the routine dimensions that support management control systems and processes. In this context, to deal with high level environmental uncertainties, organizations must have flexible operating systems and less formal structures (Patel, 2011).

For firms that work in environments without significant changes, rigid responses may be the best course of action; however, for firms that deal with high volatility, flexibility is a core strategic competency (Yu et al., 2015).

Furthermore, the construct inertia is the antonym of flexibility in the organizational literature. Inertia is defined as the relatively slow speed of organizational change in response to environmental changes (van der Steen, 2009). However, in the conventional sense, inertia has different definitions in the literature. For example, according to Boyer and Robert (2006, p. 325), inertia is prevalent despite “frequent calls for change and flexibility by different stakeholders.” For Näslund and Pemer (2012, p. 90), “organizational inertia may also imply an unwillingness to alter residing stories and language, as from a narrative approach organizational change entails conversational shift.”

Among the internal factors that contribute to organizational inertia, Hannan and Freeman (1984) highlight the importance of political coalitions. Moreover, Berry et al. (2009, p. 10) state that we have to understand the role of the coalitions of interest and the power of the change processes.

Godkin and Allcorn (2008) propose that the organization is impacted in a restrictive way through some of the dimensions of inertia, namely, insight, action, and psychological. For them, psychological inertia refers to the resistance of individuals to changes that impact in the worker’s life. Additionally, insight inertia occurs when organizational members are not skilled to understand adequately the environment or fail to explain the cause behind a given phenomenon. And finally, action inertia appears when managerial analysis of the internal and external environment is followed by slow, incomplete or deficient solutions (Godkin and Allcorn, 2008). In short, an organization has a kind of inertia volume, from
different perspectives, which can be an adaptation of the concept of inertia outsourcing as defined by Mol and Kotabe (2011).

Management accounting and control and PMS can promote organizational inertia, and according to Micheli and Manzoni (2010), create “ossification” (i.e. organizational paralysis). This may be a problem for organizations competing in stable markets, being a serious issue for firms operating in very dynamic environments.

2.3 Institutional logics

Institutional logics is based on the assumption that institutions, individuals and organizations are constituted by interconnected elements of dynamic social systems, and they are influenced by several internal and external factors (i.e. from multiple institutional logics) (Thornton et al., 2012). For Thornton and Ocasio (1999, p. 804), institutional logics is “the socially constructed, historical patterns of material practices, assumptions, values, beliefs, and rules by which individuals produce and reproduce their material subsistence, organize time and space, and provide meaning to their social reality.” This understanding is founded on Friedland and Alford’s (1991) concept of institutional logic that explores the interrelationships between individuals, organizations, and society; rejecting individualistic and rational choice theories and macro-structural perspectives.

Typically, organizations are established of different groups that have various institutional logics, and where some use the power they have to make their points of view prevail (Lambert and Sponem, 2011). The institutional logic is influenced and partially determined by these key elements (Lambert and Sponem, 2011). Thus, institutional logics are presented and translated into the actors’ day-to-day actions and “material” elements.

Organizations can accommodate different institutional logics, which may coexist in conflict or harmony (Besharov and Smith, 2014). For example, Carlsson-Wall et al. (2016) showed that in a football club, sport and business logics are sometimes in conflict, but in other teams they may be in harmony.

If a firm needs to change the prevailing institutional logic, it needs to change its organizational discourse as well. For such companies, the institutionalization process is not completed in this situation. Likewise, organizational ideas, discourses, and techniques are not consistent in this process (Dambrin et al., 2007). Furthermore, Dambrin et al. (2007, p. 174) stated that “when a discourse dealing with the new logics cannot be heard by organizational actors, this does not prevent control techniques that are coherent with the new ideal from developing, meaning that organizational taboos can be by passed by techniques.” Therefore, coherent control techniques have a relevant role in overcoming organizational restrictions.

As observed by Damayanthi and Gooneratne (2017), the institutional logic approach has enabled a theoretical development in the areas related to management control. For example, in performance analysis and the budgeting process. Damayanthi and Gooneratne (2017) also note in their literature review that most studies on institutional logics, in the area of management accounting, have limited their theoretical foundations to a single dimension or, at most, to theories from sociology. As a result, the authors suggest that future research should be carried out based on a broader theoretical perspective.

Stål (2015) argues that institutional logic can be used to analyze change. He also claims that knowing what causes inertia would enable more effective efforts to overcome this phenomenon. However, Holm (1995, p. 398) asked “how can actors change institutions if their actions, intentions, and rationality are all conditioned by the very institution they wish to change”. Indeed, more research is needed to understand how institutions influence individuals’ actions and how these actions may contribute to more stable or changeable environments. Hannan and Freeman (1984) argue that more research on the microfoundations of institutional logics, highlighting the interplay between individuals,
organizations, and institutions, is needed. In this context, these authors highlight four mechanisms of change: institutional entrepreneurs, structural overlap, event sequencing, and competing institutional logics. Institutional entrepreneurs are agents with access to the resources that are able to create and modify institutions. Structural overlap occurs when individual roles, organizational structures, and functions, which were previously distinct, are forced into association triggering a change in institutional logics guiding the firm (Thornton, 2004). Event sequencing is an occurrence that changes the interpretation and meaning of cultural symbols and social and economic structures such as the emergence of new resources or sources of power, reinforcing or eroding the dominance of the incumbent logic. Finally, a wide variety of mechanisms are taken into consideration to explain how alternative logics can compete. Also, how institutional logics become institutionalized and deinstitutionalized is an important research topic.

3. Research methodology

3.1 Case study

In this paper, a fictitious name was used to ensure the anonymity of the company. PARTSCORP is a Brazilian subsidiary of MNC, a US company that develops and produces automotive components for automakers. The parent company is present in over ten countries and hires more than 5,000 employees worldwide. PARTSCORP was a family-owned company until the acquisition in the mid of the 1990s.

At the time of this study, the Brazilian economic scenario was characterized by a high degree of strategic uncertainty, which requires significant flexibility from manufacturing companies in order to achieve sustainable competitive advantages. The volatility and the strong influence of the government’s economic policies (interest rates, fixing differentiated sales taxes for the vehicle industry, among others) characterized the automotive components industry in Brazil.

In three years, vehicle sales have declined, which has affected the income and the level of profitability in the industry (Table I). In addition to the rather rigid labor legislation in relation to employee rights, the presence of labor unions has been strong, which has considerably increased labor costs in the industry.

The external environmental variables that affected PARTSCORP at the time of the study were: policy of exemption from taxes on sales of new vehicles; the interest rate defined by the Brazilian central bank for financing vehicles purchases; strikes at assemblers (i.e. PARTSCORP’s customers); and uncertainty in the policy of tax relief on wages affecting PARTSCORP’s direct and indirect labor costs.

3.2 Data collection

In this research, several steps were followed during the fieldwork. First, the choice of a supplier company in the automotive industry took into account the economic difficulties that the sector has been suffering in recent years in Brazil. The chosen company had already been the subject of previous research, and this facilitated the access. The authors also obtained the support from the controller who mediated the interaction between the research team and

<table>
<thead>
<tr>
<th></th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicles Sales (units)</td>
<td>3.8 million</td>
<td>3.5 million</td>
<td>2.6 million</td>
</tr>
<tr>
<td>Net Sales Partscorp&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>100</td>
<td>92</td>
<td>68</td>
</tr>
<tr>
<td>EBITDA Partscorp&lt;sup&gt;b&lt;/sup&gt;</td>
<td>100</td>
<td>88</td>
<td>68</td>
</tr>
<tr>
<td>EBITDA Partscorp&lt;sup&gt;b&lt;/sup&gt;</td>
<td>100</td>
<td>80</td>
<td>-454</td>
</tr>
</tbody>
</table>

*Notes:* *a* 2013 basis (100 points); *b* PARTSCORP figures are confidential. Only changes observed from 2013 can be reported.
the interviewees, but did not participate in the interview process, except in her own interview. Second, the interviews were conducted by two or three researchers who always explained the research objectives as part of their initial speech. All the interviews followed a well-defined protocol and a script. Almost all interviews were conducted in a private room to minimize distractions. Third, in addition to recording the interviews, the researchers took notes of the respondents’ answers. The interviews were also done in sufficient time and gave to the researchers the opportunity to form a good understanding of the responses.

In addition to the interviews, the researchers attended some meetings, particularly those related to continuous improvement projects.

Thus, the data-collection strategy was based on verbal reports structured through open-ended questions (Yin, 2014, p. 122), combined with a semi-structured set of questions (Sekaran, 2000). The face-to-face conversation favored the dialogue (Gustavsson, 2007, p. 227). In this research approach, the interviewer and the respondents established a productive conversation in which the primary objective was to understand the meaning of problems, the changes that occurred in the company, and the proposals made in the context of the unit of analysis. Although the interviewers had a list of questions to guide the data-collection process in this direct and face-to-face method, the main advantage of this collecting data method is that it enables the free exchange of ideas between participants (i.e., interviewers and respondents). The main goal of the dialogue technique is to understand the social reality of the participants rather than to explain it (Gustavsson, 2007, p. 235).

3.3 Data analysis

The analysis of the interview transcriptions was carried out recursively, and the interviews were often read again in order to produce relevant findings.

Moreover, in this process, the content analysis (Krippendorff, 2012; Miles et al., 2013) was the main technique used to both interpret and understand the constructs.

We also followed several procedures adopted by Brenner and Ambos (2013). First, we triangulated the data from different respondents to identify shared concepts. Second, we carefully defined the coding (Saldaña, 2016) to analyze and relate concepts and dimensions in the first and second order analysis steps, and we establish the theoretical definitions of the constructs using the literature. Finally, we went back to check the findings that emerged from our analyses with our key interviewees. In Table II, we provide the coding used for the analysis of the interviews.

The directors and managers provided different views on the primary constructs under study. The interviewees were chosen to have a complete picture and the different perspectives of flexible budgeting in the company. Table III presents general information on the interviews.

The interviews were conducted in São Paulo, Brazil, from July 2 to September 17, 2015. The interviews lasted on average one hour. They were recorded and later transcribed, thereby increasing the accuracy of data analysis and the production of the findings (Yin, 2014). On average, the interviewees had more than 15 years working in the automotive industry. Therefore, these professionals were chosen based on their knowledge and different perspectives on the main constructs under study. The main areas of interface with controllership were covered, interviewing directors, managers and supervisors. Our understanding was that when the findings began to become repetitive, no more interviews were necessary to be performed, and the processes were concluded.

4. Analysis of findings

4.1 Flexible budget influence on organizational inertia

PARTSCORP adopted a set of planning controls including long-term planning and budgeting. The budget variations were monitored using tools such as forecasts and flexible budgets.
The flexible budget is a tool already institutionalized by the company and implemented under the policies of the parent company that enables the isolation of volume variations. In this research, many interviewees attributed volume change to a non-controllable cause. This may suggest that the use of flexible budgets may contribute to organizational inertia.

<table>
<thead>
<tr>
<th>First-order concept</th>
<th>Second-order dimension</th>
<th>Third-order construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>The speech of what efforts are being made</td>
<td>Psychological inertia</td>
<td>Organizational inertia</td>
</tr>
<tr>
<td>The assertion that the whole market is in this bad situation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mentions of periods of bonanza</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The lack of aggressive posture in the commercial area</td>
<td>Action inertia</td>
<td></td>
</tr>
<tr>
<td>The approval of capex always by the matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The financial aid by the matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The support of the matrix engineering staff</td>
<td>Insight inertia</td>
<td></td>
</tr>
<tr>
<td>The constant monitoring of the matrix executives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The failure in two previous investment projects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The difficulty of placing employees in unpaid day-off</td>
<td>Workforce flexibility</td>
<td>Flexibility</td>
</tr>
<tr>
<td>The difficulty of managing reductions in electrical energy</td>
<td>Material handling flexibility</td>
<td></td>
</tr>
<tr>
<td>Deadline for placing orders at steel supplier</td>
<td>Volume flexibility</td>
<td></td>
</tr>
<tr>
<td>The definition of the product by the customer</td>
<td>Products mix flexibility</td>
<td></td>
</tr>
<tr>
<td>The sales success of the car model of the customers and tooling</td>
<td>Equipment flexibility and material handling flexibility</td>
<td></td>
</tr>
<tr>
<td>The redefinition of teams and work shifts</td>
<td>Workforce flexibility</td>
<td></td>
</tr>
<tr>
<td>Training to improve each employee’s flexibility matrix</td>
<td>Workforce and equipment flexibility</td>
<td></td>
</tr>
<tr>
<td>Review of service supplier contract scopes</td>
<td>Fixed cost flexibility</td>
<td></td>
</tr>
<tr>
<td>The weekly meetings involving the controllership and finance department</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The creation of an empathy environment provoked by the production director</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The sales volume drop</td>
<td>Flexible budget</td>
<td>Flexible budget</td>
</tr>
<tr>
<td>The separation of what is controllable within our competence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The variation of operations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The attempt to compensate for variations in volume losses</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Source:** Elaborated by the authors based on Brenner and Ambos (2013)

<table>
<thead>
<tr>
<th>Interviewee</th>
<th>Academic background</th>
<th>Function in the company</th>
<th>Interview date</th>
<th>Interview length (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Engineering</td>
<td>Customer service director</td>
<td>June 23, 2015</td>
<td>79</td>
</tr>
<tr>
<td>2</td>
<td>Engineering</td>
<td>Production director</td>
<td>June 23, 2015</td>
<td>51</td>
</tr>
<tr>
<td>3</td>
<td>Engineering</td>
<td>Production manager</td>
<td>June 23, 2015</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>Engineering</td>
<td>Production manager</td>
<td>June 23, 2015</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Business administration and finance</td>
<td>Controller</td>
<td>September 17, 2015</td>
<td>69</td>
</tr>
<tr>
<td>6</td>
<td>Engineering</td>
<td>Purchasing manager</td>
<td>July 2, 2015</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>Business administration and finance</td>
<td>Chief financial officer</td>
<td>June 23, 2015</td>
<td>70</td>
</tr>
<tr>
<td>8</td>
<td>Engineering</td>
<td>Maintenance supervisor</td>
<td>July 2, 2015</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>Engineering</td>
<td>Quality control supervisor</td>
<td>July 2, 2015</td>
<td>75</td>
</tr>
</tbody>
</table>

**Table III.** Information on the interviews
This argument can be observed in the discourse of the customer service director who points out that “[…] there are variations that we do not control, for example, the market size.” The same opinion as the purchasing manager who explained, “[…] I see that the most difficult task today is [making] a coherent budget forecast […] because we do not have […] volume stability.” Consistent with this view, the chief financial officer (CFO) also pointed out: “[Unfortunately], we had this question of volume. If the projects reached the volume that they had promised […], we would have more chances to bring money to the company.” In the same direction, the customer service director attributed the performance problems to the decrease in volume caused by the customer, “[…] the automaker needs to sell the car [so that] we can sell [our] product,” or the process manager stating that “volume variance is not controllable.”

However, we can identify other factors that stimulate organizational inertia in its different forms. Regarding psychological inertia, we noticed the sense that “what is possible is being done” (customer service director), or “good times those we were struggling for the customer not to claim that we stopped his production line […]” (customer service director). Or the speech of the CFO in which he attributes the problem of instability to the whole market and not just to the company.

Concerning action inertia, some interviewees (e.g. the controller and the production manager) mentioned that they would expect more aggressive initiatives in the commercial area, in search of new market opportunities. Other factors that contribute to action inertia are the lack of autonomy of PARTSCORP, especially in the decisions of capital expenditures (CAPEX), where final decisions are made by the parent firm. Additionally, when PARTSCORP has financial problems, the parent firm provides additional funds, creating a relationship of dependency and submission. There is also a constant monitoring of performance by the executives of the parent firm either through a weekly conference call or on-site visits. This is used by PARTSCORP to justify the difficulties to accomplish the planned budget and the absence of proactiveness.

Regarding insight inertia, some contributing factors are the support of the engineering staff of the parent firm that are continuously involved in local operations as well as the failure of two previous projects that consumed a substantial amount of resources. Thus, organizational inertia exists and can be related to the use of flexible budgets.

4.2 Low organizational flexibility

In terms of organizational flexibility, several initiatives were identified despite the difficulties encountered. For example, the challenge of putting employees on unpaid leave as well as the difficulty of managing reductions in the cost of electricity is notorious. By contrast, through managerial initiatives of the purchasing area, the teams and the work shifts were relocated, which permitted savings in the transportation costs.

Nevertheless, the different forms of organizational flexibility appear to be weak what favors inertia. Namely, concerning the material handling flexibility, it can be considered low because of the bargaining power of steel suppliers, which require that a purchase order must be placed 90 days in advance. On the other hand, the volume flexibility depends on the sales success of the automakers’ models. Therefore, this kind of flexibility is almost non-existent and also impedes to have product mix flexibility.

These difficulties are summarized in the following speech:

The room [of] maneuver to reduce fixed cost[s] is very limited […] almost null in the face of market variations. The electricity is contracted in fixed amounts; whether the volume of production falls by a half, you will pay the full bill the same way. The workmanship is the same. For example, an 800-ton machine has five employees in two production shifts, but you cannot release the surplus shift labor whether production falls (Customer service director).
4.3 Alignment between the controller and the production director

Contrary to the expectation, the isolation of volume variation acted as an incentive for production to compensate the losses in volume with cost savings and efficiency improvements through continuous improvement initiatives such as kaizen projects. The production director stated that:

I need to know how much are my losses in volume and which actions I need to [take] to sustain the losses in volume [in order] to maintain the initial budget. [...] Whether such compensation through changes in the production planning is not possible or [sufficient], then we will search opportunities for cost savings and generate more proposals and actions [for] improvement (Production director).

The reasons and consequences of the negative variation that resulted from the decrease in sales volume were understood in a different way when the production director is compared with the other interviewees. Thus, also the role of the flexible budget may be interpreted in a different way. The production director’s professional experience and training had an impact on this perspective:

My belief is [based on] what I had learned in my former professional experience when I joined the company X. I began in the financial area. I worked as a controller. The budget is the budget. I had a boss there who liked to say: “The budget is the budget, which you never change and should never forget.” You update the figures, but always remembering that there is a budget to comply with (Production director).

This vision of the production director was in line with the controller’s view that the organization was “a little bit accommodated” to negative results. In his words:

[…] we have to think and change something because the bottom line is not good. […] We need to have a more aggressive stance […] go [to] other clients that we do not have today […] enter [other] markets […] (Controller).

Thus, in this case study, at the moment, the production and controllership areas are aligned, and that happens because the normative alignment of the controller and the production director. But, in the past, the controller was not happy with her own performance because she could not effectively work with the production department. Though this situation only improved with the hiring of the current production director in 2011. In their views, “[…] the controllership and production began to cooperate [with] each other” (production director), and with “[…] the arrival of the production director, he brought more organization to the operational area, mainly in processes, Kaizen, and cost control” (controller).

In the past, there was a problem relating to interaction and collaboration. This situation was pointed out by the production manager:

“[…] The manufacturing was [one] world and the financial department was another one.” Today this interaction is conducted on a daily basis, as pointed out by the production manager who called this process the “daily report.” In his view: “[…] whether we have, for example, a reduction in the volume, we […] immediately [have] a meeting to discuss cutting expenses in the same proportion” (Production manager).

4.4 Continuous improvement initiatives contribute to organizational flexibility

Trying to minimize losses due to the decrease in sales volume, the production director implemented several continuous improvement projects (Kaizen). In 2015, the company had about 120 Kaizen projects in progress and became the benchmark within the Group. The projects’ approval focused on cost savings and also non-monetary recognition was given to all those who proposed project ideas.

Even though we noticed a low degree of flexibility due to the rigid labor legislation, we also saw some initiatives that improved the level of flexibility, particularly of equipment and materials.
For example, the decision to internalize the manufacturing of some components and tooling, with effects on labor exploitation and fixed costs savings.

We also noted that the training to improve each employee’s skillset impacted on the workforce and equipment flexibility. After this project, more than 70 percent of the production employees could work in another production cell thanks to their differentiated training. In short, the firm has developed a matrix of flexibility.

According to the process manager, the decision of manufacturing their tools instead of buying them from suppliers has improved production flexibility. In his words:

We internalized the production of many devices. Before, we depended on sending raw materials to the supplier who would produce the component. Today, we rely only on ourselves. So, if we have available raw material […] we can make our planning and production more flexible. This change was a significant gain (Process manager).

The production manager also explained the development of the reasoning behind this situation:

We realized that our sales volume was decreasing […] and the expenses with suppliers did not fall proportionately. So, since we have equipment and idleness here, we decided to bring those parts to be produced internally. Therefore, we reduced the idleness, and the workforce had virtually not increased. […] Whether we have raw material, we are able to manage it better in the production stage (Production manager).

5. Discussion
This research shows that to address the contradictory role of the use of flexible budgets and the latent conflict between organizational inertia and flexibility, one viable approach is to focus on continuous improvement projects, which can help minimize profit reduction. Even though the impact on profit was not significant in this case, continuous improvement initiatives helped to maintain a high employee morale. However, this was only possible because the controller supported those projects and helped the person in charge of them (i.e., the production director in this case study). In short, these results show that continuous improvement programs can make the organization more flexible (Chan and Tay, 2018) while breaking down the organizational inertia. By contrast, these findings are not aligned with Moilanen (2008), wherein contact was infrequent between the parent firm and the subsidiary in the Soviet Union, in a context where budgetary controls were also emphasized. But the results are consistent with Brenner and Ambos (2013) in that face-to-face contact is typical in an organizational culture denominated by the builder type. Our findings are also in line with Schäffer et al. (2014, p. 32) findings that argue that “visits and the discussions and meetings of headquarters management with local management might be more effective than autonomous discussions on the part of subsidiary managers based on outputs generated by parent PMS.”

Furthermore, the new role of controllership in PARTSCORP confirms Tillmann and Goddard’s (2008) findings that controllers require organizational learning and socialization to improve organizational sensemaking. It is interesting to note that this new controlling role was made possible by the entry of a professional from the production area and not the financial department, as pointed out by Goretzki et al.’s (2013) study.

Similarly, this finding is also in line with Kihn (2011) who added new factors that are relevant to both managers and the budgeting process: position in the organizational structure, type of function, education and previous professional backgrounds. However, Kihn’s (2011) findings showed that the controllers and those who are controlled could interpret the organizational budgetary processes differently. Such a situation may happen because they have personal visions of the budgetary targets.

This case study shows the opposite, i.e., under certain conditions, different managers can be aligned and share a similar interpretation of the flexible budget. This result differs from
Kihn’s (2011) analysis that points out the difficulty to find similar understandings of budget targets within an organization because these understandings are very personal and subjective, and even contradictory.

In this case, paradoxically, two organizational agents who are typically in conflict – the controller and the production director – showed strong alignment. This was mainly motivated by the production director who changed the controllership’s role by pressing the controller to act as a business partner of the production area. The findings go beyond the proposition of Kihn (2011) who states that professionals with different normative profiles and academic backgrounds have different budget views and can contribute to stimulating different institutional logics. That is, professionals with experience in companies that have rigid budget control tend to have opinions opposite to those who have not been influenced by such rigid approaches. As a result, these two groups tend to challenge organizational structures in order to make sure their institutional logic prevail.

This study is consistent with Näslund and Pemer’s (2012) findings, in which the institutionalization of the flexible budget by the parent company, a tool that isolated the negative budgetary variations resulting from the sales volume drop, helped to legitimize the discourse of some managers. We note that the manager’s speech attributes the poor performance presented by the company to external variables. Thus, constituting a dominant institutional logic (Dambrin et al., 2007; Lambert and Sponem, 2011; Tan and Wang, 2011) and leading to organizational inertia (Näslund and Pemer, 2012). On the other hand, the shift in sales volume should not be an excuse to the company does not invest in changing the status quo such as launching new products in the market.

The results also showed that two organizational actors were averse to this discourse, which is an example of a political coalition (Hannan and Freeman, 1984). In this case, the production director and the controller did not accept this discourse and began to lead a group whose beliefs went against the dominant institutional logic (Dambrin et al., 2007). The movements of these two actors helped stimulate organizational flexibility, offsetting the adverse changes caused by the sales volume through continuous improvement projects and intensive budget monitoring.

The production director and the controller had at least two points in common. First, the cognitive-cultural training based on a more rigorous budgeting style. These two professionals were accustomed to rigorous control practices because they had already worked in companies with this profile. Second, that PARTSCORP must have an essential role in bridging control and production.

It is interesting to note that with the arrival of the production director, the integration between production and the controller department has increased considerably. As a result, the involvement of these actors in the process of learning and socialization in the production area has reinforced the role of controllership in shaping organizational sensemaking (Tillmann and Goddard, 2008).

This compromise of the production director regarding the budget’s bottom line seems to be one of the largest motivations for pursuing continuous improvement through Kaizen techniques and projects. Moreover, it changed how management reports were organized. Additionally, the introduction of Kaizen or continuous improvement initiatives, as a routine, was aimed at cost savings. Such monitoring helped the company to experience greater flexibility in managing resources, especially those related to human resources. These controls were enhanced via the strict control of cash flow. For example, the cash flow was tightly controlled by the weekly meetings, which were attended by key managers and were monitored in meetings with the parent company’s executives via teleconferencing. These procedures encouraged initiatives such as inventory management, negotiating delivery schedules with suppliers and so on. For example, the difficulties in the management of the raw material inventories were due to orders with small delays in deliveries by the suppliers.
6. Conclusions
The objective of this paper was to investigate how the use of flexible budgets influences organizational inertia and flexibility in a firm that has experienced a substantial drop in sales volume.

The results of this work are useful to practitioners, particularly for firms that adopt the flexible budget and are experiencing a significant drop in sales volume. Additionally, the results show that the use of this tool, when legitimized by the parent firm, can induce organizational inertia. However, in the case studied the interaction and alignment between the production and controllership areas minimized such inertia through initiatives that improve organizational flexibility, especially through continuous improvement projects (Kaizen projects).

The contribution to the literature of both management accounting and operations is threefold. First, this is an early study that investigates how organizational inertia and flexibility, two conflicting constructs, are influenced by the adoption of flexible budgets, particularly when a firm is experiencing a significant drop in sales volume. Second, the results show that the controller and the production director, who are typically in conflict, can work together to improve flexibility and interaction through Kaizen projects. Third, the research shows how the flexible budget, a managerial control tool widely used in practice but under-researched, is used.

This study was based on only one case study and, therefore, the results cannot be generalized. Nevertheless, the findings contribute to the literature of both management accounting and operations.

For future research, we may highlight the study of the implementation of the flexible budget in other environments. Finally, the lessons learned from this case study are useful for practitioners and firms in similar situations, mainly to make the flexible budget more effective and efficient in parent-subsidiary relationships.

References


Corresponding author
José Carlos Tiomatsu Oyadomari can be contacted at: oyadomari@mackenzie.br

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
A transformed fixed effect stochastic frontier approach for productivity evaluation in Indian electricity sector

Pavan Khetrapal
Department of Electrical and Electronics Engineering, Krishna Institute of Engineering and Technology School of Management, Ghaziabad, India

Abstract

Purpose – The purpose of this paper is to evaluate the utility-level productivity changes in Indian electricity sector during a period that witnessed structural reforms through several landmark regulatory changes.

Design/methodology/approach – A transformed fixed-effect stochastic frontier panel approach accounting for time-invariant unobserved heterogeneity is employed to evaluate the productivity changes, and the inefficiency level in 98 utilities spanning over the years 2001–2010. A flexible translog production model is modeled and estimated, and decomposition of productivity into components of changes in efficiency, scale technology and price effect is computed.

Findings – The empirical findings obtained from the present study suggest that the utility-level productivity in Indian electricity sector has generally declined during the observed period of 2001–2010 specifically after the implementation of Electricity Act 2003. Also, it is estimated that the state-level un-bundling of the electricity sector is not significantly associated with utility-level efficiency change. Furthermore, efficiency improvements attributable to increased competition are observed only in the case of smaller gas-based generating utilities.

Originality/value – Earlier studies on the productivity evaluation of Indian electricity industry have applied the non-parametric data envelopment analysis approach, which has several limitations. The novelty of the paper lies in the fact that this paper is one of the first attempts that implement transformed fixed-effect stochastic frontier panel approach and thus disentangle unobserved heterogeneity from inefficiency. Furthermore, it is the only paper that analyzes 98 utilities (51 generating utilities, 38 transmission and distribution licensees and 9 vertically integrated utilities) in a single framework during the period 2001–2010.

Keywords Reform, Unobserved heterogeneity, Stochastic frontier analysis, Firm-level panel data, Indian electricity sector, Total factor productivity

Paper type Case study

1. Introduction

In the past two decades, the Indian power sector has witnessed structural reforms through several landmark regulatory changes. Along the lines of power sector reforms elsewhere internationally, primarily reforms in India also emphasized un-bundling of vertically integrated utilities into functionally separate entities dealing with generation (production) and transmission and distribution (T&D) (service). The reforms also attempted to attract private capital to the sector. The primary policy objectives of initiating reforms in the sector were anticipated efficiency gains and cost reduction. Therefore, an empirical assessment of utility-level productivity changes in the generators and T&D utilities shall reveal the extent to which reforms have influenced the sector in the intended direction. However, pan-India measurement of productivity changes across the power sector value-chain poses two key challenges. First, utility-level heterogeneity due to diversity in geography, local regulations, technology employed and other unobserved factors makes pan-sector (and pan-India) measurements prone to omission-bias. Second, due to relatively lax regulatory requirements of central collection and maintenance of utility-level operating data in India, productivity measurements have to rely on data collated from multiple sources or estimated from aggregate numbers. Hence, a majority of extant research studies investigating efficiency...
changes in utilities in the Indian power sector have been confined in scope to specific geography, firm or technology.

In this context, an attempt to estimate pan-India firm-level productivity changes in the generating utilities, T&D licensees and a few remaining vertically integrated utilities is made. In the present work, the empirical strategy to measure pan-sector efficiency change is to structurally control for utility-level heterogeneity. Causal inference is crucial specifically in studies attempting to attribute observed efficiency differences to explanatory factors. For instance, investigating the influence of un-bundling on the performance of Indian power sector, Cropper et al. (2011) employ the difference-in-difference econometric technique. The method adjusts for omitted variable bias caused due to unobserved variables aggregated at the state and time period level. Therefore, by design, the method provides the regression coefficients on explanatory variables close to causal interpretation. In the present study, so as to enable causal inference in panel SFA models that jointly estimate inefficiency and the exogenous determinants of inefficiency (models following B&C, 1995), it is critical to control for unobserved heterogeneity. In a recent comparative investigation of methods, Kopsakangas-Savolainen and Svento (2011) measure inefficiencies in the Finnish electricity distribution utilities using five different SFA models. The study reports that models accounting for unobserved heterogeneity produce lower inefficiency measures and considerably different inefficiency rank orders. Thus, ignoring unobserved heterogeneity could result in confounded regression coefficients with severely limited causal interpretation. Greene (2005a, b) suggests two new methods for controlling unobserved heterogeneity in panel SFA models: the “true fixed-effects” model and the “true random-effects” model. The problem of identification is addressed in these newer models by structurally constraining the positive inefficiency term to be time-varying and the unobserved heterogeneity to be time-invariant. However, Wang and Ho (2010) point out that the “true fixed-effects” SFA model suffers from the problem of incidental parameters (Neyman and Scott, 1948; Lancaster, 2000) that might contaminate estimates of other model parameters when simultaneous estimation of fixed effects and the inefficiency variance parameter is attempted. Wang and Ho (2010) suggests a solution to this problem by developing a model that enables elimination of unobserved fixed-effect variables (either by within or difference transformation) prior to estimation of inefficiency[1].

In the present study of the Indian power sector during the reforms period, Wang and Ho (2010) transformed SFA model is therefore adopted to disentangle unobserved firm-level heterogeneity from technical inefficiency. An empirical investigation about the nature of productivity changes in 98 utilities operating in the Indian power sector during the period of 2001–2010 is undertaken in the present work. The sample in the present study represents 51 generating utilities, 38 T&D licensees and 9 vertically integrated utilities with a total of 542 firm-year observations. The unbalanced panel of sample contains observations of utilities that are under the ownership of central government, state government and private investors. Furthermore, the sample is fairly representative and accounts for 45.7 percent of total electricity generated and 59.4 percent of total electricity consumed in India during the period of 1999–2010. Using a flexible translog production specification, the measure of productivity changes is decomposed into components of changes in technology, efficiency, scale and price effects. Based on the utility-level sample, it is estimated that after Electricity Act 2003, there had been no improvements in utility-level productivity. In addition, bulk of the productivity change is attributable to technology change (newer capacity addition), whereas efficiency is observed to be generally declining.

2. Deregulation and measuring utility-level productivity change
The global wave of restructuring since the early 1990s systematically brought about significant changes in industry structure, ownership pattern and mode of regulation in the
electricity sector in several countries. A common feature in many of these reforms initiatives is the dismantling of vertically integrated utilities, thus, separating generation of electricity (production) from T&D (service), such that coordination of demand–supply, after restructuring, happens over a specialized market-based mechanism. It is suggested that the introduction of competition and market-based transactions in the sector is made on *ex ante* anticipation of improvement in technical efficiency, reduction in operating costs and hence positive welfare gains (Joskow and Schmalensee, 1983). Wolfram (2005) argues that restructuring would lead to efficiency gains because of: the new incentives faced by the incumbents to improve efficiencies; takeover of inefficient older plants and the arrival of new entrants with newer technologies; and competition driving efficient allocation of factors of production. Thus, in the restructured industry setting, *ex post* measurement of efficiency improvements for utilities across the electricity value-chain provides a basis for critical empirical scrutiny of the reforms policy.

In this context, several extant studies investigate the influence of restructuring on firm-level productivity changes in the electricity sector. Among alternative methods, the non-parametric data envelopment analysis (DEA) is a popular technique employed for the measurement of efficiency in the power sector, e.g. Vaninsky (2006). However, with access to panel data, the SFA method presents a natural way to incorporate information obtained from multiple observations of the production unit spread over time. And among other things, it also allows for a richer production specification and formal statistical testing of hypotheses (Hjalmarsson *et al.*, 1996). SFA is used by Knittel (2002), using a Cobb–Douglas specification, and by Hiebert (2002), using the more flexible translog specification for measuring efficiency of US power generators. Knittel (2002) partly controls for firm-level fixed effects by incorporating plant vintage information in the SFA model. The SFA model used by Hiebert (2002) follows Battese and Coelli (1995) to jointly measure inefficiency and the influence of firm-level factors associated with the observed heterogeneity in measured inefficiency. The study identifies capacity utilization and ownership as factors influencing efficiency, overall mixed results on efficiency gains are seen across the restructured US states.

out that there are substantial state-level differences in improvements attributable to heterogeneity in historic as well as political context. Cropper et al. (2011) find that while un-bundling has not resulted in improvements in thermal efficiency, there has been improvements in capacity utilization (+4.6 percent) and reduction in forced outages (−2.9 percent). Table I summarizes the different studies that analyze the efficiency and/or productivity of the Indian electricity sector.

It is concluded from the above discussion that the extant research has focused either at the aggregate state level or to level of generating plants. However, economically important decisions of investment in capacity, technology and choice of factor allocations are made at the level of the utility that operates several productive assets under its ownership and control. Especially after un-bundling of generation from distribution and transmission, the role of the utility as the decision-making entity is more salient. Hence, the present empirical work focuses on the utility as the unit of analysis to measure dynamic changes in efficiency at the utility level. The study also evaluates the extent to which the changes are driven by factors such as ownership differences, vintage of assets, un-bundling status of the state in which the utility is operating and competition.

3. Indian power sector reforms

The Indian Electricity Act 1910, promulgated primarily to ensure safety, was the earliest legislative attempt to regulate the working of the Indian electricity industry. After independence, the Indian Constitution accorded concurrent status to the electricity sector, thus, placing it simultaneously under the responsibility of the central and state governments. Subsequently the Electricity (Supply) Act 1948 came into effect that paved the way for the formation of vertically integrated government owned agencies: the State Electricity Boards (SEBs), entrusted with the responsibility of generation, T&D of electricity in the respective states. However, the power sector continued to be characterized by capacity underutilization, inefficient operations and financially imprudent pricing policies. This consequently leads to chronic shortages, poor quality of supply, frequent breakdowns and bankruptcy of the SEBs. The genesis of reforms in the power sector can be broadly traced to these deteriorating conditions. Arun and Nixson (1998) present a detailed discussion on the nature and genesis of reforms, beginning the amendments to the Electricity Act 1910 and Supply Act 1948 in the year 1991, which primarily opened up the sector to private investments. Subsequently, in 1998, the electricity Regulatory Commissions Act was enacted resulting in the setting up of the Central Electricity Regulatory Commission (CERC) and other state-level regulatory bodies. While primarily CERC was concerned with the regulation of firms owned and operated by the central government, it also regulated coordination activities for entities spanning multiple states. However, these early attempts hardly made any substantial impact on the growth and recovery of the power sector. In a critical examination of this early phase of reforms D’Sa et al. (1999) and Dubash and Rajan (2001) highlight that the piecemeal approach to reforms failed to rein in the progressively languishing state of the power industry.

In contrast to these earlier fragmented reform attempts, a paradigm shift in the power sector reform process was brought about by the legislation of the Electricity Act 2003 (GOI, 2003) in July 10, 2003. The Electricity Act 2003 replaced and consolidated the existing legislations aiming for substantial structural changes in the Indian electricity industry. The salient features of the Act included de-licensing of thermal and captive generation, de-licensing of distribution in rural areas, open access to transmission, phased open access to distribution, multiple licenses in distribution and recognition of electricity trading as a distinct activity enabling the creation of electricity markets. A detailed exposition of the implications of Electricity Act 2003 for generation, transmission, distribution and electricity trading can be found in Bhattacharyya (2005),

1660
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings of the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yadav et al. (2013)</td>
<td>Operating and maintenance (O&amp;M) cost (Rs. million); number of employees</td>
<td>Energy sold (million units); number of customers; distribution line length (circuit kilometer); transformer capacity (kVA); average duration of interruption per feeder (hours)</td>
<td>29 Electricity Distribution Utilities (EDU) in Uttarakhand state in 2007</td>
<td>DEA</td>
<td>The study diagnoses the efficiency changes of EDU so as to help regulators to formulate more effective policies on deregulation and disintegration and to determine the appropriate efficiency improvement measures when imposing yardstick competition on electric utilities</td>
</tr>
<tr>
<td>Shrivastava et al. (2012)</td>
<td>Coal consumption – primary fuel (thousand tons); fuel oil consumption – secondary fuel (kilo liter); auxiliary power consumption (MWh)</td>
<td>Electricity generation (MWh)</td>
<td>60 coal-fired power plants in India for the year 2008-2009</td>
<td>DEA</td>
<td>The result indicates poor performance of few power plants due to over use of input resources. Finding further reveals that efficiency of small power plants is lower in comparison to medium and large category and also performance of state-owned power plants is comparatively lower than central and privately owned</td>
</tr>
<tr>
<td>Yadav et al. (2011)</td>
<td>Operating and maintenance (O&amp;M) cost (Rs. million); number of employees</td>
<td>Energy sold (million units); number of customers; distribution line length (circuit kilometer); transformer capacity (kVA); average duration of interruption per feeder (hours); total sanctioned load/square kilometer</td>
<td>29 Electricity Distribution Divisions (EDDs) of an Indian state – Uttarakhand</td>
<td>DEA</td>
<td>The results indicate that the performance of several EDDs is sub-optimal, suggesting the potential for cost reductions and possible reduction in number of employees. The empirical findings also reveal that plain area divisions are relatively efficient and have higher potential to influence the performance of inefficient EDDs</td>
</tr>
<tr>
<td>Cropper et al. (2011)</td>
<td>Nameplate capacity (MW); generation (MUs); age (years)</td>
<td>Forced outages (%); planned maintenance (%); availability (%); capacity utilization (%)</td>
<td>82 thermal power plants during the year 1994–2008</td>
<td>SFA</td>
<td>The results suggest that the gains from un-bundling of generation from transmission and distribution were limited to the states that reformed before the Electricity Act of 2003</td>
</tr>
<tr>
<td>Yadav et al. (2011)</td>
<td>Operating and maintenance (O&amp;M)</td>
<td>Energy sold (million units); number of</td>
<td>29 Electricity Distribution</td>
<td>DEA</td>
<td>The results indicate that numerous divisions have scope for improvement in overall efficiency, and most of the utilities are</td>
</tr>
</tbody>
</table>

Table 1. Summary of studies on Indian electricity sector
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings of the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sen and Jamasb (2010)</td>
<td>cost (Rs. million); number of employees</td>
<td>customers; distribution line length (circuit kilometer); transformer capacity (kVA)</td>
<td>Divisions (EDDs) of an Indian state – Uttarakhand</td>
<td>SFA</td>
<td>inefficient due to their scale inefficiency rather than technical inefficiency</td>
</tr>
<tr>
<td></td>
<td>Plant load factor (%); gross generation (MUs); T&amp;D losses (%)</td>
<td>State industrial GDP (Rs. million); state industrial GDP per capita (Rs.); ratio of industrial to domestic price of electricity (Rs.); ratio of industrial to agricultural price of electricity (Rs.); industrial consumption of electricity in a state (MUs)</td>
<td>19 state utilities for the year 1991–2007</td>
<td></td>
<td>The empirical findings deduce that Indian electricity reforms have negatively affected the efficiency of the distribution sector</td>
</tr>
<tr>
<td>Singh (2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>It is a qualitative study undertaken by the researcher to explore the emergence of competition in the Indian power sector</td>
</tr>
<tr>
<td>Thakur et al. (2006)</td>
<td>Model A: total cost (Rs. million); number of employees</td>
<td>Model A: energy sold (Mkwh); customers (million); distribution line length (circuit km)</td>
<td>26 State-Owned Electric Utilities (SOEUs) in India</td>
<td>DEA</td>
<td>The study provides an overview of the status of competition in various segments of the power sector and reviews the phased program for open access outlined by various State Electricity Regulatory Commissions, including their status and impact on competition. The paper also highlights a number of residual issues to enable wholesale and retail competition, and discusses the approaches to address the same</td>
</tr>
<tr>
<td></td>
<td>Model B: adjusted cost (Rs. million); number of employees</td>
<td>Model B: energy sold (Mkwh); customers (million); distribution line length (circuit km)</td>
<td></td>
<td></td>
<td>The results indicate that the performance of several SOEUs is sub-optimal, suggesting the potential for significant cost reductions. The empirical findings also reveal that the bigger utilities display greater inefficiencies and have distinct scale inefficiencies</td>
</tr>
<tr>
<td>Thakur et al. (2005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>It is a qualitative study undertaken by the researchers to analyze the impact of Electricity Act 2003 on various segments (i.e. generation, transmission and distribution) of Indian electricity sector</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The study analyzes the probable impacts and implications of the Electricity Act 2003 that are likely to have far-reaching consequences for the Indian power sector</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Inputs</th>
<th>Outputs</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings of the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shanmugam and Kulshreshtha (2005)</td>
<td>Power generated (GWh)</td>
<td>Capital employed; specific coal consumption in tons; specific secondary oil consumption; auxiliary power consumption in % terms</td>
<td>56 coal-based thermal power stations, distributed over various geographical regions of India over the period 1994–1995 to 2001–2001</td>
<td>SFA</td>
<td>The empirical findings indicate that there exists substantial scope for increasing thermal power generation in the country, with improved application of existing technology and without employment of additional resources. The results also signify that the western region is technically more efficient than other regions and young plants are more efficient than their old counterparts</td>
</tr>
<tr>
<td>Khanna et al. (1999)</td>
<td>Age of the BTU (years); ownership of the plant; manufacturer of the plant; capacity utilization factor</td>
<td>Annual plant-level auxiliary consumption (MkWh); annual heat energy input (billion kcal)</td>
<td>63 coal-based power plants in India for the four-year period, 1987–1988 to 1990–1991</td>
<td>SFA</td>
<td>The study explores the sources and magnitude of energy inefficiency in the electricity-generating sector in India and its implications for carbon emissions. The analysis also demonstrates the potential for institutional and economic policy reforms that provide incentives for the adoption of efficiency-enhancing production practices to reduce carbon emissions</td>
</tr>
<tr>
<td>Khanna et al. (1999)</td>
<td>Ownership of the power plant; age of the power plant; non-utilized capacity factor</td>
<td>Annual cost of electricity generated; annual net electricity generation by a plant in (kWh)</td>
<td>66 power plants in India for the period 1987–1988 to 1990–1991</td>
<td>SFA cost frontier model</td>
<td>The empirical analysis shows that publicly owned plants are more inefficient than privately owned plants and that capacity utilization is a significant determinant of inefficiency in the electricity-generating industry in India</td>
</tr>
</tbody>
</table>
While the reforms started in the early 1990s, structurally significant changes where set in motion only after the enactment of the Electricity Act 2003, especially in terms of potential to influence the technological efficiency of firms operating in the electricity sector. The Act specifically articulates intention to instill competition in the industry and outlines the framework to transit from vertically integrated monopoly structure to a multi-buyer and multi-seller model. With the establishment of wholesale electricity market, the institutional framework for a competitive industry structure got further augmented. Similar to observations by Ranganathan (2004) and Singh (2010), the author also expects that the series of structural changes in the electricity sector, especially during the decade starting year 2001, demonstrates potential to incentivize firms to improve technological as well as operational efficiency. In addition, given wide variation in ownership structure, local regulation and historical context, substantial pan-India heterogeneity in firm-level productivity outcome is expected, in response to these institutional incentives. It is in this overarching institutional context that the present study empirically investigates a substantial duration of the reform period (2001–2010), attempting to measure the extent of overall productivity improvement in the sector and identify exogenous observable factors responsible for utility-level heterogeneity in outcomes.

4. Data variables
A sample data set of Indian power-generating utilities and T&D licensees for the period of 2001–2010 is created in the present work. The sample represents about 46 percent of total generation and about 60 percent of total electricity consumption in India during the period. The sample spans across 19 states and represents ownership under central government, state government and private investors. Variable definition, unit of measurement and respective sources of data are summarized in Table II.

Power-generating firms are classified as “coal-based,” “gas-based” or “mixed” depending on the type of fuel consumed most. Utilities with generating assets using coal, gas and other sources with no one dominant fuel type is classified under the “mixed” category. Similarly, utilities engaged only in T&D function are classified as “distribution utilities” and utilities operating generating companies as well as engaged in T&D are classified as “vertically integrated.” The distribution of utilities across the various categories is described in Table III and summary statistics for all the data variables is shown in Table IV.

The unit of fuel input is normalized to energy equivalent GWhr units. From Table III, the ratio of electricity generated to fuel input shows an aggregate input–output efficiency of about 28 and 26 percent for coal and gas-based generators, respectively. Transmission loss estimated from the distribution utilities is about 28 percent. These estimates of aggregate efficiencies conform well with other estimates based on plant-level measurements like CEA (2008).

5. Methodology
Transformed fixed-effect stochastic frontier model
Initially, a primal stochastic production frontier using a deterministic kernel $f(x_{nit}, t; \beta_n)$ producing a scalar output $y_{it}$ is shown as given in the following equation:

$$y_{it} = f(x_{nit}, t; \beta_n).\exp(\epsilon_{it}),$$  (1)

for the $i$th producer $i = 1, \ldots, I$ during time period $t = 1, \ldots, T$ using inputs $x_{ni}, n = 1, \ldots, N$, where $e_{it}$ represents producer-specific time-varying stochastic inefficiency term. A flexible translog form, developed in Christensen et al. (1971) and Christensen and Jorgenson (1973), to express the time-varying stochastic production function in Equation (1) is chosen.
<table>
<thead>
<tr>
<th>S. No.</th>
<th>Variable</th>
<th>Units</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Electricity output</td>
<td>GWhr</td>
<td>Total electricity generated or distributed. Computed by dividing the reported revenue from operations by yearly average region-wise electricity for each type of generating technology. In case of T&amp;D and vertically integrated utilities the region-wise yearly average retail electricity prices are used.</td>
<td>(a) Company operating revenue reported in annual reports obtained from CMIE PROWESS. (b) Electricity retail prices obtained from TEDDY year-book³.</td>
</tr>
<tr>
<td>2</td>
<td>Capital deployed</td>
<td>Indian million rupees (INR)</td>
<td>Gross fixed assets (real) deployed. Current values deflated by GDP (1999 = 100). Computed for a period by adding Net-fixed assets of the period with cumulated depreciation till that period. Computed by dividing the total reported employee expenditure by the yearly average estimated wages in the power sector in India.</td>
<td>(a) Company assets and depreciation reported in annual reports obtained from CMIE PROWESS. (b) GDP obtained from World Bank Development indicators³.</td>
</tr>
<tr>
<td>3</td>
<td>Labor employed</td>
<td>Numbers</td>
<td>Computed by dividing the total reported employee expenditure by the yearly average estimated wages in the power sector in India.</td>
<td>(a) Company total employee expenditure reported in annual reports obtained from CMIE PROWESS. (b) Yearly average wages in power sector estimated by a smaller sample of firm data reporting both number of people employed and the total expenditure on labor. This smaller sample is obtained from CMIE PROWESS and DATASTREAM financial data.</td>
</tr>
<tr>
<td>4</td>
<td>Fuel consumed: coal</td>
<td>GWhr energy equivalent of coal used</td>
<td>Computed by dividing the total reported expenditure on fuel/raw-material by the yearly average estimated purchase price of coal in each region. An average calorific value of 4,000 KCal/Kg or 4,648.9 KWhr/metric ton is assumed for coal.</td>
<td>(a) Company total expenditure on fuel/raw-material reported in annual reports obtained from CMIE PROWESS. (b) Yearly average purchase price of coal in power sector estimated by a smaller sample of firm data reporting both quantity of fuel and the total expenditure on labor. This smaller sample is obtained from CMIE PROWESS and DATASTREAM financial data.</td>
</tr>
<tr>
<td>5</td>
<td>Fuel consumed: natural gas</td>
<td>GWhr energy equivalent of natural gas used</td>
<td>Computed by dividing the total reported expenditure on fuel/raw-material by the yearly average estimated purchase price of gas in each region. An average calorific value of 40 Mjoule/m³ or 11.11 KWhr/m³ is assumed.</td>
<td>(a) Company total expenditure on fuel/raw-material reported in annual reports obtained from CMIE PROWESS. (b) Yearly average purchase price of natural gas in power sector estimated by a smaller sample of firm data reporting both quantity of fuel and the total expenditure on fuel from each region. This smaller sample is obtained from CMIE PROWESS.</td>
</tr>
<tr>
<td>6</td>
<td>Electricity input</td>
<td>GWhr</td>
<td>Computed by dividing the total reported expenditure on fuel/electricity purchased by the yearly average sale price to utilities in each region.</td>
<td>(a) Company total expenditure on fuel/electricity purchase reported in annual reports obtained from CMIE PROWESS. (b) Yearly average sale price of electricity by coal- and gas-based generators separately. Estimated by a smaller sample of firm data, from CMIE PROWESS, reporting both quantity of electricity sales and the total revenue from electricity sales from each region. The smaller sample of firm data obtained from CMIE PROWESS.</td>
</tr>
<tr>
<td>7</td>
<td>Coal price</td>
<td>INR per metric ton</td>
<td>Region-year average purchase price of coal in power sector estimated by a smaller sample of utility data reporting both quantity of fuel and the total expenditure on fuel from each region.</td>
<td>(continued)</td>
</tr>
<tr>
<td>S. No.</td>
<td>Variable</td>
<td>Units</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>8</td>
<td>Gas price</td>
<td>INR per cubic meter</td>
<td>Region-year average purchase price of natural in power sector estimated by a smaller sample of utility data reporting both quantity of fuel and the total expenditure on fuel from each region</td>
<td>The smaller sample of firm data obtained from CMIE PROWESS</td>
</tr>
<tr>
<td>9</td>
<td>Sale price of electricity by coal-based generators</td>
<td>INR per KWhr</td>
<td>Region-year average sale price of electricity by coal-based generators. Estimated by a smaller sample of firm data reporting both quantity of electricity sales and the total revenue from electricity sales from each region</td>
<td>The smaller sample of firm data obtained from CMIE PROWESS</td>
</tr>
<tr>
<td>10</td>
<td>Sale price of electricity by gas-based generators</td>
<td>INR per KWhr</td>
<td>Region-year average sale price of electricity by gas-based generators. Estimated by a smaller sample of firm data reporting both quantity of electricity sales and the total revenue from electricity sales from each region</td>
<td>The smaller sample of firm data obtained from CMIE PROWESS</td>
</tr>
<tr>
<td>11</td>
<td>Retail price of electricity</td>
<td>INR per KWhr</td>
<td>State-year average sale price of electricity by utilities</td>
<td>Electricity retail prices obtained from TEDDY yearbook</td>
</tr>
<tr>
<td>12</td>
<td>Capital price</td>
<td>Percentage</td>
<td>Computed as: (a) Price of Capital = Expense of Capital/Gross Fixed Assets; (b) Expense of Capital = Interest Share of Capital + Depreciation; (c) Interest Share of Capital = (Annual Interest on Long-Term Debt) × (Fixed Assets)/(Long-Term Debt)</td>
<td>Company financial data in annual reports obtained from CMIE PROWESS</td>
</tr>
<tr>
<td>13</td>
<td>Vintage</td>
<td>Year</td>
<td>The year of incorporation of the company is taken as a proxy for the approximate vintage of the firm’s productive assets</td>
<td>Company annual reports</td>
</tr>
<tr>
<td>14</td>
<td>Time since unbundling</td>
<td>Year</td>
<td>Years past since the vertically integrated State Electricity Boards (SEBs) were unbundled and separated as generators and distribution utilities in the respective states</td>
<td>TEDDY yearbook</td>
</tr>
<tr>
<td>15</td>
<td>Index of competition in power sector in the state</td>
<td>Index number</td>
<td>Measure of the competitiveness of the state power sector. Score of 0–40, higher score indicating more sustainable revenue model</td>
<td>Ministry of Power, Government of India</td>
</tr>
</tbody>
</table>

**Notes:** *CMIE PROWESS: the PROWESS database of Indian companies details maintained by the Center for Monitoring of Indian Economy, Mumbai India; TEDDY: TERI Energy Data Directory and Yearbook, annual publication of The Energy and Resource Institute, New Delhi India; World Bank Development Indicators: maintained at http://data.worldbank.org/data-catalog/world-development-indicators; Ministry of Power Report: www.powermin.nic.in/indian_electricity_scenario/pdf/Final_Report_Rating.pdf*
Table III. Description of sample

Table IV. Summary statistics of data variables
The translog production function is a local second-order approximation to any arbitrary production function, and thus, displays several desirable properties for empirical estimation. The translog specification places no prior functional constraints on returns to scale, elasticity of substitution between inputs and homotheticity. Christensen and Jorgenson (1973) discuss the aforementioned and other related properties of the translog production function in detail. Additionally, Diewert (1976) shows the translog form to be "exact" for the Divisia (1925) index. Subsequently, the present study used the Divisia index to estimate efficiency change and productivity change over the period of study. For the sample of unbalanced panel data on \( I \) producers over \( T \) time periods considered in the present study, the production function in the translog form is expressed as given in the following equation:

\[
\ln E_{it} = \alpha_i + \beta_{K} \ln K_{it} + \beta_{L} \ln L_{it} + \beta_{F} \ln F_{it} + \frac{1}{2} \beta_{KK} \ln K_{it}^2 + \frac{1}{2} \beta_{LL} \ln L_{it}^2 \\
+ \frac{1}{2} \beta_{FF} \ln F_{it}^2 + \beta_{KL} \ln K_{it} \ln L_{it} + \beta_{KF} \ln K_{it} \ln F_{it} + \beta_{LF} \ln L_{it} \ln F_{it} \\
+ \frac{1}{2} \beta_{tt} t^2 + \beta_{tt} t \ln K_{it} + \beta_{tt} t \ln L_{it} + \beta_{tt} t \ln F_{it} + \epsilon_{it},
\]

(2)

where \( K \) indicates the capital deployed, \( L \) the number of employees and \( F \) the fuel used. The inefficiency term is defined as \( \epsilon_{it} = v_{it} - u_{it} \), where \( v_{it} \sim N(0, \sigma^2_v) \) is the noise component and \( u_{it} \) is the non-negative stochastic technical inefficiency component. Similar to Kumbhakar and Lovell (2003), in this translog specification as well, time \((t)\) proxies technical change in the stochastic production frontier as well as represents technical efficiency change in the error component. Subsequently, distributional and model specification restriction to econometrically disentangle the two effects is imposed. An attempt is made in the present study to separate the firm-specific unobserved heterogeneity, like Greene (2005b), by introducing the \( \alpha_i \) fixed-effect term. The consequent technical challenges in econometric estimation of such a specification arise broadly due to: the increased computational burden on account of introduction of additional unknown parameters for estimation (one-additional parameter for each firm in the sample in case of fixed-effect model); the problem of incidental parameters (Neyman and Scott, 1948; Lancaster, 2000) contaminates estimates of other model parameters when simultaneous estimation of \( \alpha_i \), the inefficiency parameter, is attempted. The former of the two issues is addressed to some extent by recent developments in computer algorithms (Greene, 2005a). However, in a sample with fixed \( T \) and where \( I \to \infty \), the later problem of incidental parameters results in inconsistent estimates of the variance–covariance matrix (Wang and Ho, 2010). Since, the inefficiency parameters of interest are contained in the variance–covariance matrix, so it is inappropriate to ignore inconsistency of estimates produced by a maximum likelihood estimator (MLE). Wang and Ho (2010) proposed a transformation for the panel stochastic frontier model that allows tractable MLE estimation of the "true fixed-effects" model. The present study adopts this approach to estimate the parameters of the stochastic production function specified in Equation (2). Here MLE tractability is achieved by the use of "scaling property" (Alvarez et al., 2006; Wang and Schmidt, 2002) to represent the inefficiency component \( u_{it} \) as a product of a positive non-stochastic time-varying function and a stochastic time-invariant inefficiency term as given in the following equations:

\[
u_{it} = h_{it} \cdot u_{it}^*,
\]

(3)

\[
 h_{it} = g(z_{hit} \delta_k),
\]

(4)

where \( u_{it}^* \sim N^+(0, \sigma^2_u) \) is assumed to be non-negative half-normal and \( z_{hit} \) represents
\(k = 1, \ldots, K\) exogenous and non-stochastic determinants of inefficiency. In this “composed error” representation \(e_{it} = v_{it} - u_{it}\), the noise component \(v_{it}\) is assumed to be iid and distributed independently of \(u_{it}\). Furthermore, both \(u_{it}\) and \(v_{it}\) are assumed to be independent of \(\{x_{nit}, z_{kid}\}\) for all \(T\) observations of the \(i\)th utility. The scaling property lends several theoretically appealing properties enabling versatile model specifications for empirical work (Alvarez et al., 2006; Wang and Schmidt, 2002; Wang and Ho, 2010). Specifically, the “within” transformation of stochastic frontier models made tractable by the scaling property is adopted in the present study. The “within” transformation removes the individual fixed-effect (incidental parameter) \(\alpha_i\) from the model, and thus, the estimator is immune to the incidental parameters problem. The “within” transformation model adopted in the present paper for MLE estimation is described in greater detail in the Appendix. Furthermore, the time-varying component of the inefficiency term, \(h_{it}\), is specified as given in the following equation:

\[
h_{it} = \exp(z_{kit}\delta_k),
\]

where the influence of following different class of exogenous factors is investigated: vintage, proxied by the year of incorporation of the firm; ownership dummies, identifying a firm to central government, or state government ownership class and private ownership is the reference class; external environmental factors, primarily the time since electricity sector is unbundled in the state in which the firm has majority of operations and extent of competitiveness enabled by institutional conditions; and time trend, all other dis-embodied factors proxied by a quadratic time specification. The inefficiency determinant in Equation (5) is thus specified as given in the following equation:

\[
z_{kit}\delta_k = \text{time} \cdot \delta_t + \text{time}^2 \cdot \delta_{tt} + V \text{int age} \cdot \delta_V + \text{Central Gov. Dummy} \cdot \delta_{CG} + \text{Private Gov. Dummy} \cdot \delta_{PG} + \text{Unbundled Dummy} \cdot \delta_{Udl} + \text{Competition} \cdot \delta_{Comp}.
\]

Estimating productivity changes

Differentiating the production specification with respect to time, following Kumbhakar and Lovell (2003), yields various components of TFP change. The rate of shift in production frontier or technical change is given by the following equation:

\[
\Delta T = \frac{\partial \ln f(x, t; \beta)}{\partial t},
\]

and the change in technical efficiency is obtained by the following equation:

\[
\Delta T = \frac{\partial u}{\partial t}.
\]

The Divisia (1925) index of productivity change is defined for a scalar output as follows:

\[
TFP = \frac{d \ln y}{dt} - \frac{d \ln X}{dt} = \frac{y \cdot \dot{X} - \sum_n (S_n \dot{x}_n)}{y},
\]

where \(S_n = w_n x_n / \sum_m w_n x_m\) is the observed expenditure share of the input \(x_n\). Substituting for \(y\) in Equation (9) obtained by totally differentiating Equation (1) yields with minor algebraic manipulation the following decomposition of productivity change, as given
in the following equation:

\[
\text{TFP} = \Delta T + \Delta TE + (\Gamma - 1) \sum_n \left( \frac{\gamma_n}{\Gamma} \right) x_n + \sum_n \left( \frac{\gamma_n}{\Gamma} - S_n \right) x_n,
\]

(10)

where the elasticity of output with respect of input \( x_n \) is defined as \( \gamma_n = x_n (\delta f/\delta x_n) \). The returns to scale characterizing the production function is then expressed as \( \Gamma = \sum_n (\gamma_n) \). The third term in Equation (10) as given in the following equation:

\[
\Psi = (\Gamma - 1) \sum_n \left( \frac{\gamma_n}{\Gamma} \right) x_n,
\]

(11)

represents the contribution of scale effects due to expansion or contraction of inputs toward total productivity change. Evidently under constant returns to scale (\( \Gamma = 1 \)) there is no contribution of scale effects. However, under increasing/decreasing returns to scale (\( \Gamma > 1/\Gamma < 1 \)) input expansion positively/negatively contributes to productivity change. The allocative efficiency (or price effect) given by the following equation:

\[
\Omega = \sum_n \left( \frac{\gamma_n}{\Gamma} - S_n \right) x_n,
\]

(12)

represents productivity changes that are resulting from factor prices being at deviance from their respective marginal contribution to production. Thus, in case of factor prices reflecting perfect marginal costs (\( (\gamma_n/\Gamma - S_n) = 0 \)), the contribution due to price effect would be nil (\( \Omega = 0 \)). TFP change and its decomposition, derived in Equation (10), can be computed using the parameter estimates of the production function in Equation (2) as follows:

\[
\Delta \hat{T}_n = \hat{\beta}_t + \hat{\beta}_{it} t,
\]

(13)

\[
\Delta \hat{TE}_{it} = -\hat{u}_{it} \frac{dh_{it} dt}{d} \approx -\hat{u}_{it} \frac{(h_{it} - h_{it-1})}{t - t-1},
\]

(14)

\[
\tilde{\gamma}_{nit} = \hat{\beta}_n + \sum_k \hat{\beta}_{nk} \ln x_{it} + \hat{\beta}_{nt} t,
\]

(15)

\[
\hat{\Gamma}_{it} = \sum_n \tilde{\gamma}_{nit}.
\]

(16)

6. Results and discussions

The empirical investigation of the present work is guided by two key objectives: first, to determine the distribution and nature of productivity change in the power sector, and second, the identification of sources of inefficiency. An attempt is made to fulfill these objectives by first, jointly estimating inefficiency and the exogenous determinants of inefficiency, and then the estimated TFP changes are decomposed into constituents of: technical change, efficiency change, scale effects and price effects, to understand the nature of change.

For different classes of utilities in the power sector, the primal production model is estimated using the transformed fixed-effects SFA method. Furthermore, the study employed the MLE technique to fit the model with empirical data. The estimated parameters of the model are shown in Table V and TFP and its decompositions are shown in Table VI.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>T&amp;D</th>
<th>Vertically integrated</th>
<th>Mixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(K)</td>
<td>0.856 (0.32)</td>
<td>1.214 (0.13)</td>
<td>1.311</td>
</tr>
<tr>
<td>ln(L)</td>
<td>0.255 (0.15)</td>
<td>0.326 (0.18)</td>
<td>0.230</td>
</tr>
<tr>
<td>ln(F)</td>
<td>-0.282 (0.16)</td>
<td>-0.192 (0.13)</td>
<td>-0.050</td>
</tr>
<tr>
<td>(1/2) ln(K) ln(K)</td>
<td>0.262 (0.15)</td>
<td>0.326 (0.18)</td>
<td>0.230</td>
</tr>
<tr>
<td>(1/2) ln(L) ln(L)</td>
<td>-0.050 (0.13)</td>
<td>-0.026 (0.18)</td>
<td>-0.013</td>
</tr>
<tr>
<td>(1/2) ln(F) ln(F)</td>
<td>-0.013 (0.13)</td>
<td>-0.026 (0.18)</td>
<td>-0.013</td>
</tr>
<tr>
<td>ln(K) ln(Time)</td>
<td>-0.013 (0.01)</td>
<td>-0.026 (0.18)</td>
<td>-0.013</td>
</tr>
<tr>
<td>ln(L) ln(Time)</td>
<td>-0.026 (0.18)</td>
<td>-0.013 (0.01)</td>
<td>-0.013</td>
</tr>
<tr>
<td>ln(F) ln(Time)</td>
<td>-0.013 (0.13)</td>
<td>-0.026 (0.18)</td>
<td>-0.013</td>
</tr>
<tr>
<td>Time</td>
<td>0.255 (0.15)</td>
<td>0.326 (0.18)</td>
<td>0.230</td>
</tr>
<tr>
<td>Asset Vintage</td>
<td>-0.282 (0.16)</td>
<td>-0.192 (0.13)</td>
<td>-0.050</td>
</tr>
<tr>
<td>Owner: Central Govt.</td>
<td>-0.282 (0.16)</td>
<td>-0.192 (0.13)</td>
<td>-0.050</td>
</tr>
<tr>
<td>Owner: State Govt.</td>
<td>0.081 (0.10)</td>
<td>0.034 (0.09)</td>
<td>0.013</td>
</tr>
<tr>
<td>Unbundled</td>
<td>0.081 (0.10)</td>
<td>0.034 (0.09)</td>
<td>0.013</td>
</tr>
<tr>
<td>Competition</td>
<td>-0.282 (0.16)</td>
<td>-0.192 (0.13)</td>
<td>-0.050</td>
</tr>
</tbody>
</table>

**Notes:** Standard error (in parenthesis) computed using delta method. *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

**Table V.** Maximum likelihood estimates of the translog production model parameters.
Table VI. Power sector: TFP decomposition and changes in TFP

<table>
<thead>
<tr>
<th>Year</th>
<th>TFP</th>
<th>( \Delta T )</th>
<th>( \Delta T E )</th>
<th>( \Psi )</th>
<th>( \Omega )</th>
<th>( \Gamma )</th>
<th>Generator: coal</th>
<th>TFP</th>
<th>( \Delta T )</th>
<th>( \Delta T E )</th>
<th>( \Psi )</th>
<th>( \Omega )</th>
<th>( \Gamma )</th>
<th>Generator: gas</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001–2002</td>
<td>0.212</td>
<td>-0.034</td>
<td>0.158</td>
<td>0.005</td>
<td>-0.008</td>
<td>0.554</td>
<td>0.156</td>
<td>0.148</td>
<td>-0.076</td>
<td>0.092</td>
<td>-0.008</td>
<td>-0.258</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002–2003</td>
<td>0.207</td>
<td>-0.18</td>
<td>0.125</td>
<td>0.006</td>
<td>0.0094</td>
<td>1.714</td>
<td>0.016</td>
<td>0.146</td>
<td>-0.093</td>
<td>0.015</td>
<td>-0.052</td>
<td>0.916</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003–2004</td>
<td>0.068</td>
<td>0.015</td>
<td>0.084</td>
<td>-0.023</td>
<td>-0.007</td>
<td>1.550</td>
<td>-0.017</td>
<td>0.122</td>
<td>-0.076</td>
<td>-0.066</td>
<td>0.003</td>
<td>0.794</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004–2005</td>
<td>0.063</td>
<td>0.039</td>
<td>0.043</td>
<td>-0.008</td>
<td>-0.012</td>
<td>1.781</td>
<td>0.054</td>
<td>0.096</td>
<td>-0.058</td>
<td>0.040</td>
<td>-0.025</td>
<td>0.632</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005–2006</td>
<td>-0.115</td>
<td>0.065</td>
<td>-0.003</td>
<td>-0.151</td>
<td>-0.026</td>
<td>1.831</td>
<td>-0.392</td>
<td>0.073</td>
<td>-0.044</td>
<td>-0.383</td>
<td>0.038</td>
<td>0.812</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006–2007</td>
<td>-0.001</td>
<td>0.100</td>
<td>-0.024</td>
<td>-0.058</td>
<td>-0.019</td>
<td>1.360</td>
<td>0.009</td>
<td>0.036</td>
<td>-0.022</td>
<td>0.103</td>
<td>-0.07</td>
<td>0.546</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007–2008</td>
<td>0.190</td>
<td>0.136</td>
<td>-0.071</td>
<td>0.124</td>
<td>0.001</td>
<td>0.268</td>
<td>0.013</td>
<td>0.004</td>
<td>-0.001</td>
<td>0.03</td>
<td>-0.021</td>
<td>0.354</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008–2009</td>
<td>0.244</td>
<td>0.161</td>
<td>-0.120</td>
<td>0.212</td>
<td>-0.008</td>
<td>0.662</td>
<td>-0.001</td>
<td>-0.014</td>
<td>0.015</td>
<td>0.022</td>
<td>-0.024</td>
<td>0.718</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009–2010</td>
<td>0.091</td>
<td>0.188</td>
<td>-0.158</td>
<td>0.474</td>
<td>-0.112</td>
<td>0.469</td>
<td>-0.112</td>
<td>-0.034</td>
<td>0.045</td>
<td>-0.112</td>
<td>0.011</td>
<td>0.548</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean\textsubscript{2001–2005}</td>
<td>0.138</td>
<td>0.001</td>
<td>0.103</td>
<td>0.018</td>
<td>0.017</td>
<td>1.400</td>
<td>0.052</td>
<td>0.128</td>
<td>-0.076</td>
<td>0.021</td>
<td>-0.21</td>
<td>0.521</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean\textsubscript{2005–2010}</td>
<td>0.082</td>
<td>0.130</td>
<td>-0.075</td>
<td>0.120</td>
<td>-0.093</td>
<td>0.918</td>
<td>-0.080</td>
<td>0.013</td>
<td>-0.001</td>
<td>-0.068</td>
<td>0.024</td>
<td>0.592</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Year</th>
<th>ΔT</th>
<th>ΔTE</th>
<th>Ψ</th>
<th>Ω</th>
<th>Γ</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001–2002</td>
<td>0.167</td>
<td>0.062</td>
<td>0.083</td>
<td>0.034</td>
<td>0.003</td>
</tr>
<tr>
<td>2002–2003</td>
<td>0.292</td>
<td>0.111</td>
<td>0.016</td>
<td>0.045</td>
<td>0.005</td>
</tr>
<tr>
<td>2003–2004</td>
<td>0.517</td>
<td>0.021</td>
<td>0.012</td>
<td>0.036</td>
<td>0.004</td>
</tr>
<tr>
<td>2004–2005</td>
<td>0.538</td>
<td>0.022</td>
<td>0.025</td>
<td>0.039</td>
<td>0.004</td>
</tr>
<tr>
<td>2005–2006</td>
<td>0.232</td>
<td>0.006</td>
<td>0.017</td>
<td>0.032</td>
<td>0.003</td>
</tr>
<tr>
<td>2006–2007</td>
<td>0.421</td>
<td>0.019</td>
<td>0.009</td>
<td>0.021</td>
<td>0.002</td>
</tr>
<tr>
<td>2007–2008</td>
<td>0.432</td>
<td>0.027</td>
<td>0.004</td>
<td>0.011</td>
<td>0.001</td>
</tr>
<tr>
<td>2008–2009</td>
<td>0.067</td>
<td>0.024</td>
<td>0.001</td>
<td>0.004</td>
<td>0.001</td>
</tr>
<tr>
<td>2009–2010</td>
<td>-0.004</td>
<td>0.032</td>
<td>0.010</td>
<td>0.016</td>
<td>0.002</td>
</tr>
<tr>
<td>Mean2001–2005</td>
<td>0.139</td>
<td>0.017</td>
<td>0.009</td>
<td>0.013</td>
<td>0.002</td>
</tr>
<tr>
<td>Mean2005–2010</td>
<td>-0.392</td>
<td>-0.058</td>
<td>0.022</td>
<td>0.007</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Notes: ΔT, technology change; ΔTE, technical efficiency change; Ψ, scale effect; Ω, price effect; Γ, returns to scale; "Mean year = year changes."
For all the estimated models, except “Mixed Generators,” the inefficiency component, \( \ln(\sigma_u) \), is significant and larger that the stochastic noise component \( \ln(\sigma_v) \). Therefore, for these models the data show existence of stochastic inefficiency different from noise. Also, it is observed that \( \sum \beta_n \neq 1 \) and \( \beta_{nk} \neq \beta_{nt} \neq 0, \forall n,k \). Thus, the production technology does not conform to the linearly homogeneous and simpler Cobb–Douglas specification. This justifies our choice of the flexible translog specification and also implies that the scale component, \( \gamma_n \), varies across utilities and through time. Furthermore, it is expected that during this period TFP changed differently for the generators, T&D and integrated utilities.

Generators: coal and gas
For the coal-based generators only ownership is significantly causing efficiency differences. It is estimated that the central government owned generators to be about 57 percent \((=1-e^{-5.63})\) less inefficient than state government owned or private-owned ones. As expected, the assets vintage influences reduction in inefficiency for newer plants, about 0.25 percent \((=1-e^{-0.239/(10 \times (2008-1913)^3)})\) reduction for assets newer by one year; however, the effect is not statistically significant. No influence of competition, un-bundling or time trend in inefficiency is observed. During this period the average per year TFP change observed is 11 percent. The empirical findings observed that after Electricity Act 2003, after 2004, there had been increase in technical change (shift in frontier), 13 percent per year, while efficiency had been declining at \(-7.5\) percent per year. The mean returns to scale, \( \Gamma = 1.15 \), indicates that coal-based power generation shows increasing returns to scale.

For gas-based generators, increased state-level competition is reducing inefficiency such that for every one index point increase in competition there is about 25 percent reduction in inefficiency. Inefficiency also shows a significant time trend. The quadratic terms indicate that inefficiency is increasing till the year 2007 and there is improvement subsequently. No influence of asset vintage, un-bundling or ownership differences on inefficiency is observed. There is an average reduction in TFP of \(-1.4\) percent per year, and the decline is mostly after year 2005 (see figure). After 2005, it is observed that the technical change has reduced from 12.8 percent per year to 1.3 percent, whereas efficiency change has improved from \(-7.6\) to \(-0.1\) percent per year. The mean returns to scale, \( \Gamma = 0.56 \), indicates that gas-based power generation shows decreasing returns to scale.

T&D and integrated utilities
For the T&D utilities, assets vintage has a significant influence as seen by reduction in inefficiency for newer plants. About 1.6 percent \((=1-e^{-1.496/(10 \times (2008-1913)^3)})\) inefficiency reduction for assets newer by one year is observed. Inefficiency also shows a significant time trend. The quadratic terms indicate that inefficiency is increasing at a reducing rate till the year 2008 and there is no decline subsequently (see Figure 1). No influence of competition, un-bundling or ownership differences on inefficiency is observed. TFP changed at a mean rate of 46 percent per year. After 2004, technical change reduced from 13.8 to 8.4 percent per year and efficiency change marginally worsened from \(-7.3\) to \(-8.1\) percent per year. The mean returns to scale, \( \Gamma = 20 \), indicates that T&D firms show increasing returns to scale.

Only ownership is observed to be significantly associated with inefficiency differences for the integrated utilities. The state government owned utilities are observed to be significantly inefficient compared to the private utilities. No influence of competition, un-bundling or time trend in inefficiency is observed. TFP is changing at a mean rate of about \(-11\) percent per year. After 2004, technical change declined from 17.2 to \(-5.6\) percent per year, while efficiency improved from \(-13.2\) to \(-3.3\) percent per year. The mean returns to scale, \( \Gamma = 8 \), indicates that integrated utilities show increasing returns to scale.

Figure 1 shows the technical efficiency trend in India electricity sector and TFP change in the sector is indicated in Figure 2.
7. Conclusion
The empirical findings obtained from the present study suggest that utility-level productivity in the Indian power sector has generally declined during the observed period of 2001–2010. Also, it is estimated that the state-level un-bundling of the electricity sector is not significantly associated with utility-level efficiency change. Furthermore, efficiency improvements attributable to increased competition are observed only in the case of smaller gas-based generating utilities. During the period after Electricity Act 2003, positive technology change is observed while simultaneously a decline in efficiency is observed for the coal-based generators. Improvement in efficiency over time is observed only for
gas-based generators and integrated utilities, whereas T&D licensees show a decline in both technical change and efficiency.

The results obtained in the present empirical investigation are consistent with earlier findings. For instance, Cropper et al. (2011) find no statistically significant improvement in thermal efficiencies after unbundling, while Sen and Jamasb (2010) and Cropper et al. (2011) find a significant improvement in plant-load factors (capacity utilization). A similar effect reflected in the increase in mean scale change effect from 1.8 to 12 percent per year after year 2004 is also observed in the present study. However, contribution to TFP improvement from increased capacity utilization is offset partially by the declining efficiencies.

These results are indicative of the piecemeal approach to power sector reforms in India. The emphasis of reforms in India had been toward unbundling of utilities and opening up the
sector to private independent power producers. However, market for power remains under-developed, tariff reforms are not initiated and fuel remains short in supply. These anomalies are likely to create skewed incentives for firms. The generators are governed by rate of return regulation and generally do not face retail competition. Therefore, de-licensing investment in generation creates incentives for private investors to invest in large capital-intensive projects and is observed in the form of increased technical change only in the coal-based generators that is a result of increased investments in newer and larger plants. From a policy perspective, the results so obtained point toward the need for tariff reforms to encourage increased participation of independent power producers. For the T&D licensees, controlled and low retail prices hardly make up for cost recovery and create disincentive for private investments. For the large generating utilities, there is a lack of market incentives to reduce costs or improve efficiency; therefore, strengthening the electricity markets and introduction of retail competition are possible policy alternatives to pursue.

Note
1. A well-known problem with conventional fixed-effects SFA models with the assumption of time-invariant inefficiency is that it is not possible to distinguish inefficiency from unobserved heterogeneity captured by the fixed-effect term (Schmidt and Sickles, 1984).

References


Further reading


Appendix

SFA model specification
The within transformed SFA model (Wang and Ho, 2010) used in this paper is discussed here. Consider an SFA model with the following general specifications:

\[ y_{it} = x_{it} + x_{i} + \beta + \varepsilon_{it} \quad i = 1, \ldots, I, t = 1, \ldots, T, n = 1, \ldots, N, \]  

(A1)
here \( x_{it} \) is a vector of \( N \) production factor variables (or explanatory variables in general) and \( \alpha_{it} \) represents unobserved fixed effect corresponding to the \( i \)th utility. \( v_{it} \sim N(0, \sigma^2_v) \) is the noise component and \( \pi_{it} \) is the non-negative stochastic technical inefficiency component. While \( \mu \) is set to zero (i.e. \( \mu = 0 \)) and assumes a half-normal distribution for the inefficiency component. The vector \( z_{kit} \) represents \( K \) exogenous variables determining inefficiency.

**Transformed specification**

The within transformation is obtained by subtracting the sample mean of each panel from every individual observation in the panel. The transformation, by de-meaning, removes time-invariant fixed effects from the model. The model specification (Equations (A1)–(A6)) post-transformation may be represented as follows:

\[
\tilde{y}_i = \tilde{x}_i + \tilde{\beta} + \tilde{\alpha}_i, \quad (A7)
\]

\[
\tilde{e}_i = \tilde{v}_i + \tilde{\pi}_i, \quad (A8)
\]

\[
\tilde{\pi}_i \sim N(0, \Pi), \quad (A9)
\]

\[
\tilde{\alpha}_i = \tilde{h}_i + \tilde{\alpha}_i, \quad (A10)
\]

\[
\Pi \sim N^+(\mu, \sigma^2_\pi), \quad (A11)
\]

here, the mean of individual utility over the panel is denoted by \( y_{it} = \frac{1}{T} \sum_{t=1}^{T} y_{it} \), and the mean differenced value by \( \tilde{y}_{it} = y_{it} - y_{it} \). The full panel as a vector stack is represented as \( \tilde{y}_i = (y_{i1}, y_{i2}, \ldots, y_{iT}) \). The variance–covariance matrix of \( \tilde{\alpha}_i \) (Equation (A9)) is given as follows:

\[
\Pi = \begin{bmatrix}
\sigma^2_v (1-1/T) & \sigma^2_v (1-1/T) & \cdots & \sigma^2_v (1-1/T) \\
\sigma^2_v (1-1/T) & \sigma^2_v (1-1/T) & \cdots & \sigma^2_v (1-1/T) \\
\vdots & \vdots & \ddots & \vdots \\
\sigma^2_v (1-1/T) & \sigma^2_v (1-1/T) & \cdots & \sigma^2_v (1-1/T)
\end{bmatrix} . \quad (A12)
\]

**Log–likelihood function**

For the model described above, Wang and Ho (2010) derive the marginal log–likelihood function of the \( i \)th panel as follows:

\[
\ln L_i = -\frac{1}{2} T^{-1} \ln(2\pi) - \frac{1}{2} T^{-1} \ln(\sigma_v^2) - \frac{1}{2} \tilde{\pi}^T \tilde{\Pi}^{-1} \tilde{\pi} + \frac{1}{2} \left( \frac{\mu^2}{\sigma_v^2} \right) + \ln \left( \sigma_1 \Phi \left( \frac{\mu_1}{\sigma_1} \right) \right) - \ln \left( \sigma_1 \Phi \left( \frac{\mu_1}{\sigma_1} \right) \right), \quad (A13)
\]
where $\Pi^*$ is the generalized inverse of $\Pi$, $\phi$ the normal density function, $\Phi$ the cumulative density function, and:

$$
\mu_1 = \frac{\mu_1/\sigma_1^2 - \tilde{h}_{1a} \Pi^* \tilde{h}_{1a}}{\tilde{h}_{1a} I \tilde{h}_{1a} + 1/\sigma_1^2},
$$

(A14)

$$
\sigma_1^2 = \frac{1}{\tilde{h}_{1a} I \tilde{h}_{1a} + 1/\sigma_1^2}
$$

(A15)

$$
\tilde{e}_{1a} = \tilde{y}_{1a} - \tilde{x}_{1a} \beta.
$$

(A16)

The log-likelihood function of the model $L$ is obtained by summing the marginal likelihood over $t = 1, \ldots, I$:

$$
L = \sum_{t=1}^{I} L_t.
$$

(A17)

### Inefficiency and fixed-effect estimation

The inefficiency index of observation/utility, $i$, during period, $t$, can be estimated as the expectation of $u_{it}$ conditional on the model residue, $\tilde{e}_{1a}$:

$$
E(u_{it} | \tilde{e}_{1a}) = h_{it} \left[ \mu_1 + \Phi(\mu_1/\sigma_1)\sigma_1 \right].
$$

(A18)

The fixed effects, $\alpha$s, can be recovered from the estimates of parameters obtained:

$$
\tilde{z}_i = y_{ia} - x_{ia} \hat{\beta} + \tilde{\mu}_2 \hat{h}_{ia} + \tilde{\sigma}_2 \hat{h}_{ia} \Phi(\tilde{\mu}_2/\tilde{\sigma}_2)
$$

(A19)

where:

$$
\tilde{\mu}_2 = \frac{\tilde{\sigma}_2^{-2} - \tilde{\sigma}_2^{-2T} \tilde{\sigma}_u^{-2T} \tilde{h}_{it} \tilde{h}_{it} \tilde{\sigma}_2^{-2T}}{\tilde{\sigma}_u^{-2T} \tilde{h}_{it} \tilde{h}_{it} \tilde{\sigma}_2^{-2T} \tilde{\sigma}_u^{-2T}}
$$

(A20)

$$
\tilde{\sigma}_2^2 = \frac{\tilde{\sigma}_2^{-2T} \tilde{\sigma}_u^{-2T} \tilde{\sigma}_u^{-2}}{\tilde{\sigma}_u^{-2T} \tilde{h}_{it} \tilde{h}_{it} \tilde{\sigma}_2^{-2T} \tilde{\sigma}_u^{-2T}}
$$

(A21)

A routine in the R, statistical language, is written to estimate the maximum likelihood function. Additional routines compute inefficiency indices following Equation (A18) and the firm fixed effects following Equation (A19).

### Corresponding author

Pavan Khetrapal can be contacted at: pavan.jsj@gmail.com

---

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com
The influence of time on employee engagement in the SA business environment

Nico Martins
Department of Industrial and Organisational Psychology, University of South Africa, Pretoria, South Africa, and
Hester Nienaber
Department of Operations Management, University of South Africa, Pretoria, South Africa and
School of Hotel and Restaurant Management, Northern Arizona University (NAU), Flagstaff, Arizona, USA

Abstract

Purpose – The purpose of this paper is twofold: first, to investigate the influence of time on the results of the dimensions of employee engagement; and second, to determine whether there are any significant differences between the levels of engagement of the different demographic groups, so as to determine specific future interventions to improve employee engagement.

Design/methodology/approach – This study adopted a quantitative approach employing a survey which used a questionnaire to collect data from the same convenience sample, over a three-year period. The differences were tested by measuring change through an analysis of variance.

Findings – Three dimensions, namely, team commitment, team orientation and organisational strategy and implementation were significantly higher in the third than first period. Africans and respondents on lower job grades reported significantly lower levels of engagement than white respondents and top management.

Research limitations/implications – The limitation of the study is the low participation rate of some groups.

Practical implications – Top management can foster engagement in addition to introducing effective interventions, based on sound measurement, to improve employees’ engagement levels.

Social implications – Engaged employees are happy/healthy, which can be expected to spill over to their lives outside of the workplace and thus favourably influence society.

Originality/value – Limited longitudinal research in connection with employee engagement is published. This study provides evidence of a valid barometer for a multicultural, developing economy, against which employee engagement can be measured.

Keywords Employee engagement, Longitudinal study, Change over time, Demographic group

1. Introduction

Occupational diseases, especially occupational stress, are on the rise – to such an extent that it is a concern for many organisations (Anthony-McMann et al., 2017; Beehr and Newman, 1978; Brough et al., 2014; Elavainio et al., 2015), as it adversely affects organisational performance. In addition, workplace stress is more likely to spill over to the external environment such as the family rather than the other way round (Giorgi et al., 2015) thus affecting society. It is therefore necessary for leadership and management to reduce occupational stress by creating an environment that nurtures occupational well-being, which will allow employees to perform at their peak (Anitha, 2014; Anthony-McMann et al., 2017; Goetsch, 2010; Gutermann et al., 2017; Mone et al., 2011; Shuck and Rose, 2015). One way of countering workplace stress is to create an environment that fosters “engagement”, for example, by the use of group techniques such as world café methodology or appreciative inquiry (Viljoen, 2015). Engagement is assumed to be a positive and stable indicator of occupational well-being (Al Mehrzi and Kumar Singh, 2016; Anitha, 2014; Anthony-McMann et al., 2017; Gutermann et al., 2017; Kahn, 1990; Saks, 2006;
Schaufeli et al., 2002). It was shown to benefit the organisation and its employees (Anitha, 2014; Bakker et al., 2011; Harter et al., 2013; Mone et al., 2011; Reijseger et al., 2017).

The literature identifies various kinds of engagement (Kahn, 1990; Saks, 2006; Schaufeli et al., 2002), that is, engagement at the individual level, labelled as work engagement (Schaufeli et al., 2002) and engagement at a broader level like the team, department, business unit, organisation, or even the profession, labelled as employee engagement (Macey and Schneider, 2008; Saks, 2006; Schaufeli and Salanova, 2011). Regardless of the kind of engagement, it is a multi-levelled, multi-dimensional construct (Al Mehrzi and Kumar Singh, 2016; Anitha, 2014; Kahn, 1990; Macey and Schneider, 2008; Nienaber and Martins, 2015; Rich et al., 2010; Saks, 2006; Schaufeli and Salanova, 2011; Schaufeli et al., 2002) and not directly observable, which makes it complex and, consequently, difficult to measure (Lockwood, 2007; Mills et al., 2012). In addition, various authors conceptualise engagement differently, regardless of the level of engagement, which results in a variety of measurement instruments (see e.g. Al Mehrzi and Kumar Singh, 2016; Anitha, 2014; Anthony-McMann et al., 2017; Barrick et al., 2015; Kahn, 1990; Macey and Schneider, 2008; Nienaber and Martins, 2015; Rich et al., 2010; Saks, 2006; Schaufeli et al., 2002; Shuck et al., 2016). Moreover, it has been observed that the validity of some measurement instruments is less than optimal (Byrne et al., 2016; Mills et al., 2012; Rothmann and Rothmann, 2010; Viljevac et al., 2012), which is not uncommon for multifaceted, multilevel constructs (Aguinis and Edwards, 2014), like engagement. Instruments for the measuring of engagement reports less than optimal validity may lead to invalid findings, which may limit the usefulness of the measurement (Reio and Shuck, 2015; Steenkamp and Baumgartner, 1998; Strassheim, 2011) resulting in ineffective remedial interventions.

Generally, studies reporting on the measurement of engagement use cross-sectional data (see e.g. Anitha, 2014; Barrick et al., 2015; Klassen et al., 2012; Reijseger et al., 2017; Rich et al., 2010; Saks, 2006; Shuck et al., 2016; Viljevac et al., 2012). Cross-sectional study designs show whether certain variables are associated in ways proposed by the theory, with limited inferences of causality. There are not many studies reporting on the measurement of engagement, irrespective of level, that also report on the construct validity of these measures, as reflected in invariance testing across different demographic variables (see Martins, 2015, 2016; Nienaber and Martins, 2015; Klassen et al., 2012; Seppälä et al., 2009) and over time (Mäkikangas et al., 2016; Seppälä et al., 2009). Of these studies, engagement at the individual level is more likely to report on invariance testing (Klassen et al., 2012; Mäkikangas et al., 2016; Seppälä et al., 2009) than those at a broader level of engagement (Martins, 2015, 2016; Nienaber and Martins, 2015). According to Viljevac et al. (2012), longitudinal studies are required to provide more definitive conclusions about the cause-and-effect relationships, with regard to employee engagement measures. In addition, longitudinal studies can provide information about the temporal order of the relationships underlying employee engagement and show how the presumed outcomes change over time; and whether this change can be attributed to the assumed independent variables (Hasset and Paavilainen-Mäntymäki, 2013; Menard, 2002).

Authors are not unanimous about the definition or nature of longitudinal research (Baltes and Nesselroade, 1979; Hasset and Paavilainen-Mäntymäki, 2013; Menard, 2002; Wall and Williams, 1970). This type of research is generally seen to measure the same variables at least twice in distinct periods for the same (or at least similar) set of participants, while the analysis involves some comparisons across the periods to permit for the measurement of change from one period to the next (Hasset and Paavilainen-Mäntymäki, 2013; Menard, 2002; Taris and Kompier, 2014). Consequently, researchers should take into consideration the theories on the specific relations under investigation, previous empirical studies on these relations and practical considerations in choosing an appropriate longitudinal design for the study to be useful (Baltes and Nesselroade, 1979; Taris and Kompier, 2014). One of the most important considerations in longitudinal studies is the time intervals between the studies (Taris and Kompier, 2014). There is no norm for appropriate
time intervals between studies (Taris and Kompier, 2014). Thus, it is recommended that researchers consider the type of cause and effect studied, as well as the development and context of the process that is being examined (Taris and Kompier, 2014). The first aim of this paper is to investigate the influence of time on the results of the various dimensions of employee engagement measure over a three-year period (2013–2015). Three years were chosen because engagement is considered a stable construct (Al Mehrzi and Kumar Singh, 2016; Anitha, 2014; Anthony-McMann et al., 2017; Gutermann et al., 2017; Kahn, 1990; Saks, 2006; Schaufeli et al., 2002). The next section elaborates on the relevant literature in connection with employee engagement, followed by the method. The results are presented in Section 4, and the discussion of the results is in Section 5. In Section 6, the authors present conclusions, limitations and recommendations for further research.

2. Engagement

The previous section introduced the construct of engagement, while this section provides additional clarifying information. Engagement was first introduced in the literature by Kahn (1990) and from the 2000s gained momentum with an increased number of studies published on the topic (Anthony-McMann et al., 2017; Martins, 2016). Despite this increased attention on engagement, authors are not in agreement in their treatment of engagement (see e.g. Anitha, 2014; Anthony-McMann et al., 2017; Barrick et al., 2015; Nienaber and Martins, 2015; Rich et al., 2010; Shuck et al., 2016). The disagreement on the treatment of engagement results in various definitions and consequently various engagement measurement scales (Anthony-McMann et al., 2017; Nienaber and Martins, 2015). According to Kahn (1990), Saks (2006) and Schaufeli and Salanova (2011), there are various types of engagement, for example, engagement at the individual level, commonly referred to as work engagement; and engagement at a broader level, generally known as employee engagement. Work engagement receives far more attention in the literature than employee engagement (Byrne et al., 2016), despite recommendations to study engagement at a broader level as it may be a better way of studying the link between engagement and organisational performance (Harter et al., 2002; Nienaber and Martins, 2015; Saks, 2006).

Some authors maintain that engagement at the individual and organisational level should not be separated, for at least two reasons. One is that the organisational level of measurement includes aspects of engagement that cannot be captured by merely aggregating individual/unit measures (Pugh and Dietz, 2008). Second, each individual position with its accompanying role in the organisation is specifically designed to accomplish the goals of the organisation, via strategy implementation (Nienaber and Martins, 2015). These goals are impacted by organisational factors influencing employee engagement (see Al Mehrzi and Kumar Singh, 2016; Anitha, 2014; Nienaber and Martins, 2015). For the strategy to be effective, it should be founded on competitive advantage (Grant, 2016; Porter, 1985). In essence, competitive advantage means that an organisation does something better than the competition in attracting customers based on value offered by combining the resources at its disposal (Peteraf and Barney, 2003; Porter, 1985). Of all the resources, employees are the most important (Lockwood, 2007) due to their competence (knowledge, skills, experience, health and well-being, attitudes and behaviours) (Craig and Lopez, 2016; García-Granero et al., 2015), which can change owing to changes in the workplace (Endres and Mancheco-Smoak, 2008; Fawcett et al., 2004; Pfeffer, 2010; Pierson, 2007).

Previous research has shown that engagement is associated with competitive advantage (Cheese et al., 2008; Klassen et al., 2012; Lockwood, 2007; Mills et al., 2012), as well as strategy implementation via innovation and risk-taking (Nienaber, 2017; Reijseger et al., 2017). Innovation and risk-taking are also central to engagement (Reijseger et al., 2017). Hence, it stands to reason that employee engagement is important in organisational performance (Anitha, 2014; Mone et al., 2011; Nienaber, 2017; Reijseger et al., 2017). According to Jarzabkowski and Spee (2009), people “do strategy”. Therefore, when pursuing organisational
goal achievement via strategy implementation, it is important to heed the observation of Anitha (2014) and Dyer (2009), namely, that employees who are engaged at organisational level are familiar with the purpose of their organisation and can accurately communicate its competitive advantage, while caring passionately for its customers in pursuit of organisational goals. This observation links with that of Kahn (1990), Saks (2006) and Cheese et al. (2008) that individual employees choose to engage themselves in varying degrees in response to the organisational environment in which they operate.

As was pointed out earlier, it is the duty of leadership and management to create an environment that fosters engagement and facilitates strategy implementation and thus high organisational performance. An environment conducive to engagement can be accomplished by attending to the psychological conditions of meaningfulness, safety and availability as suggested by Kahn (1990, 1992). Meaningfulness refers to an employee’s sense of how significant it is for them to bring their authentic self to the workplace in view of the “return” they receive to do so. Meaningfulness is affected by task characteristics, role characteristics and work interactions. Safety refers to how safe it is for the employee to bring their authentic self to the workplace without fear of negative consequences to their self-image, status or career. Interpersonal relationships, group dynamics, management style, processes and organisational norms influence safety. Availability refers to how available the employee is to bring their authentic self to the workplace, because of their personal resources at a particular moment. Availability is affected by the physical and emotional energy of employees, feelings of insecurity and their outside lives.

The psychological conditions of meaningfulness, safety and availability correspond to engagement as condition and outcome, as proposed by Shuck and Rose (2015), while these conditions also fit the dimensions of vigour, dedication and absorption as proposed by Schaufeli et al. (2002). Moreover, these variables are also reflected in varying degrees in the engagement conceptualisations of Al Mehrzi and Kumar Singh (2016), Anitha (2014) and Nienaber and Martins (2015), among others.

The engagement instrument of Nienaber and Martins (2015) was developed and validated for the South African multicultural, developing economy context, building on existing theory and taking many factors into consideration related to both construct representation and nomothetic span (Nienaber and Martins, 2015). The researchers investigated the business and organisational psychology databases for articles (in English) pertaining to employee engagement at both the individual and organisational levels. The databases searched were ProQuest, EBSCOhost, Emerald and SABINET as well as sources from the Chartered Institute of Personnel Development were consulted. The search yielded a total of 921 possible articles, of which 53 met the inclusion criteria for the construct “engagement” at the individual, unit and/or organisational levels and/or its measurement and/or validation of engagement measuring instruments. A further 170 textbooks and 76 dissertations/theses were yielded (Martins, 2016, p. 54). As such, the engagement instrument captures the complexity of employee engagement. Construct validity was assessed by examining exploratory factor analysis, and an independent sample using confirmatory factor analysis. The model fit and loading magnitude indicated the adequacy of the proposed factor structure. The engagement questionnaire consists of six factors representing engagement at individual, team and organisational level. The individual level comprises of trait, state and behavioural engagement, which are influenced by factors at team/unit level such as work design, leadership and trust. These, in turn, are influenced by factors at organisational level such as vision, mission, goals and strategy, which are anchored in the competitive advantage of the organisation (Nienaber and Martins, 2015). All six factors making-up employee engagement explain a significant proportion of variation in employee engagement (> 25 per cent according to Heiman, 2014) Organisational strategy and implementation accounts for most (70.9 per cent) of the variation in engagement, while the items associated with innovation and
risk-taking scored the lowest (Nienaber, 2017), and this factor was scored the most unfavourable by respondents (see Table II).

The Nienaber and Martins (2015) instrument may be considered solid as each of the six factors consists of at least five or more items (Hinkin, 1998), with factor loadings of at least 0.40 (Costello and Osborne, 2005; Henson and Roberts, 2006). The validity, discriminant validity, convergent validity and reliability of the instrument have been established (Martins, 2015, 2016; Nienaber and Martins, 2015). Cronbach’s α for the six dimensions ranged between 0.813 and 0.942. A recent study indicated that the fit indices of the base model (GFI: 0.809; RMSEA: 0.048; NFI: 0.860; IFI: 0.947; TLI: 0.942; CFI: 0.947) were all at levels recommended by the relevant scholars in the field (Gallant, 2017). The six dimensions are discussed in the section Measuring instruments.

The instrument consisted of 69 questions, 9 of which collected biographical information (gender, qualifications, job grade and tenure) and 60 required responses to statements about engagement at individual, team/unit and organisational level, using a five-point Likert-type scale, widely used to measure opinion, belief and attitude (DeVellis, 2013), such as engagement.

According to the IMD (2017), South Africa’s ranking alternated between 53rd and 52nd out of 63 countries, in 2013 to 2017, reflecting a weak competitive position, indicating a potentially unfavourable situation for employee engagement. In examining South Africa’s competitive profile, it transpired that the “human factor” was a major contributor to this “sorry state” of ranking. Of specific concern were factors such as low employment levels, resulting from a lack of obligation to skills development, as well as low productivity and efficiency levels. Unfavourable institutional frameworks, business legislation, the societal framework, the labour market, attitudes and values and infrastructure, including health issues (IMD, 2017), compound these factors. Moreover, the war for talent and technological change may also adversely affect employee engagement levels. Not surprising employee engagement in South Africa measured relatively low (Nienaber and Martins, 2015), which is no different from the global situation (Aon Hewitt, 2017). The factors reflecting a weak competitive position as highlighted in the IMD (2017) are consistent with the findings of studies investigating strategy implementation in a South African context, namely, that leadership down the line is important in strategy implementation (Jooste and Fourie, 2009; Poisat, 2014). These factors further reflect that the unavailability of resources, especially human resources (Tait and Nienaber, 2010; Van der Merwe and Nienaber, 2015), contributes to implementation failure, adversely impacting organisational performance. These findings have implications for employee engagement, and correspond with those of international studies (Anitha, 2014; Barrick et al., 2015; Gutermann et al., 2017). This leads to the second aim, namely, to determine whether there are any significant differences between the levels of engagement in the various demographic groups during the last year of study, to determine specific interventions that may be recommended to enhance the levels of employee engagement.

3. Research method
The data from three employee engagement studies were used to fulfil the research objectives.

3.1 Objectives of the study
The first research question aimed to determine whether there was a significant difference between the results of the engagement dimensions over a three-year period (2013–2015). The aim was to determine if the engagement levels of the participants changed over the three-year period in correspondence with South Africa’s GDP, which declined from 2 per cent in 2013, to 1.6 per cent in 2014 and to 1.3 per cent in 2015 (IMD, 2017). Essentially, the outcome of this can indicate whether employee engagement is a stable construct, as purported in the literature.
The second research question of the study aimed to determine whether there were any significant differences in the engagement levels between the various demographic groups during the last year of study. This could help to propose future interventions, based on the most current results, to improve employee engagement.

3.2 Research participants and research approach
The participants in all three studies (2013–2015) were convenience samples from a database consisting of 285,000 business people from various industries. The samples were drawn from government institutions and different businesses that provide jobs that reflect the profile of the South African working population. The database was permissioned – that is, everybody on the database participated in the online surveys after giving their permission first. Electronic surveys administered by the iFeedback.co.za online data collection portal were initiated by means of a mass e-mail invitation. Each potential participant received a personalised e-mail stating the purpose of the investigation. The participants were also informed that the survey would take approximately 15 min to complete. They were invited to participate in the survey on a voluntary, confidential and anonymous basis (Martins, 2015). Due to the confidentiality clause of the surveys it was not possible to invite the same participants to participate in the three surveys. Although the same database was used to collect data, it was not possible to perform sample attrition analyses. Hence, it is possible that the samples obtained in the three periods were not identical, as there is no guarantee that the same participants responded to the survey in the three periods. However, participants are considered similar, and thus acceptable for longitudinal comparisons (Hasset and Paavilainen-Mäntymäki, 2013; Menard, 2002; Taris and Kompier, 2014). This research is no different from other longitudinal studies which report differences in respondents (see Kuypers et al., 2018). Our approach is in line with the recommendations of Hasset and Paavilainen-Mäntymäki (2013), Hirschfeld et al. (2013), Maloney et al. (2010), Menard (2002), Kuypers et al. (2018) and Taris and Kompier (2014), to increase the chance of detecting relationships actually existing.

3.3 Statistical analysis
The first research question aimed to determine whether there was a significant difference between the results of the engagement dimensions over a three-year period. This was tested by means of the one-way analysis of variance (ANOVA) with the post hoc Scheffe test. According to Caruana et al. (2015), univariate (ANOVA) and multivariate (MANOVA) analysis of variance is often adopted for longitudinal analysis. In both cases, the assumption of equal interval lengths and normal distribution in all groups and that only means are compared, sacrificing individual-specific data. To ensure that the data fit the one-way ANOVA model a number of assumptions were tested. The researchers ensured that the first assumption of one dependent variable was met with two or more categorical, independent groups. The second assumption, the researchers adhered to, was independence of observations. The researchers also ensured that no significant outliers in the groups of the independent variable in terms of the dependent variable were included. Finally, the researchers investigated if the dependent variable was normally distributed, for each independent variable. The limitation of using a sample which could have changed due to attrition of participants and that not the same participants participated in all three surveys was acknowledged by the authors. On the other hand, the use of a multi-sample design, longitudinal data and sophisticated statistical methods can be considered the strengths of this study (Seppälä et al., 2009). It should also be noted that this study was the first to examine the effect of time on employee engagement in a South African context. Differences are reported at the customary $\alpha$ coefficient of 0.05 to determine if there are any differences between the three years of study. The second phase of analyses involved the various demographic groups. The one-way ANOVA was also utilised to determine if there are any significant differences between the various demographic groups.
groups. Statistical analysis was carried out by means of the Statistical Package for Social Sciences (version 23).

3.4 Measuring instrument


The data were gathered by means of the same employee engagement questionnaire (Martins, 2016), from the same convenience sample. The respondents had to rate the items on a five-point Likert scale: 1 = strongly disagree; 2 = disagree; 3 = unsure; 4 = agree; 5 = strongly agree. The scale contains 45 Likert-scaled items, phrased in English, the business language of South Africa, and is deemed to contain conceptually clear items at all levels (demographic level, organisational level, team level and individual level). The scale is administered as a self-report, which is the most appropriate method when subjective perceptions and experiences of individuals are required (Conway and Lance, 2010).

The questionnaire covered the following employee engagement constructs:

- **Team orientation**: this dimension refers to the team members’ willingness to render support; to be well-organised; to take personal ownership of their job responsibilities; and represents engagement at individual level.

- **Organisational satisfaction**: this dimension includes aspects such as a meaningful job; an enjoyable, inspiring job; a general feeling of job satisfaction; and represents engagement at individual level.

- **Effectiveness of managerial tasks**: this dimension refers to the effectiveness of managerial tasks such as trust, support, feedback and performance evaluation; and represents engagement at team/unit level. The dimension name was adapted from immediate manager to reflect the meaning more clearly.

- **Organisational commitment**: this dimension involves commitment to the organisation: to take pride in the work that is being done for the organisation; congruence between personal and organisational values; positivity about the organisation’s future; and represents engagement at the organisational level.

- **Organisational strategy and implementation**: the aim of this dimension is to determine if employees are encouraged to develop new ideas, and if initiative is encouraged. It also aims to determine if employees are involved in implementing strategy; and represents engagement at organisational level.

- **Team commitment**: this dimension refers to cooperation; collective solving of problems; the ability to adapt to change; doing more than is expected; and represents engagement at team/unit level.
The validity, discriminant validity, convergent validity and reliability of the instrument were established in previous studies (Martins, 2016; Nienaber and Martins, 2015). The reliability results of the extracted constructs demonstrate strong internal consistency, well above 0.70, as illustrated by Cronbach’s $\alpha$ coefficients. Cronbach’s $\alpha$ coefficients ranged between 0.816 (strategy and implementation) and 0.947 (organisational commitment). This is typically what is expected of an established questionnaire. According to Nunnally (1978), Cronbach’s $\alpha$ scores for an established scale should be at least above 0.70.

4. Results
4.1 Descriptive statistics
The results were reported on an aggregated level by dimension and biographical groups only, that is, no results could be tied down to any individual (Martins, 2015). The demographic profile of the participants is reflected in Table I.

It is important to note that the participation percentage per demographic group was, generally, very consistent across the three years of study. An important observation is the large percentage of management cadre participants in the three surveys. These results

<table>
<thead>
<tr>
<th>Item</th>
<th>2013 (1,073)</th>
<th>2014 (4,125)</th>
<th>2015 (4,099)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>542 (50.5%)</td>
<td>1,994 (48.3%)</td>
<td>2,387 (58.2%)</td>
</tr>
<tr>
<td>Female</td>
<td>527 (49.1%)</td>
<td>2,131 (51.7%)</td>
<td>1,712 (41.8%)</td>
</tr>
<tr>
<td>No response</td>
<td>4 (0.4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of service</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–1</td>
<td>54 (5.0%)</td>
<td>281 (6.8%)</td>
<td>166 (4.0%)</td>
</tr>
<tr>
<td>2–3</td>
<td>171 (15.9%)</td>
<td>650 (15.8%)</td>
<td>456 (11.1%)</td>
</tr>
<tr>
<td>4–5</td>
<td>149 (13.9%)</td>
<td>510 (12.4%)</td>
<td>429 (10.5%)</td>
</tr>
<tr>
<td>6–10</td>
<td>273 (25.4%)</td>
<td>1,113 (27.0%)</td>
<td>1,006 (24.5%)</td>
</tr>
<tr>
<td>10 and longer</td>
<td>380 (35.4%)</td>
<td>1,571 (38.1%)</td>
<td>2,042 (49.8%)</td>
</tr>
<tr>
<td>No response</td>
<td>46 (4.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qualification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std 6 (Grade 8) and below</td>
<td>1 (0.1%)</td>
<td>1 (0.0%)</td>
<td>7 (0.2%)</td>
</tr>
<tr>
<td>Std 7 and 8 (Grades 9 and 10)</td>
<td>13 (1.2%)</td>
<td>40 (1.0%)</td>
<td>27 (0.7%)</td>
</tr>
<tr>
<td>Std 9 and 10 (Grades 11 and 12)</td>
<td>173 (16.1%)</td>
<td>565 (13.7%)</td>
<td>466 (11.4%)</td>
</tr>
<tr>
<td>Certificate</td>
<td>173 (16.1%)</td>
<td>615 (14.9%)</td>
<td>495 (12.1%)</td>
</tr>
<tr>
<td>Diploma</td>
<td>283 (26.4%)</td>
<td>1,014 (24.6%)</td>
<td>952 (23.2%)</td>
</tr>
<tr>
<td>First degree</td>
<td>163 (15.2%)</td>
<td>666 (16.1%)</td>
<td>745 (18.2%)</td>
</tr>
<tr>
<td>Post-graduate qualification</td>
<td>265 (24.7%)</td>
<td>1,218 (29.5%)</td>
<td>1,404 (34.3%)</td>
</tr>
<tr>
<td>No response</td>
<td>2 (0.2%)</td>
<td>6 (0.1%)</td>
<td>3 (0.1%)</td>
</tr>
<tr>
<td>Race group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>239 (22.3%)</td>
<td>934 (22.6%)</td>
<td>778 (19.0%)</td>
</tr>
<tr>
<td>Coloured</td>
<td>125 (11.6%)</td>
<td>413 (10.0%)</td>
<td>329 (8.0%)</td>
</tr>
<tr>
<td>Indian</td>
<td>112 (10.4%)</td>
<td>345 (8.4%)</td>
<td>338 (8.2%)</td>
</tr>
<tr>
<td>White</td>
<td>570 (53.1%)</td>
<td>2,332 (56.5%)</td>
<td>2,543 (62.0%)</td>
</tr>
<tr>
<td>Other</td>
<td>4 (0.4%)</td>
<td>19 (0.5%)</td>
<td>19 (0.5%)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>20 (1.9%)</td>
<td>77 (1.9%)</td>
<td>82 (2.0%)</td>
</tr>
<tr>
<td>No response</td>
<td>3 (0.3%)</td>
<td>5 (0.1%)</td>
<td>10 (0.2%)</td>
</tr>
<tr>
<td>Generation group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1946 and 1964</td>
<td>–</td>
<td>1,285 (31.2%)</td>
<td>1,458 (35.6%)</td>
</tr>
<tr>
<td>Born between 1965 and 1977</td>
<td>–</td>
<td>1,690 (41.0%)</td>
<td>1,730 (42.2%)</td>
</tr>
<tr>
<td>Born between 1978 and 2000</td>
<td>–</td>
<td>1,150 (27.9%)</td>
<td>911 (22.2%)</td>
</tr>
<tr>
<td>Job grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top management</td>
<td>60 (5.6%)</td>
<td>526 (12.8%)</td>
<td>734 (17.9%)</td>
</tr>
<tr>
<td>Senior management</td>
<td>231 (21.5%)</td>
<td>855 (20.7%)</td>
<td>1,149 (28.0%)</td>
</tr>
<tr>
<td>Manager</td>
<td>304 (28.3%)</td>
<td>1,041 (25.2%)</td>
<td>1,175 (28.7%)</td>
</tr>
<tr>
<td>Supervisor</td>
<td>141 (13.1%)</td>
<td>442 (10.7%)</td>
<td>349 (8.5%)</td>
</tr>
<tr>
<td>Employee</td>
<td>335 (31.2%)</td>
<td>1,255 (30.4%)</td>
<td>688 (16.8%)</td>
</tr>
<tr>
<td>No response</td>
<td>2 (0.2%)</td>
<td>6 (0.1%)</td>
<td>4 (0.1%)</td>
</tr>
</tbody>
</table>

*Source*: Authors’ compilation based on survey results

Table I.
Demographic results of the three years of study
indicate that many respondents occupy executive-, senior- or management-level positions at their respective organisations. The split between male and female was almost equal, as reflected in the general population. Notably, in 2015, the majority (77.8 per cent) of the sample was born before 1977, making them 38 years and older. A large percentage reported a post-graduate qualification. This implies that the sample consisted of a large number of well-educated and older individuals who hold senior positions at their companies. The biographic information supports the observations of Fulton (2016), with the exception of the high response rate from (senior) managers, who usually do not respond to surveys owing to time pressures. Fulton (2016) argued that a high response rate from senior executives is desirable because they know their organisations best and thus are the best source of information. This can have a cascading effect on the employees at other levels of the organisation. The low response rate from African participants may be attributed to cultural differences (Lyness and Kropf, 2007).

As was mentioned before, to ensure that the data fit the one-way ANOVA model, the following assumption is tested to determine normality. An inspection of the data indicates that the data are skewed left, indicating rather higher than lower scores (Figure 1). The box outliers are calculated by means of the box plot formula, for extreme outliers (outliers more than three box lengths of the median). The results indicated only 14 extreme outliers with counts of more than 1.16. The authors are in agreement with researchers such as Ghasemi and Zahediasl that these outliers will not have any impact on a sample size of +9,000 (total sample). In summary, the violation of assumption of normality is not a problem for this sample. The authors thus continued with the ANOVA model testing.

4.2 Inferential statistics
The first aim of the study was to investigate the possible impact of time on the engagement levels of the respondents, focussing on the measured engagement dimensions. Table II portrays the results of this analysis.
It is noteworthy that only three dimensions, namely, team commitment, organisational strategy and implementation and team orientation were significantly more positive in 2015 than in 2013. The overall means of all the other dimensions also indicated higher engagement levels, but not significantly higher. The highest levels of engagement were noted for the dimensions of team commitment (engagement at organisational level) and team orientation (engagement at individual level).

The results of the various demographic groups were subsequently investigated with the Scheffe post hoc test, to determine which of the groups in the last study (2015) were significantly more engaged, and for which dimensions. These results are depicted in Tables III–VIII. Only the results of the demographic groups with significant differences ($p < 0.05$) are portrayed and discussed. Given the large number of management members who participated in the surveys, the overall results pertaining to the job grades were analysed first (Table III).

These results indicate that top management were the most engaged in their organisations in all three years of study. A more detailed analysis by the engagement dimension indicates similar tendencies (Table IV).

The results of the top management job grade are significantly more engaged than those of the other job grades for all dimensions, except for the dimension of team commitment and team orientation in the case of senior management.

Table V portrays the results for the generation groups.

---

### Table II. Overall comparative results of the three years of study

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>2013 Respondents</th>
<th>2014 Respondents</th>
<th>2015 Respondents</th>
<th>Overall mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organisational strategy and implementation</td>
<td>1,069</td>
<td>3.2679</td>
<td>0.86955</td>
<td>6.026</td>
</tr>
<tr>
<td>Team commitment</td>
<td>1,072</td>
<td>3.9824</td>
<td>0.68970</td>
<td>3.260</td>
</tr>
<tr>
<td>Organisational satisfaction</td>
<td>1,073</td>
<td>3.711</td>
<td>0.81039</td>
<td>2.881</td>
</tr>
<tr>
<td>Organisational commitment</td>
<td>1,073</td>
<td>3.5843</td>
<td>0.90638</td>
<td>1.983</td>
</tr>
<tr>
<td>Effectiveness of managerial tasks</td>
<td>1,072</td>
<td>3.4826</td>
<td>0.97419</td>
<td>1.499</td>
</tr>
<tr>
<td>Team orientation</td>
<td>1,069</td>
<td>4.1066</td>
<td>0.76513</td>
<td>3.195</td>
</tr>
</tbody>
</table>

Notes: *$p < 0.05$; **$p < 0.01$
These results indicate that those born between 1978 and 2000 (Generation Y) are in all instances the most engaged generation and significantly more engaged than the other generations.

Table VI portrays the results of the years of service groups, which first indicate that the zero to one years of service groups are the most engaged group in the dimension organisation strategy and implementation, and effectiveness of managerial tasks. Second, the group with experience of 10 years and longer is significantly more engaged in the dimension organisational satisfaction.

Only one dimension was significantly different between the race groups, namely, organisational strategy and implementation (Table VI). The white participants are the most engaged group and significantly more so than the African participants.

The results of the qualification groups (Table VIII) showed significant differences in two dimensions.

It is important to note that the employees with higher qualifications displayed higher levels of engagement and were also in two instances significantly more engaged than those

<table>
<thead>
<tr>
<th>Job level</th>
<th>Mean</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organisational strategy and implementation</strong></td>
<td></td>
<td>77.896</td>
<td>0.000**</td>
</tr>
<tr>
<td>Top management</td>
<td>3.83++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior management</td>
<td>3.41+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>3.27+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisor</td>
<td>3.08+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee</td>
<td>3.12+</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Team commitment</strong></td>
<td></td>
<td>26.277</td>
<td>0.000**</td>
</tr>
<tr>
<td>Top management</td>
<td>4.08++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior management</td>
<td>4.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>4.00+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisor</td>
<td>3.94+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee</td>
<td>3.88+</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Organisational satisfaction</strong></td>
<td></td>
<td>99.153</td>
<td>0.000**</td>
</tr>
<tr>
<td>Top management</td>
<td>4.18++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior management</td>
<td>3.86+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>3.70+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisor</td>
<td>3.54+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee</td>
<td>3.40+</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Effectiveness of managerial tasks</strong></td>
<td></td>
<td>34.438</td>
<td>0.000**</td>
</tr>
<tr>
<td>Top management</td>
<td>3.88++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior management</td>
<td>3.58+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>3.45+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisor</td>
<td>3.34+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee</td>
<td>3.58+</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Organisational commitment</strong></td>
<td></td>
<td>83.312</td>
<td>0.000**</td>
</tr>
<tr>
<td>Top management</td>
<td>4.14++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior management</td>
<td>3.69+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>3.56+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisor</td>
<td>3.37+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee</td>
<td>3.34+</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Team orientation</strong></td>
<td></td>
<td>24.808</td>
<td>0.000**</td>
</tr>
<tr>
<td>Top management</td>
<td>4.32++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior management</td>
<td>4.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>4.14+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervisor</td>
<td>4.02+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee</td>
<td>4.00+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table IV.
Overall comparative results of job grades by dimension for 2015

Notes: ++ Post hoc test significantly more positive than +. *p > 0.05; **p > 0.01
Source: Calculated from survey results
Influence of time on employee engagement

<table>
<thead>
<tr>
<th>Generation</th>
<th>Mean</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organisational strategy and implementation</td>
<td>15.478</td>
<td>0.000**</td>
<td></td>
</tr>
<tr>
<td>Born between 1978 and 2000</td>
<td>3.48++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1965 and 1977</td>
<td>3.30+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1946 and 1964</td>
<td>3.33+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Team commitment</td>
<td>10.704</td>
<td>0.000**</td>
<td></td>
</tr>
<tr>
<td>Born between 1978 and 2000</td>
<td>4.12++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1965 and 1977</td>
<td>4.02+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1946 and 1964</td>
<td>3.99+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisational satisfaction</td>
<td>37.812</td>
<td>0.000**</td>
<td></td>
</tr>
<tr>
<td>Born between 1978 and 2000</td>
<td>3.92++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1965 and 1977</td>
<td>3.72+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1946 and 1964</td>
<td>3.64+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness of managerial tasks</td>
<td>5.759</td>
<td>0.003**</td>
<td></td>
</tr>
<tr>
<td>Born between 1978 and 2000</td>
<td>3.60++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1965 and 1977</td>
<td>3.48+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1946 and 1964</td>
<td>3.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisational commitment</td>
<td>23.281</td>
<td>0.000**</td>
<td></td>
</tr>
<tr>
<td>Born between 1978 and 2000</td>
<td>3.78++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1965 and 1977</td>
<td>3.59+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1946 and 1964</td>
<td>3.54+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Team orientation</td>
<td>5.649</td>
<td>0.004**</td>
<td></td>
</tr>
<tr>
<td>Born between 1978 and 2000</td>
<td>4.22++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1965 and 1977</td>
<td>4.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born between 1946 and 1964</td>
<td>4.11+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: ++ Post hoc test significantly more positive than +. *p > 0.05; **p > 0.01
Source: Calculated from survey results

<table>
<thead>
<tr>
<th>Years of service</th>
<th>Mean</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy and implementation</td>
<td>5.030</td>
<td>0.000**</td>
<td></td>
</tr>
<tr>
<td>0–1 year</td>
<td>3.58++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2–3 years</td>
<td>3.32+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4–5 years</td>
<td>3.31+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–10 years</td>
<td>3.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 years and longer</td>
<td>3.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisational satisfaction</td>
<td>4.291</td>
<td>0.002**</td>
<td></td>
</tr>
<tr>
<td>0–1 year</td>
<td>3.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2–3 years</td>
<td>3.66+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4–5 years</td>
<td>3.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–10 years</td>
<td>3.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 years and longer</td>
<td>3.81++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness of managerial tasks</td>
<td>3.969</td>
<td>0.003**</td>
<td></td>
</tr>
<tr>
<td>0–1 year</td>
<td>3.82++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2–3 years</td>
<td>3.52+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4–5 years</td>
<td>3.49+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–10 years</td>
<td>3.51+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 years and longer</td>
<td>3.55+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: ++ Post hoc test significantly more positive than +. *p > 0.05; **p > 0.01
Source: Compiled from statistical analysis

Table V. Results of generation groups by dimension for 2015

Table VI. Results of years of service groups by dimension for 2015
with certificates. The group with post-graduate qualifications was significantly more engaged in the dimension organisational strategy and implementation, while the group with a diploma was significantly more engaged in the dimension organisational satisfaction.

The results of the gender groups are displayed below in Table IX.

Table VII.
Results of race groups by dimension for 2015

<table>
<thead>
<tr>
<th>Race group</th>
<th>Mean</th>
<th>$F$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4.549</td>
<td>0.000*</td>
</tr>
<tr>
<td>Organisational strategy and implementation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>3.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coloured</td>
<td>3.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indian</td>
<td>3.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>3.42++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>3.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: ++ Post hoc test significantly more positive than +; *$p < 0.05$; **$p < 0.01$

Source: Compiled from statistical analysis

Table VIII.
Results of years of qualifications by dimension for 2015

<table>
<thead>
<tr>
<th>Qualifications</th>
<th>Mean</th>
<th>$F$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4.193</td>
<td>0.002**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisational strategy and implementation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std 9 and 10 (Grades 11 and 12)</td>
<td>3.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Certificate</td>
<td>3.27+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diploma</td>
<td>3.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First degree</td>
<td>3.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-graduate qualification</td>
<td>3.44++</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.716</td>
<td>0.005**</td>
</tr>
<tr>
<td>Organisational satisfaction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std 9 and 10 (Grades 11 and 12)</td>
<td>3.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Certificate</td>
<td>3.66+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diploma</td>
<td>3.82++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First degree</td>
<td>3.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-graduate qualification</td>
<td>3.79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: *$p < 0.05$; **$p < 0.01$

Source: Compiled from statistical analysis

Table IX.
Results of gender groups by dimension for 2015

<table>
<thead>
<tr>
<th>Gender group</th>
<th>Mean</th>
<th>$F$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>23.957</td>
<td>0.000**</td>
</tr>
<tr>
<td>Organisational satisfaction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>3.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness of managerial tasks</td>
<td></td>
<td>4.760</td>
<td>0.029*</td>
</tr>
<tr>
<td>Male</td>
<td>3.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisational commitment</td>
<td></td>
<td>5.959</td>
<td>0.015*</td>
</tr>
<tr>
<td>Male</td>
<td>3.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>3.60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: *$p < 0.05$; **$p < 0.01$

Source: Calculated from survey results
Table IX indicates that significant differences between the gender groups were displayed in three dimensions, namely, organisational satisfaction, effectiveness of managerial tasks and organisational commitment. In all three instances, the males were significantly more engaged.

5. Discussion
The results of the current study make several contributions to the current body of knowledge. With reference to the first research question about the influence of time (three years) on the results of the levels of employee engagement dimensions, the results indicated that engagement significantly improved in three dimensions, namely, team commitment, organisational strategy and implementation and team orientation – in 2015 than in 2013 (see Table II). Although the overall means of all other dimensions of engagement also indicated higher scores, however, they were not significantly higher. The results of this study thus suggest that engagement is, generally, relatively stable, as suggested by the literature (Kahn, 1990; Saks, 2006; Schaufeli et al., 2002) and as measured by this instrument drawing on the definition of Kahn (1990). The highest levels of engagement were reported for the dimensions of team orientation and team commitment (see Table II). Despite the significant improvement in engagement in the dimension organisational strategy and implementation (engagement at the organisational level) over the time, the mean is still the lowest of all the dimensions making-up employee engagement, while effectiveness of managerial tasks (engagement at team/unit level) scored the second lowest. If these dimensions are not improved, they may adversely affect employee well-being as warned in the literature (Anthony-McMann et al., 2017; Beehr and Newman, 1978; Brough et al., 2014; Elovainio et al., 2015; Giorgi et al., 2015), leading to employees not performing at their peak (Anthony-McMann et al., 2017; Goetsch, 2010; Gutermann et al., 2017; Mone et al., 2011) and to the detriment of organisational performance.

Given this observation, it would seem that leadership and management participating in this survey are not entirely fulfilling their duty to create an environment that nurtures engagement (Al Mehrzi and Kumar Singh, 2016; Anitha, 2014; Gutermann et al., 2017). This is to the detriment of competitiveness (IMD, 2017) and society at large. In spite of the declining GDP, which declined from 2 per cent in 2013, to 1.6 per cent in 2014 and to 1.3 per cent in 2015, the engagement levels of the participants were very stable over this period and as indicated, some engagement dimensions improved significantly. In view of the unfavourable economic growth, which prevailed at the time of the surveys, it is surprising that employee engagement improved, as the negative economic conditions are considered to bear unfavourably on engagement. However, the (slight) improvement in employee engagement in this period can be explained by a range of reasons. First, the same convenience sample was utilised for the study and participants/respondents’ exposure to the construct might have impacted testing effects, as they could have been more sensitive to the construct. Second, in view of the competitive pressures, leadership and management might have intensified their competitive efforts to ensure survival, impacting on participants and their roles in organisations. Finally, the participants’ attitudes might have been more optimistic, since in times of uncertainty employees are more inclined to stay in their current jobs and to contribute in their work environment.

The second research question of the study aimed to determine whether there were any significant differences between the engagement levels of the various demographic groups during the last year of study. The demographic variable job grade showed that top management was significantly more engaged, in all six dimensions, in comparison to other job levels (see Table IV). In the case of generation, the group born between 1978 and 2000 (Generation Y) was significantly more engaged than the other generations in all six dimensions of employee engagement (Table V). In terms of years of service (Table VI) the group with less than one years’ service was significantly more engaged in the dimensions...
organisational strategy and implementation and effectiveness of managerial tasks, than the other service groups. The group with 10 years and more service was significantly more engaged in the dimension organisational strategy and implementation, than the other service groups. In the race group, whites were significantly more engaged in organisational strategy and implementation than the other groups (Table VI). Regarding qualification, the group with post-graduate qualifications was significantly more engaged in the dimension organisational strategy and implementation, than the other groups, while the group with diplomas was significantly more engaged in the dimension organisational satisfaction, than the other groups (Table VIII).

Given these observations, in an effort to improve employee engagement levels, leadership and management should attend to the dimension strategy and implementation because it was scored the lowest of all dimensions, while it explains most of the variance (70.9 per cent) in employee engagement; and it has the potential to improve significantly as demonstrated by the results of this study. Moreover, engagement is associated with competitive advantage which is the foundation of a sound strategy (Craig and Lopez, 2016; Cheese et al., 2008; García-Granero et al., 2015; Klassen et al., 2012; Lockwood, 2007; Mills et al., 2012). Engagement is also associated with innovation and risk-taking, which drive strategy implementation (Nienaber, 2017; Reijseger et al., 2017). It is also important in organisational performance (Anitha, 2014; Barney, 1991; Grant, 2016; Peteraf and Barney, 2003; Mone et al., 2011; Nienaber, 2017; Porter, 1985; Reijseger et al., 2017).

Employees execute strategy (Jarzabkowski and Spee, 2009) and if they are not engaged, it can be expected that there will be a performance gap (i.e. goal achievement will fall short of planned performance). Leadership and management play an important role in strategy implementation (Jooste and Fourie, 2009; Poisat, 2014) by creating an environment that nurtures engagement (Kahn, 1990; Saks, 2006; Schaufeli et al., 2002; Schuck and Rose, 2015). Nurturing of engagement cannot be accomplished without utilising employees competence (Craig and Lopez, 2016; Endres and Manchego-Smoak, 2008; Fawcett et al., 2004; García-Granero et al., 2015; Pfeffer, 2010; Piersol, 2007; Tait and Nienaber, 2010; Van der Merwe and Nienaber, 2015).

In addition, this study supports the notion that employee engagement is a multi-dimensional, multi-levelled construct (Al Mehrzi and Kumar Singh, 2016; Anitha, 2014; Kahn, 1990; Macey and Schneider, 2008; Nienaber and Martins, 2015; Rich et al., 2010; Saks, 2006; Schaufeli and Salanova, 2011; Schaufeli et al., 2002).

6. Conclusion and implications
From the results, it is clear that the engagement levels of the participants in the study improved during the three years (2013–2015). One of the reasons for this improvement might be a greater commitment of employees to their current jobs owing to the prevailing unfavourable economic environment, which made it difficult for them to find alternative employment. Other reasons might be the longitudinal nature of the study creating a sensitivity to the construct; the profile of the respondents (mostly managerial participants and persons with a post-graduate qualification) (see Fulton, 2016); and length of service (those with less than a year service being more engaged). The results show that these groups are engaged. However, the results suggest that the management groups may not cascade their engagement to the lower job levels, as expected. There might be various reasons for this, such as insufficient quality leadership (Poisat, 2014) and insufficient involvement of incumbents in lower hierarchical positions in strategy shaping, which is compounded by the inadequate compensation for these employees (Anstey, 2013).

7. Recommendations
In addition, the 2015 study yielded some interesting results, which gave rise to a number of recommendations. Employee engagement improved the most in the case of organisational
strategy and implementation, yet a number of demographic groups were not very engaged in this factor. Generation X and the Baby Boomers, persons with more than one year of service, all race groups except whites, all qualification groups except those with post-graduate qualifications and the job levels supervisors, and the employee. From the engagement survey results, it would appear that there is a lack of supervisor and employee involvement in organisational strategy development and implementation.

According to Jarzabkowski and Spee (2009), employees “do strategy” which can positively improve organisational performance. It thus makes good sense for senior management to enhance strategy implementation by creating an environment that fosters engagement (Shuck and Rose, 2015), especially for Africans and employees at the lower hierarchical levels. Management needs to remain mindful of the value of the engagement of employees at all levels in strategy implementation. In the South African context and given the results of this study, it might be worthwhile for management to investigate the use of group techniques such as world café methodology or appreciative inquiry (Viljoen, 2015), to involve employees of all demographic groups, of all qualifications, at all job levels and regardless of their years of service. The comparative results indicate very high levels of team commitment, and also a significant improvement for 2015. This is an indication that organisations can leverage group techniques as an intervention. Both the aforementioned group techniques, namely, world café methodology and appreciative inquiry, can be applied using an African approach (Viljoen, 2015). This can especially be done to explore various ways of implementing strategy. It thus seems that if organisations intend to improve the engagement levels of all their employees, they will need to be multi-culturally sensitive in the way they manage and implement employee engagement.

Although the study provided new insights into the longitudinal effects of employee engagement and the engagement levels of the various demographic groups, the study is not without its limitations. Longitudinal studies are expensive and time-consuming, which might have an impact on similar future studies. It is also very difficult to survey the same sample group over a period of years, due to the mobility of the sample members as well as testing effects whether respondents lose interest or are more sensitive to the construct in question. This limitation is deemed to be overcome because populations members are deemed similar (Hasset and Paavilainen-Mäntymäki, 2013; Menard, 2002; Taris and Kompier, 2014). The researchers propose that future longitudinal research starts with a specific selected sample which can be used consecutive surveys, after obtaining approval from the participants.

References


Influence of time on employee engagement


Further reading

About the authors
Nico Martins holds a PhD degree in Industrial Psychology and is with the Department of Industrial and Organisational Psychology at the University of South Africa. Nico specialises in the field of Organisational Psychology and is employed as Research Professor. His research focusses on organisational culture/climate, organisational trust, employee engagement and organisational diagnoses. He is Chief Editor of the *Southern African Business Review*. Nico Martins is the corresponding author and can be contacted at: martinsn@mweb.co.za

Hester Nienaber fulfils the role of Research Professor at the College of Economic and Management Sciences at the University of South Africa; and Visiting Research Professor at the School of Hotel and Restaurant Management, Northern Arizona University, Flagstaff, Arizona. Her research theme is “shaping competitive advantage by unlocking human competence”. She has a number of publications in peer-reviewed journals (sole and co-authored with colleagues around the globe); participates in peer-reviewed conferences; and reviews papers for journals and conferences. Currently, she is NRF-rated Researcher; recipient of Emerald Literati Network Best Paper Award (2011) and recipient of Excellence in Tuition Award (2010).
Prioritizing warehouse performance measures in contemporary supply chains

Tritos Laosirihongthong
Department of Industrial Engineering, Thammasat University, Bangkok, Thailand
Dotun Adebanjo
Business School, University of Greenwich, London, UK
Premaratne Samaranayake
School of Business, Western Sydney University, Penrith, Australia
Nachippan Subramanian
School of Business Management and Economics, University of Sussex, Brighton, UK, and
Sakun Boon-itt
Thammasat Business School, Thammasat University, Bangkok, Thailand

Abstract
Purpose – Due to the importance of efficiency and responsiveness measures rather than just efficiency measures, this research recognizes both measures when considering overall performance of warehouse operations. Thus, the purpose of this paper is to prioritize overall performance measures associated with warehouse operations in manufacturing, third-party logistics service provider and retail industry supply chains.

Design/methodology/approach – The study uses an integrated approach that involves the Q-sort method to group measures into four categories. Fuzzy analytical hierarchy process was then used to prioritize individual performance measures within each category and integer liner programming model was used to validate prioritized categories, using the judgment of multiple decision makers across three industries.

Findings – The result shows that the financial category is a dominating performance category in managing warehouse operations across all three industries selected. Within the financial category, cost of insurance accounted for 25 percent of total weight of the category, and is considered to be a powerful measure. The financial category is verified by multiple decision makers across three industries, as the most important performance category.

Research limitations/implications – As part of adopting the proposed methodology in practice, it needs to be guided by overall methodology appropriate for industry-specific contexts.

Originality/value – Key novel aspects of this study are to categorize warehouse operations measures and analyze their perspectives in different industries, understand dominant categories of warehouse operations measures in the contemporary supply chain and finally to explore to what extent current practices lead to achieving efficiency and responsiveness in the selected industries.

Keywords AHP, Fuzzy sets, Priorities, Performance categories

Paper type Research paper

1. Introduction

In today’s business environment characterized by increasing globalization, intense competition and customer sophistication, firms continue to change their global business operations to improve overall performance. In this context, warehouse operations play a significant and critical role in achieving better performance through various improvement

The authors would like to thank Ms Sujika Boonyarit for her contribution on data collection and preliminary data analysis. The authors would also like to thank Thailand Research Fund (TRF), for the under Grant No. BRG6080012, the Center of Excellence in Operations and Information Management, Thammasat University and the Bualuang ASEAN Chair Professorship grants for their supports.
methods such as manufacturing postponement and centralized distribution (Nair, 2005; Tse et al., 2012), adopting autonomous vehicle-based storage and retrieval systems (Roy et al., 2012) and optimally coordinating and integrating the interrelated decisions of production sequencing and vehicle routing (Park and Hong, 2009). It is apparent that many smaller warehouses are being replaced by fewer large warehouses to realize economies of scale (de Koster et al., 2007). Furthermore, Bowersox et al. (2013) emphasize the need for integration of global manufacturing with logistics operations including service capabilities and transport support for efficient and effective business performance. The need for integration is exacerbated by increasing competition among businesses, in particular for maintaining required levels of competitive advantage through responsiveness, cost leadership and differentiation (Zailani and Rajagopal, 2005; Cagliano et al., 2006). In addition, businesses are faced with increasing trends of strategic partnerships, outsourcing, virtual logistics and green logistics (de Koster and Warffemius, 2005) for improving overall performance – implying the need for integrating global operations is much more needed than ever before.

In addition to the definitions of performance measures, other aspects discussed in literature include key logistics activities, logistics service requirements and the relationships between various factors including the effect of logistics capability on overall performance. Lu and Yang (2010) identify crucial logistics capability dimensions and classify international distribution center operators. de Koster and Warffemius (2005) distinguish a number of performance aspects associated with warehouse operations, but limited to international operations in Europe, while Cao and Jiang (2013) propose a service capability maturity model for optimum resource configuration solution in public warehouse operations.

More recently, Staudt et al. (2015) through a comprehensive literature review of warehouse performance measurement identified a variety of indicators and tools to measure warehouse performance while recognizing the lack of clear definitions for some of those measures. While there is much work on global operations within the context of integration and the impact of overall performance, in particular improving inventory through integration of production sequencing and vehicle routing (Park and Hong, 2009), network design of integrated e-supply chains (Dotoli et al., 2007; Xia and Tang, 2011) and improving global planning through integration of logistics functions (Wang et al., 2012), there is very limited research on performance measures at firm level and relative importance of each performance measure across large supply chains of integrated manufacturing and distribution networks. Furthermore, it has been reported that a performance measurement system consisting of a single measure is inadequate, not inclusive and ignores the interaction among important supply chain characteristics and critical aspects of organizational strategic goals (Beamon, 1999). In addition, Cuthbertson and Piotrowicz (2011) through a common framework for the empirical analysis of performance management systems conclude that performance measurement is a context-dependent process, tailored to specific supply chain requirements. Therefore, it is recommended that for performance measurement systems to be successful, they should include three types of measures – resource measures, output measures and flexibility measures. In the context of measurement and improvement of warehouse performance, most of the research has focused on either warehouse/storage performance at the organization level, limited types of performance measures (e.g. only operational performance) (Sharma and Shah, 2016; Sharma and Shah, 2015) and performance measurement systems for measuring operational performance during disaster response in humanitarian supply chains (Santarelli et al., 2015). There has been limited research on the impact of warehouse performance on overall performance in global supply chains, which are increasingly characterized by supply chain disruptions (Samaranayake et al., 2011;
Lee and Rha, 2016), increasing competitiveness with customers demanding quick response and speedy deliveries (Nair, 2005). This study aims to address these issues. Consequently, the key research questions are twofold:

RQ1. What are the key warehouse performance measures and categories?

RQ2. What are the priorities of those categories from a contemporary supply chain perspective?

Thus, the purpose of this study is to categorize warehouse operations measures and analyze their perspectives in different industries, understand dominant categories of warehouse operations measures in the contemporary supply chain and finally to explore to what extent current practices lead to achieving efficiency and responsiveness measures in the selected industry context. Consequently, this research examines how overall performance measures are influenced by individual performance measures by analyzing the relative importance of each one within each category and verifying ranking/priorities of categories through the judgment of multiple decision makers from three industries (manufacturing, third-party logistics (3PL) and retail) selected. Overall performance is considered by incorporating a combination of efficiency and responsiveness-related performance measures associated with warehouse operations. The research is based on the comparison of three types of organizations operating in Asia – manufacturers, retailers and 3PL companies. It is important to study all three types as they all operate warehouses distinctly even though they may work in the same supply chain and thus, their operational perspectives and motivations may be different. The research adapts a three-stage methodology that uses Q-sort method to categorize warehouse performance measures, fuzzy-based analytical hierarchy process (AHP) approach to prioritize measures within each group, and validate prioritized (ranked) categories, using the judgment of multiple decision makers across three industries. The proposed three-stage approach is an extension of commonly adopted hybrid approach of Q-sort method and AHP for prioritizing performance measures. It incorporates three types of measures (resource, output and flexibility) and validates ranking/priorities of categories using judgments from multiple decision makers. This study provides important theoretical contributions to the literature on broader warehouse management and, in particular, selecting the right combination of performance measures from a range of efficiency and responsive-related measures for not only improving global operations at firm level, but also making right choices of performance measures as required by the competitive pressures of the supply chain that warehouse operations are engaged in. Using practice-based view (Bromiley and Rau, 2014), this study suggests what current practices companies need to reconfigure to achieve both efficiency and responsive measures in the warehouse operations context. Within the context of warehouse management, this study contributes to an understanding of the measurement of warehouse performance. According to Faber et al. (2013), warehouses have become an increasingly complex context to manage and this study, therefore, investigates performance measures within an increasingly complex environment. The remainder of the paper is structured as follows. Literature review on broader warehouse management and specific performance measures are presented next, followed by the research methodology and data analysis. The research findings section is followed by discussion and conclusion which include research implications for industrial practitioners as well as limitations and future work.

2. Literature review
Improving warehouse performance in global operations is a demanding task, particularly within an environment of increasing competition, customer sophistication and uncertainty in demand and supply in large supply chain networks. This is evident from recent trends in
increasing demand for value-added services, automated processing and information technology (IT) (Min, 2006). In addressing competitive pressures of business, Nair (2005) claim that firms are adopting policies such as manufacturing postponement and centralized distribution to improve their operational performance. By using a conceptual model of operational policies and performance, the study of Nair (2005) shows that there is a positive relationship between operational policies and performance.

2.1 Performance measures and evaluation

The measurement of performance was highlighted by Chia et al. (2009) as an important activity within the context of supply chains as it helps to drive strategic performance. This may be because metrics can help to understand the outcomes of organizational activity (Jothimani and Sarmah, 2014). Warehouses not only play an important connecting role in the supply chain, but they also impact cost and have become complex entities to manage. It is important, therefore, to continuously investigate how their performance is measured (Faber et al., 2013). Various studies have identified a number of performance aspects in different contexts of warehouse management, including: productivity, flexibility and outbound logistics (de Koster and Warffemius, 2005); productivity, delivery competence and responsiveness (Nair, 2005); service capability through storage, transportation, cost control and time control (Cao and Jiang, 2013); economic and technical-related performance measures (Johnson and McGinnis, 2011); inventory accuracy, timely delivery service, individual order fulfillment, flexible value-added service and responsiveness to special customer requests (Min, 2006); three critical logistics service capabilities (Lu and Yang, 2010); and cost, throughput, space utilization and service (Gu et al., 2010).

Performance evaluation has been considered from various perspectives including warehouse design and operation (Gu et al., 2010), improvements of warehouse operations, in particular, postponement operations (Tse et al., 2012) and relationships between operational policies and performance (Nair, 2005). Beamon (1996) presented characteristics which make a performance measurement system effective and they are inclusiveness, universality, measurability and consistency.

There are many performance measures identified by various studies. Lu and Yang (2010) indicated two different types of performance measures: financial and non-financial measures. Apart from broader performance measures such as those mentioned above, Lu and Yang (2010), based on a comprehensive literature review, identified seven common measures which are profit rate, sales growth rate, reduced operation cost, return on investment, market share growth, customer relationship and customer satisfaction.

Given the large number of performance measures associated with warehouse operations, improvements have been sought in various contexts including the need for developing logistics capabilities (Lai, 2004), direct linkage between quality customer service and supplier’s performance (Sharma et al., 2004) and the impact of the complexity of a warehouse on the warehouse’s performance (Faber et al., 2002).

2.2 Improvement practices and performance

Since the inception of various concepts on logistics systems and distribution practices, many theoretical and empirical studies have reported outcomes of various practices to achieve superior performance with lower investment (Chen et al., 2007). Improvements in warehouse performance are achieved in many ways including practices that enhance operations efficiency and flexibility using autonomous vehicle-based storage and retrieval systems (Roy et al., 2012), and adopting manufacturing postponement and centralized distribution techniques with value chain flexibility (Nair, 2005). In addition, Min (2006) proposed a warehouse management system that is designed to speed up order turnaround time, improve inventory accuracy, provide instant order status information, manage
warehouse space and enhance labor productivity. Similarly, Tse et al. (2012) proposed a hybrid intelligent system for improving postponement operation in warehouses, by integrating case-based reasoning and fuzzy logic.

Many practices or systems proposed so far concentrated on cost reduction, time reduction and reliability aspects of performance measures, such as delivery of products in a cost effective manner through manufacturing postponement and centralized distribution (Pagh and Cooper, 1998) and meeting expected demand with the lowest possible cost with inter-facility mold transfers (Aghezzaf, 2007). Others include lower operational cost and improved mass customization flexibility through postponement operation in warehouse (Tse et al., 2012), inventory accuracy, timely delivery service and individual order fulfillment through warehouse management system (Min, 2006), service capability and optimization of warehouse configuration through warehouse product service system (Cao and Jiang, 2013) and meeting delivery deadlines with reduced inventory using trans-shipments (Lau and Nakandala, 2012). Studies have also reported on technical efficiency aspects of warehouse operations (Johnson and McGinnis, 2011), performance at firm level, distribution centers (Lu and Yang, 2010; Roy et al., 2012), warehouse design and performance evaluation (Gu et al., 2010).

The literature presented shows that there has been much academic interest in the performance of warehouses and the use of performance measures. However, it has also been shown that supply chains and warehousing operations have become more complex. While there is significant research on understanding practices to improve certain measures, it is not well known how to categorize performance measures and to target key performance measures at warehouse level within specific industries. Warehousing operations are now carried out by different types of organizations in the supply chain including manufacturers, suppliers and 3PL organizations. There is yet no direct ranking or comparison of warehousing performance measurement across these types of organizations and particularly from a quantitative perspective. This is a significant gap bearing in mind that the different positions that these types of organizations occupy in the supply chain may affect their priorities when it comes to managing warehousing operations. In particular, a better understanding of the differences in weighting and prioritization given to warehousing performance measures could provide insights into how to improve warehousing operations from a balanced resilience and an integrated supply chain perspective. This study, based on the practices and experiences of organizations in four South East Asian countries (Thailand, Vietnam, Malaysia and Indonesia), addresses this gap. This is achieved through a mixed research design approach with both model development and experimental investigation. The proposed hybrid approach categorizes performance measures and captures the main performance measures associated with global warehouse operations and relative importance based on prior knowledge using situational analysis of fuzzy logic as part of a broader approach. Relative importance of key performance measures is based on fuzzy AHP and is compared with weights obtained from fuzzy-based group prioritization approach using fuzzy pairwise comparison matrices.

3. Research methodology and data analysis
The proposed research methodology comprises of three stages. The main focus of the first stage was to determine the potential performance measures applicable in managing warehouse operations. This stage involved, content validity analysis using identification of performance measures from a comprehensive literature review on broader supply chain performance and categorizing those measures using Q-sort method (Rajesh et al., 2011) through interviews of practitioners involved in managing warehouse operations. The second stage involved prioritization of performance measures within categories using fuzzy AHP method. At the third stage, priorities/ranking of performance categories are validated
using judgments of multiple decision makers. The three-stage research methodology outlined above is summarized and is presented in a schematic view as shown in Figure 1.

3.1 Categorization of performance measures
The literature review on performance measures focusing on warehouse operations identified 32 performance measures as presented in Table I.

In total, 19 warehouse managers from organizations in South East Asia (Thailand, Vietnam, Malaysia and Indonesia) were interviewed and asked them to classify each performance measure into one of four different categories. This activity was carried out during a Corporate Annual Meeting held in Bangkok, Thailand where the senior management staff representing regional offices and the three selected industries attended. The four categories were selected based on the level 1 performance metrics of supply chain operations reference model which include accuracy (reliability); flexibility and responsiveness; cost; assets (resources) (Huan et al., 2004). In this context, accuracy can be measured using metrics associated with inventory and warehousing functions of supply chains (Hou et al., 2010) and is directly influenced by the level of integration of those functions supported by computer integrated manufacturing.

Subsequent analysis of the interview findings indicated that out of 32 performance measures, 19 performance measures could be classified into four categories with an acceptable level of Cohen’s $\kappa$ coefficient (Cohen, 1960). Hence, Table II and Figure 2 show the Cohen’s $\kappa$ of three industries. These are: accuracy ($\kappa = 0.8175$), resources utilization ($\kappa = 0.7317$), financial outcomes ($\kappa = 0.7504$) and responsiveness and flexibility ($\kappa = 0.7576$), while Table III shows Cohen’s $\kappa$ of manufacturing, retail and 3PL industries.

In the second stage, the study applied an integrated Chang’s (1996) extent analysis on fuzzy AHP to determine the relative weight of importance among performance measures and the four categories identified above. Relative weights of importance of performance measures, based on fuzzy AHP, are compared with those of fuzzy group prioritization method. Chang’s extent analysis method is recognized as unsuitable when there are cases of

![Figure 1. Three-stage research methodology](image-url)
| Warehouse Performance Indicators                      | (a) | (b) | (c) | (d) | (e) | (f) | (g) | (h) | (i) | (j) | (k) | (l) | (m) | (n) | (o) | (p) | (q) | (r) | (s) | (t) | (u) | (v) | (w) | (x) | (y) | (z) | (aa) | (bb) | (cc) | Total |
|-------------------------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Inventory accuracy                                    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 9    |
| Accuracy in order picking                            |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 6    |
| Accuracy in order shipping                           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 4    |
| % product transferred without trans. errors           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 4    |
| Inventory obsolescence                                |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2    |
| % orders/lines received with shipping docs            |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2    |
| Capacity utilization                                  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 5    |
| Equipment utilization                                 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 5    |
| Labor productivity and labor utilization              |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 8    |
| Information availability and stability                |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 9    |
| Shipping cost                                         |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 13   |
| Inventory on-hand                                     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 7    |
| Product damage rate                                   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 4    |
| Insurance price                                       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2    |
| Maintenance cost                                      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 3    |
| Order cycle time                                      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 11   |
| On-time delivery                                      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 12   |
| Responsiveness to urgent deliveries                   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2    |
| Unloading/Loading time                                |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2    |
| Traceability                                          |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2    |
| Average lateness of order                            |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 7    |
| Transportation speed                                  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 5    |
| Customer query time                                   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 5    |

(continued)
| Warehouse Performance Indicators | (a) | (b) | (c) | (d) | (e) | (f) | (g) | (h) | (i) | (j) | (k) | (l) | (m) | (n) | (o) | (p) | (q) | (r) | (s) | (t) | (u) | (v) | (w) | (x) | (y) | (z) | (aa) | (bb) | (cc) | Total |
|----------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| In-transit time                  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2   |
| Number of stock-outs             | X   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 5   |
| Product variety                  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2   |
| Order size flexibility           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2   |
| Stock turnover                   | X   | X   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 8   |
| Carrier reliability              |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 3   |
| Delivery flexibility             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 3   |
| Flexibility of service systems   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 2   |

Notes: (a) Mondragon et al. (2011); (b) Baker and Halim (2007); (c) Bennett and Khug (2012); (d) Birou et al. (2011); (e) Chae (2009); (f) Colson and Dorigo (2004); (g) Evangelista et al. (2012); (h) Gu et al. (2007); (i) Gu et al. (2010); (j) Gurasekaran et al. (2001); (k) Gotzamani et al. (2010); (l) Hoek (2001); (m) Jeffers (2010); (n) Green et al. (2008); (o) de Koster and Warffemius (2005); (p) Meepetchdee and Shah (2007); (q) Milgate (2001); (r) Petersen (2002); (s) Petersen et al. (2005); (t) Poon et al. (2009); (u) Rahman and Wu (2011); (v) Banomyong and Supatn (2011); (w) Sheu et al. (2006); (x) Sohn et al. (2007); (y) Visich et al. (2009); (z) Wouters and Sportel (2005); (aa) Yang et al. (2012); (bb) Yu et al. (2012); (cc) Zailani and Rajagopal (2005)
having irrational “zero weight” of criteria which cause wrong priority weights (Wang et al., 2008). Therefore, Chang’s extent analysis method on fuzzy AHP is a reliable method except in such cases (Vaziri and Beheshtinia, 2016). Since this study did not encounter the problem of a zero weight for any criterion, Chang’s extent analysis was deemed suitable. Details are presented in the subsequent section.

<table>
<thead>
<tr>
<th>Category</th>
<th>$K_j$</th>
<th>$\text{Var}(K_j)$</th>
<th>$(K_j)/\text{SE}(K_j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.7509</td>
<td>0.0001812</td>
<td>55.785</td>
</tr>
<tr>
<td>Category A</td>
<td>0.8175</td>
<td>0.0044636</td>
<td>12.236</td>
</tr>
<tr>
<td>Category B</td>
<td>0.7317</td>
<td>0.0034087</td>
<td>12.532</td>
</tr>
<tr>
<td>Category C</td>
<td>0.7504</td>
<td>0.0049885</td>
<td>10.646</td>
</tr>
<tr>
<td>Category D</td>
<td>0.7576</td>
<td>0.0067133</td>
<td>9.247</td>
</tr>
<tr>
<td>na</td>
<td>−0.0140</td>
<td>0.0017136</td>
<td>−0.339</td>
</tr>
</tbody>
</table>

Table II. Cohen’s $\kappa$ of three industries

Prioritizing warehouse performance measures

Figure 2. Overall result from Q-sort method

Cohen’s $\kappa$ of manufacturing industry | Cohen’s $\kappa$ of retail (after 2 rounds) | Cohen’s $\kappa$ of 3PLs

<table>
<thead>
<tr>
<th>Category</th>
<th>$K_j$</th>
<th>$\text{Var}(K_j)$</th>
<th>$(K_j)/\text{SE}(K_j)$</th>
<th>$K_j$</th>
<th>$\text{Var}(K_j)$</th>
<th>$(K_j)/\text{SE}(K_j)$</th>
<th>$K_j$</th>
<th>$\text{Var}(K_j)$</th>
<th>$(K_j)/\text{SE}(K_j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.8031</td>
<td>0.0004675</td>
<td>37.144</td>
<td>0.6666</td>
<td>0.0011850</td>
<td>19.365</td>
<td>0.6881</td>
<td>0.0008755</td>
<td>23.256</td>
</tr>
<tr>
<td>Category A</td>
<td>0.8752</td>
<td>0.0157156</td>
<td>6.981</td>
<td>0.7427</td>
<td>0.0336788</td>
<td>4.047</td>
<td>0.7409</td>
<td>0.0249595</td>
<td>4.890</td>
</tr>
<tr>
<td>Category B</td>
<td>0.7349</td>
<td>0.0127920</td>
<td>6.497</td>
<td>0.7020</td>
<td>0.0309531</td>
<td>3.990</td>
<td>0.6842</td>
<td>0.0214737</td>
<td>4.669</td>
</tr>
<tr>
<td>Category C</td>
<td>0.7783</td>
<td>0.0167229</td>
<td>6.018</td>
<td>0.6621</td>
<td>0.0370067</td>
<td>3.442</td>
<td>0.7159</td>
<td>0.0261287</td>
<td>4.429</td>
</tr>
<tr>
<td>Category D</td>
<td>0.8580</td>
<td>0.0205613</td>
<td>5.984</td>
<td>0.6517</td>
<td>0.0409581</td>
<td>3.139</td>
<td>0.6804</td>
<td>0.0398800</td>
<td>3.891</td>
</tr>
<tr>
<td>na</td>
<td>−0.0133</td>
<td>0.0163802</td>
<td>−0.104</td>
<td>−0.0213</td>
<td>0.0476187</td>
<td>−0.098</td>
<td>−0.0179</td>
<td>0.0316353</td>
<td>−0.100</td>
</tr>
</tbody>
</table>

Table III. Cohen’s $\kappa$ of manufacturing, retail and 3PL industries

Dimensions of performance indicators for warehouse management

- Accuracy (0.8175)
- Resources utilization (0.7317)
- Financial outcome (0.7504)
- Responsiveness and Flexibility (0.7576)
3.2 Prioritization of performance measures using fuzzy AHP

After validating the constructs of performance measures in managing warehouse operations through Q-Sort method, analytical hierarchy structure was established to determine the relative weight of importance among constructs (categories) and measures. In this study, a panel of experts was selected based on their experience. Nahm et al. (2002) argue that the number of experts should be large enough to assure multiple perspectives, and small enough to make the research manageable. The experts came from three different industries: manufacturing, retail and 3PL provider with operations in South East Asia (Thailand, Vietnam, Malaysia and Indonesia). As shown in Table IV, 20 practitioners were interviewed during June–August 2012.

In order to determine the judgment matrix, these experts were asked to pairwise comparable categories using linguistic terms which were subsequently expressed by fuzzy numbers. Many published studies on prioritizing performance measures have developed a wide variety of models related to experts’ judgments (Reisinger et al., 2003; Roberts and Philip, 1996; Chan and Wu, 1998). Among many multiple criteria decision making methods, AHP has been widely used by many researchers (Wang et al., 2008), since AHP provides an ideal tool for a sequence of multi-objective decision-making problem (Dong, 2013). Although the traditional AHP method is suited to prioritizing performance measures in order to incorporate the opinion of experts (stakeholders or decision makers), however, it cannot reflect human thinking because of the imprecision and vagueness of decision makers’ judgments. AHP with its fuzzy extension, so-called fuzzy AHP, therefore, was developed to compensate the deficiency in traditional AHP and it has been widely used in the past by many studies for prioritization (Ayag, 2005; Chen et al., 2006). In this study, fuzzy AHP was applied to obtain more decisive judgments by weighting the performance measures in the presence of vagueness of experts’ preferences. The approach of fuzzy AHP to calculate weights of performance measures is described as follows:

1. Develop a hierarchical structure for prioritizing the performance measures.

A fuzzy AHP model (Figure 3), based on the identified potential measures and associated categories, is developed and presented. With a hierarchical structure, a complicated and

<table>
<thead>
<tr>
<th>Business type</th>
<th>Assessor No. 1</th>
<th>Assessor No. 2</th>
<th>Assessor No. 3</th>
<th>Assessor No. 4</th>
<th>Assessor No. 5</th>
<th>Assessor No. 6</th>
<th>Assessor No. 7</th>
<th>Assessor No. 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>Warehouse ass. manager</td>
<td>Engineer</td>
<td>Process analysis engineer</td>
<td>Factory manager</td>
<td>Warehouse supervisor</td>
<td>Procurement engineer</td>
<td>Warehouse manager</td>
<td>General manager</td>
</tr>
<tr>
<td>Third-party logistics providers (3PLs)</td>
<td>Manager</td>
<td>Supervisor</td>
<td>General manager</td>
<td>Supervisor</td>
<td>Assistant manager</td>
<td>Manager</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>Retails</td>
<td>Manager</td>
<td>Manager</td>
<td>Supervisor</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>3</td>
<td>14</td>
</tr>
</tbody>
</table>

Table IV.
Qualification of practitioners joining in fuzzy AHP

<table>
<thead>
<tr>
<th>Business type</th>
<th>Assessor No. 9</th>
<th>Assessor No. 10</th>
<th>Assessor No. 11</th>
<th>Assessor No. 12</th>
<th>Assessor No. 13</th>
<th>Assessor No. 14</th>
<th>Assessor No. 15</th>
<th>Assessor No. 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>Manager</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>3</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Third-party logistics providers (3PLs)</td>
<td>5.5</td>
<td>2</td>
<td>11</td>
<td>3</td>
<td>10</td>
<td>7.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
complex problem is converted to a hierarchical system of elements. The hierarchical structure systematically accommodates the use of expert judgment:

(2) Establish a fuzzy judgment matrix (or a pairwise comparison matrix).

A panel of experts was asked to make pairwise comparisons for elements and questionnaires were provided to collect information from the experts. Each expert was asked to assign linguistic terms based on his/her subjective judgment, to the pairwise comparisons by asking which one of two categories is more important and how much more important it is with respect to the preceding element. In this case, linguistic terms adopted include five different scales: equally important, moderately important, fairly important, very important (VI) and absolutely important. In decision making, each expert gave his/her preference on the categories identified in the above step (Step 1) of the procedure. In the case of pairwise comparison of categories using linguistic terms, each expert’s judgment resulted in \((n-1)\) of comparisons for \(n\) number of categories. For the four categories identified above, there are six pairwise comparisons from each expert. The fuzzy judgment matrix reflects the relative importance of the decision categories. After recoding answers from the experts in linguistic terms, these linguistic judgments are then converted to triangular fuzzy sets:

(3) Combine the opinions from several experts by using geometric mean.

The perception of each expert varies according to individual experience and knowledge. In order to get a consistent and fair outcome from several experts’ subjective judgments, the informed judgments were aggregated through the geometric mean of individual experts’ judgments. By using the geometric mean method to derive the fuzzy weight, therefore, different judgmental values can be converted to one element in the fuzzy judgment matrix. Let \(M^k_{ij}\) represent a subjective judgment of the \(k\)th expert for the relative importance of two elements (the \(i\)th element and the \(j\)th element), then the fuzzy geometric mean \(M_{ij}\) from \(m\) experts is shown in following equation.

Fuzzy geometric mean:

\[
M_{ij} = \left( M^1_{ij} \otimes M^2_{ij} \otimes \cdots \otimes M^m_{ij} \right)^{1/m}.
\]

(4) Repeat the calculation of the local priority weights for other categories of performance measures.

(5) Calculate the global priority weight of each element.

---

**Figure 3.** Hierarchy structure of warehouse performance measures
The global priority weight of each element is calculated by multiplying its local weight with its corresponding weight along the hierarchy. Tables V–IX show the results from fuzzy AHP process from the three different groups of experts, while Table X presents the overall result. There are a variety of extensions to the fuzzy AHP approach that can increase its usefulness for managerial decision making. For prioritizing the performance measures, in this study,

### Table V.
Relative weights of four categories of performance measures

<table>
<thead>
<tr>
<th>Indicator dimensions</th>
<th>Manufacturing</th>
<th>3PLs</th>
<th>Retails</th>
<th>Total 3 Business</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Accuracy</td>
<td>0.138</td>
<td>0.240</td>
<td>0.269</td>
<td>0.223</td>
</tr>
<tr>
<td>B. Resource utilization</td>
<td>0.308</td>
<td>0.248</td>
<td>0.091</td>
<td>0.222</td>
</tr>
<tr>
<td>C. Financial outcome</td>
<td>0.457</td>
<td>0.344</td>
<td>0.320</td>
<td>0.370</td>
</tr>
<tr>
<td>D. Responsiveness and flexibility</td>
<td>0.096</td>
<td>0.168</td>
<td>0.320</td>
<td>0.176</td>
</tr>
</tbody>
</table>

### Table VI.
Relative weights of five measures within "Accuracy" category

<table>
<thead>
<tr>
<th>Accuracy dimension</th>
<th>Manufacturing</th>
<th>3PLs</th>
<th>Retails</th>
<th>Total 3 businesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1. Inventory accuracy</td>
<td>0.213</td>
<td>0.182</td>
<td>0.124</td>
<td>0.178</td>
</tr>
<tr>
<td>A2. Accuracy in order picking</td>
<td>0.159</td>
<td>0.207</td>
<td>0.173</td>
<td>0.175</td>
</tr>
<tr>
<td>A3. Accuracy in order shipping</td>
<td>0.171</td>
<td>0.188</td>
<td>0.210</td>
<td>0.182</td>
</tr>
<tr>
<td>A4. % of product transferred without transaction errors</td>
<td>0.212</td>
<td>0.197</td>
<td>0.247</td>
<td>0.232</td>
</tr>
<tr>
<td>A5. % of order/lines received with correct shipping documents</td>
<td>0.245</td>
<td>0.225</td>
<td>0.247</td>
<td>0.232</td>
</tr>
</tbody>
</table>

### Table VII.
Relative weights of three measures within "Resources utilization" category

<table>
<thead>
<tr>
<th>Resources utilization dimension</th>
<th>Manufacturing</th>
<th>3PLs</th>
<th>Retails</th>
<th>Total 3 businesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1. Space utilization</td>
<td>0.165</td>
<td>0.278</td>
<td>0.043</td>
<td>0.183</td>
</tr>
<tr>
<td>B2. Equipment utilization picking</td>
<td>0.505</td>
<td>0.363</td>
<td>0.600</td>
<td>0.469</td>
</tr>
<tr>
<td>B3. Labor productivity and utilization</td>
<td>0.330</td>
<td>0.359</td>
<td>0.358</td>
<td>0.348</td>
</tr>
</tbody>
</table>

### Table VIII.
Relative weights of five measures within "Financial outcome" category

<table>
<thead>
<tr>
<th>Financial outcome dimension</th>
<th>Manufacturing</th>
<th>3PLs</th>
<th>Retails</th>
<th>Total 3 businesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1. Shipping cost</td>
<td>0.204</td>
<td>0.228</td>
<td>0.211</td>
<td>0.211</td>
</tr>
<tr>
<td>C2. Inventory holding cost</td>
<td>0.105</td>
<td>0.124</td>
<td>0.210</td>
<td>0.153</td>
</tr>
<tr>
<td>C3. Product damage rate</td>
<td>0.194</td>
<td>0.202</td>
<td>0.134</td>
<td>0.177</td>
</tr>
<tr>
<td>C4. Insurance cost</td>
<td>0.275</td>
<td>0.257</td>
<td>0.223</td>
<td>0.247</td>
</tr>
<tr>
<td>C5. Shortage cost</td>
<td>0.222</td>
<td>0.189</td>
<td>0.223</td>
<td>0.211</td>
</tr>
</tbody>
</table>

### Table IX.
Relative weights of six measures within "Responsiveness and flexibility" category

<table>
<thead>
<tr>
<th>Responsiveness and flexibility dimension</th>
<th>Manufacturing</th>
<th>3PLs</th>
<th>Retails</th>
<th>Total 3 businesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1. Responsiveness to urgent deliveries</td>
<td>0.190</td>
<td>0.232</td>
<td>0.128</td>
<td>0.178</td>
</tr>
<tr>
<td>D2. Transportation speed</td>
<td>0.110</td>
<td>0.173</td>
<td>0.195</td>
<td>0.153</td>
</tr>
<tr>
<td>D3. Customer query time</td>
<td>0.142</td>
<td>0.005</td>
<td>0.209</td>
<td>0.142</td>
</tr>
<tr>
<td>D4. Order size flexibility</td>
<td>0.217</td>
<td>0.258</td>
<td>0.169</td>
<td>0.210</td>
</tr>
<tr>
<td>D5. Delivery flexibility</td>
<td>0.143</td>
<td>0.158</td>
<td>0.169</td>
<td>0.149</td>
</tr>
<tr>
<td>D6. Service system flexibility</td>
<td>0.199</td>
<td>0.173</td>
<td>0.129</td>
<td>0.167</td>
</tr>
</tbody>
</table>
they were grouped into categories. These categories are compared pairwise first, then the performance measures are compared pairwise with respect to their criteria. The performance measures were not compared to their counterparts across the categories. In doing so, a smaller number of the pairwise comparisons are required.

Table V shows that financial outcome is the most important measure across all three industry sectors. Retailers however, place significantly less emphasis on resources utilization compared to the other two sectors. Rather, they prefer to emphasize responsiveness and flexibility.

With respect to the accuracy dimension, the percentage of products transferred without transaction errors and the percentage of orders received with the correct shipping documents were deemed to be most important. Conversely, accuracy in order picking was deemed least important across the three industry sectors.

Within the resource utilization category, equipment utilization was the most important measure across all industry sectors while space utilization was the least important.

Within the financial outcomes category, cost of insurance was seen as most important. However, it is not clear whether this is an indication that the costs are too high or if it is an acknowledgment that insurance is seen as a priority to minimize financial loss. The least important measure was inventory holding cost. This may be because of the adoption of inventory practices such as JIT that minimize inventory holding.

With respect to responsiveness and flexibility, flexibility of order size was seen as most important across all three industry sectors and particularly by the logistics service providers. This may be because different order sizes affect their loading factors and efficient use of delivery vehicles. In contrast, the logistics services providers place virtually no importance on customer query time.

(6) Validate the priorities/ranks of performance categories using judgments of multiple decision makers.

Priorities of performance categories identified by weights from AHP (Table X) are validated by minimization of deviation between the ranking of each category by individual decision
maker and weighted average of each category. Decision makers are selected from the group of 20 warehouse managers that participated in the study (Table IV). The following key variables and parameters were defined for the minimization of deviation using integer programming model:

- $DM_j = j$th decision maker ($j = 1, 2, \ldots, m$).
- $r_{ij} = \text{rank of the } i$th performance category ($i = 1, 2, \ldots, n$) by $j$th decision maker ($j = 1, 2, \ldots, m$).
- $y_j = \text{weightage assigned to } j$th decision maker, depending on the importance of decision maker for the evaluation of ranking of categories.
- $Y_i = \text{weighted average of performance category } i$.

Each decision maker is assigned with a weight (less than 1, making total weight of all decision makers to 1). Since ranking/priorities are considered from 1 (the most prominent) to $n$ (least prominent), the reciprocal of weight of each decision maker is considered for arriving at weighted average for each category ($Y_i$). Thus, $Y_i$ is computed as follows:

$$Y_i = \sum_{j=1}^{m} \left( \frac{r_{ij}}{y_j} \right) \quad i = 1, \ldots, n. \quad (1)$$

In this case, there are four categories ($i = 1-4$) and six decision makers (DM1–DM6), (three from manufacturing, two from 3PLs and one from retail) for the validation of priorities/ranks. In this case, only six decision makers are selected to represent all three entities of the selected supply chain (manufacturer, 3PL and retailer) and they were senior managers at respective organizations. For this stage, it was necessary to involve only senior executives and not all stakeholders. This is because senior managers have organizational responsibility for setting priorities regarding the choice and prioritization of organizational measures. Therefore, their roles and responsibilities in the organization are different from the experts used in the first two stages of the research.

Since reciprocal weighted average of each performance category is considered, minimum deviation between weighted average and individual ranks is achieved by maximizing the reciprocal deviation, represented by $Z$. Further, it is emphasized using the multiplication of reciprocal deviation (sum of $|r_{ij} - Y_i|$) and rank (decision variable $- X_i$) for each performance category. For example, when reciprocal deviation is largest, it will be multiplied by the largest rank (least priority). Thus, the problem formulation is given by.

Objective function:

$$\text{Max } Z = \sum_{i=1}^{n} \sum_{j=1}^{m} (r_{ij} - Y_i) |X_i|, \quad (2)$$

subjected to:

$$1 \leq X_i \leq n \quad \forall i(1, 2, \ldots, n), \quad (3)$$

$$X_i \neq X_k \quad \forall i, k \text{ such that } i \neq k, \quad (4)$$

$$X_i \text{ is integer } \forall i(1, 2, \ldots, n). \quad (5)$$

- Equation (1): weighted average performance of category $i$. 
• Equation (2): the objective function of the model which maximizes deviation between the weighted average and individual rankings from different decision makers.
• Equation (3): restricts the ranking of n performance categories from 1 to n only.
• Equation (4): ensures that no two performance categories are given the same rank by the same decision maker.
• Equation (5): integer value of the rank is ensured.

The individual ranks by each decision maker for each performance category and evaluated weighted average are given in Table XI.

The proposed integer linear programming (ILP) model using individual ranks $r_{ij}$ by decision makers $DM_j$ and evaluated weighted average $Y_i$ values (Table XI) was solved using Excel Solver. The solution $(X_i)$ values of ILP model is given by: $X_1 = 3, X_2 = 2, X_3 = 1$ and $X_4 = 4$. In this case, the performance category of financial outcomes is identified as the highest priority ($X_3 = 1$), while the performance category of responsiveness and flexibility is the least priority ($X_4 = 4$). Thus, the performance category of financial outcomes is validated by judgments of multiple decision makers as the prominent ranking/priority of warehouse performance measures in contemporary supply chains.

### 4. Research findings

This section discusses the findings of our study from several perspectives. First, the study highlights the dominant performance measures in four dimensions using the literature review. Second, the study summarizes the reflections of practitioners about grouping performance measures under various dimensions. Finally, the study presents the findings of prioritization with respect to dominant dimensions and influencing performance measures.

#### 4.1 Dominant performance measures based on literature review

Previous studies, irrespective of industry emphasized on performance measures related to three aspects – accuracy, resource utilization and financial. The dominant measures cited by researchers in the three categories are as follows (Huan et al., 2004; Shepherd and Günter, 2006; Akyuz and Erkan, 2010):

1. resource utilization measures are capacity utilization and labor productivity and labor utilization;
2. accuracy measures are on-time delivery, order cycle time, inventory accuracy and stock turnover; and
3. financial measure is shipping cost.

Interestingly, there were not many studies that gave due importance to responsiveness and flexibility measures except transportation speed.

<table>
<thead>
<tr>
<th>Weightage</th>
<th>DM1</th>
<th>DM2</th>
<th>DM3</th>
<th>DM4</th>
<th>DM5</th>
<th>DM6</th>
<th>Weighted average ($Y_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reciprocal weightage</td>
<td>0.25</td>
<td>0.20</td>
<td>0.15</td>
<td>0.20</td>
<td>0.10</td>
<td>0.10</td>
<td>112</td>
</tr>
<tr>
<td>Category 1</td>
<td>4</td>
<td>5</td>
<td>6.67</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>112</td>
</tr>
<tr>
<td>Category 2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>Category 3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>99.67</td>
</tr>
<tr>
<td>Category 4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>57.33</td>
</tr>
<tr>
<td>Category 5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>147.67</td>
<td></td>
</tr>
</tbody>
</table>

Table XI. Rank ($r_{ij}$) for the $i$th performance category by $j$th decision maker ($DM_j$) and weighted average ($Y_i$)}
4.2 Reflections of practitioners
Practitioners opined similar views with respect to previous studies, irrespective of industry and were more consistent in grouping performance measures related to accuracy, resources utilization and financial outcome categories than flexibility and responsiveness. Out of the three aspects their perception about accuracy is almost consistent. In order to make these reflections applicable to specific industries, future studies may consider expanding the approach adopted in this study across other industries. This study is based on inputs from two distinct industries: Electronics/Electrical (High Tech) and Automotive Parts/Components manufacturing. These two industries have similar warehouse operations environment in nature such as working under the lean, supply chain and total quality management.

4.3 Prioritization of performance categories and measures
Composite rating based on three industries suggests that warehouses focus more on efficiency-related measures. Out of four categories, financial outcome had prominence in all industries selected for the study (Table X). Warehouse operations until now focus on how to minimize cost by excluding non-value-added activities. Resource utilization and accuracy construct trails after financial outcome category because all the industries attempt to effectively utilize their resources without any loss. Responsiveness and flexibility category did not have prominence in the overall analysis. The focus on cost rather than responsiveness finds resonance with the findings by de Koster and Balk (2008) that Asian warehouses pay more attention to cost, resource utilization and accuracy measures. The influencing individual performance measures within each category are discussed below.

Cost of insurance, shortage cost and shipping cost are considered to be the dominant measures in understanding the financial dimension in warehousing operations. It is interesting to note that the cost of insurance is considered as the most important measure rather than shipping and shortage costs. Equipment utilization and labor utilization have been considered as dominant measures to capture resource utilization dimensions. Industrial practitioners place more value on equipment and labor compared to space utilization.

Two measures – percentage product transferred without transaction errors and percentage orders/lines received with correct shipping documents were rated as the most valued measures to capture accuracy compared to other accuracy measures related to inventory, order picking and shipping. Out of six measures considered for the study, order size flexibility is the most important followed by responsiveness to urgent deliveries. Interestingly, transportation speed mentioned by various researchers was considered as the least influencing performance measure by the practitioners. In the future, practitioners need to evolve practices to concentrate on transportation speed and responsiveness-related measures.

4.4 Industry-related findings
The prioritization of the four categories and their performance measures are discussed below with respect to industries considered in this study. Practitioners representing the manufacturing industry considered financial outcome category and resource utilization dimension as the paramount performance measures of warehousing. In terms of measures within each category, most of the results are synonymous with overall results. Few deviations with respect to overall analysis are as follows: in the accuracy category, manufacturers gave more importance to inventory accuracy than percentage product transfer without transaction errors. In the responsiveness and flexibility category, manufacturers prefer to consider flexibility of service systems to meet as an important measure than responsiveness to urgent deliveries. This study also found that representatives from 3PL service providers gave similar weight to categories as the
manufacturers did. They gave importance toward financial outcome category and resource utilization category. As suggested by de Koster and Balk (2008), Asian warehouses being primarily privately owned are more focused on cost compared to their European counterparts, which focus more on responsive measures and practices. There are few variations in the prioritization of measures within each dimension per industry when compared to the overall analysis. The deviations within each category are as follows. Accuracy in order picking was given more importance than percentage orders/lines received with correct shipping documents. Shipping cost in financial category is considered to be more important than cost of insurance. Response to urgent deliveries was considered to be more important than order size flexibility in the responsiveness and flexibility category. Finally, this study shows that the financial outcome and responsiveness and flexibility categories are the most performance categories in the retail industry. It is interesting to note that retail industry is the only one which gives due importance to the responsiveness and flexibility category. The top 2 measures within all categories indicated by the three industries are similar to the overall analysis except the responsiveness and flexibility category. Customer query time was more prominent than order size flexibility.

5. Discussions and conclusions
This study set out to achieve three key aims; first, to identify and categorize performance measures; second, to investigate and establish relative importance of individual performance measures within the context of performance categories; and third, to validate priorities of performance measure categories using judgment of multiple decision makers selected from three industries. Using integrated research methodology, the study was able to categorize 19 performance measures into four categories based on Q-sort analysis. Perhaps, the key issue is not the measures that were chosen but the measures that were not chosen by the experts. For example, if considering the “financial outcome” category, a number of seemingly important measures such as inventory obsolescence, maintenance cost, number of stock-outs and stock turnover were not selected. Clearly, all warehouse operations would incur a maintenance cost and would be subject to a stock turnover. In addition, depending on the organization and industry involved, obsolescence and stock-outs could be an issue. The exclusion of these measures could imply one of a number of things. First, it could imply that the costs associated with these measures are not large enough for the measures to be seen as important by the experts. Second, it may be because the importance of these costs could vary from industry to industry and, therefore, the collective perception of importance is not as high as the factors that were chosen. Third, it may be that some of the costs are “hidden,” not immediately obvious or are not perceived as attributable to warehouse costs. For example, stock turnover and obsolescence may have significant financial implications but can be argued to be “hidden” costs. However, the fact that they are also not associated with other performance categories (e.g. obsolescence and stock turnover may also affect resource utilization) raises an interesting question about the perception of performance measures and corresponding practices in the industry. Most of the measures that were not chosen are bottom line measures which may not influence process performance (Ganesan et al., 2009). The key suggestion here is that there is no clarity about why some performance measures are associated with certain performance dimensions and others are not. There may be several influencing factors such as individual company systems, industrial sector practices or even national cultures.

With respect to the results from fuzzy logic analysis, it is interesting that the manufacturers place the greatest emphasis on financial outcomes and resources utilization while placing little on accuracy and responsiveness/flexibility. Conversely, their customers, the retailers place as much emphasis on responsiveness/flexibility as they do on financial outcomes while also rating accuracy as important. On the contrary,
they rate resource utilization quite low. The suggestion from these findings is that in spite of the much avowed understanding of customer needs and striving to fulfill them, manufacturers continue to manage for cost by following lean practices thereby shunning flexibility. This observation suggests in future manufacturing industry should focus on lean practices to achieve both performance measures. On the other hand, retailers maintain a focus on cost by demanding flexibility and accuracy. This diametric opposition raises questions about whether the methodological and technological advances that have been developed to bridge these gaps (e.g. Vendor Management Inventory (VMI), EDI, forecasting) have had any fundamental influence on practices in both the retail and manufacturing sectors. These challenges may be, in part, due to relatively less developed acceptance of these advanced technological solutions in developing parts of Asia when compared to developed western countries. It may also suggest that manufacturers, retailers and 3PLs in this part of the world are still largely focused on their own organizations rather than overall supply chain efficiency and integration.

5.1 Individual performance measures

Analysis of individual performance measures within the context of the three sectors provides interesting insights. With respect to “Accuracy,” there appears to be a good spread of emphasis across all five measures. However, with respect to “Resource utilization” both the retailers and manufacturers focus primarily on equipment utilization and noticeably less significantly on space utilization. In contrast, the 3PL companies have a more even spread across the three measures. In particular, low emphasis of the retailers may be the result of the following two factors. The first, as discussed in the previous paragraph, is their requirement for flexibility which implies less stock-holding and consequently, a reduced focus on space utilization. The second is the fast moving nature of the retail industry and the consequent ability to hold fewer inventories in warehouses. Indeed, many large retailers mainly make use of distribution centers which hold inventory only for very little time before sending them out to different retail locations. It may well be that the core need to efficiently cross-dock products from manufacturers and 3PL organizations is a key reason why retailers focus particularly on equipment utilization. On the other hand, manufacturers and, in particular, retailers who are tasked with the flexibility requirements of the retailers would have the practice of holding more inventory and so they are more focused on space utilization.

The “Financial outcomes” category suggests that compared to retailers, manufacturers and 3PLs are more focused on product damage rate and less on inventory holding cost. The focus on product damage rate is understandable, since products carry a higher risk of damage during transportation. Finally, with respect to “Responsiveness and Flexibility,” retailers and 3PLs place more emphasis on transportation speed. This could reflect their closeness to the consumer and variability in demand. It is evident from our study that retailers and 3PLs are more concerned about significant performance measures than bottom line measures. However, it could also be due to the need to deliver products within allocated time slots. In addition, manufacturers and 3PLs place more emphasis on responsiveness to urgent deliveries and order flexibility. This reflects the requirements of the retailer and may be related with the desire of retailers to hold minimal inventory. From a theoretical framework perspective, performance categories associated with warehouse operations in the supply chain are closely related to supply chain resilience framework (Pettit et al., 2010) where performance categories such as capability dimensions contribute to balanced resilience and improved performance. For example, responsiveness and flexibility category identified as a prominent measure of 3PLs and retailers is identified as one of the capabilities of supply chain resilience, contributing to balanced resilience involving multiple tiers of suppliers and customers of contemporary supply chain (Pettit et al., 2010).
5.2 Research implications, limitations and further work

The findings from this study have important implications for industry. For industry, there is a need to understand performance measures and also identify and use performance measures that not only complement their operations but that also reflect the priorities of their supply chain partners. Manufacturers and 3PLs need to understand that retailers are particularly keen on responsiveness and flexibility and so they need to focus more on this too. The disconnection between the priorities of Asian retailers, 3PLs and manufacturers suggests that there is a need for a balanced approach to managing warehouse operations from the supply chain partners. In particular, 3PLs and manufacturers need to change the focus of their warehousing operations and management. They need to enable their processes to be more responsive and flexible in order to react quickly to changes in the market. In an environment where consumers are demanding fast response from both online and traditional retailers, the manufacturers and 3PLs that service these retailers will have to rebalance their strategies and improve their operational efficiency, responsiveness and flexibility as well as the accuracy of their operations. This may imply that they need to co-ordinate their warehousing operations more closely with those of the retailers particularly as retailers are focused significantly on reducing inventory holding costs.

There are particular implications for 3PLs which find themselves as the intermediaries between manufacturers and retailers. They need to carefully balance the conflicts between their upstream and downstream supply chain partners. For example, manufacturers are more focused on space utilization than retailers, while there are also noticeable differences in focus on inventory holding costs and inventory accuracy. Therefore, 3PLs need to play a pivotal role in smoothing out supply chain imbalances particularly with respect to inventory management. This may imply that they sometimes act as temporary buffers between manufacturers seeking efficiency and retailers seeking agility. The implication is that much of the smoothing of the supply chain will be the responsibility of 3PLs. It is, therefore, important that they pay particular attention to the cost implications of playing this role and ensure that taking on the extra risk and responsibilities do not adversely affect their business. This does not appear to be a role that Asian 3PLs have embraced so far.

For retailers, they need to appreciate the challenges that manufacturers and 3PLs face with respect to responsiveness and particularly as it affects the cost of warehouse operations. While the need for responsiveness is important for retailers to satisfy ever-demanding consumers, there is a cost associated with this, in which retailers may not wish to bear. It is important for retailers to understand that manufacturers and 3PLs cannot indefinitely bear the cost of increasing demands for flexibility and responsiveness. To address these challenges, they need to share more relevant information with the supply chain partners and where possible, seek closer integration by adopting VMI, for example.

With respect to study limitations, in this research, performance measures were considered across efficiency and responsive categories, but the responsive category is limited to very few performance measures such as responsiveness to urgent deliveries and customer query time. Furthermore, another limitation of the research is the use of enhanced version of fuzzy AHP methods rather than using standard fuzzy AHP method in the proposed hybrid methodology. Thus, to demonstrate the robustness of the proposed approach, future research should focus on using enhanced version of fuzzy-based prioritization methods for deriving group priorities/weights. For future development, warehouse managers need to consider performance measures with a focus on practices that could improve socio-technical aspects such as those related to the quality of work life and sustainability. For academia, this study has shown that there are gaps in the current understanding of the factors that underpin the choice of performance measures and practices in warehouse management and in particular, the influence of supply chain managers in the choices made.
Future studies could consider the impact of practices including organizational or national culture to understand performance in warehouses as well as investigate the primary factors that drive the understanding and selection of performance measures. In addition, it will be worthwhile to investigate the use of different technologies in warehouses and how such technologies lead to evolution and change in practices and performance measurement decisions.

References


Further reading


Corresponding author
Premaratne Samaranayake can be contacted at: p.samaranayake@westernsydney.edu.au

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Assessing Egyptian construction projects performance using principal component analysis

Mohamed Mahdy Marzouk
Structural Engineering Department, Faculty of Engineering, Cairo University, Cairo, Egypt, and
Emad Fayez Gaid
Orascom Construction, Cairo, Egypt

Abstract
Purpose – The construction sector is a major contributor to the Egyptian economy and gross domestic products, plus considered as one of its fastest-growing sectors. Various deficiencies such as low productivity, delays, cost overrun, poor quality, etc. have plagued the construction sector, leading to undesirable project performance across Egypt and several other countries. One of the best methods to measure and improve performance evaluation and consequently a step toward industry improvement is key performance indicators (KPIs). The paper aims to discuss this issue.

Design/methodology/approach – A questionnaire survey was conducted to identify the importance index for 35 KPIs. The KPIs were identified and selected based on an extensive literature review and one on one interview with expert construction professionals. Moreover, exploratory factor analysis method has been used to analyze and determine the inter-relationships between the set of indicators.

Findings – Seven indicators that have the highest importance indicator were selected to create project overall performance indicator assessment models to consistently measure the performance of different construction projects. Seven equations are introduced reflecting the output of the research while considering both organization size and project type.

Practical implications – The proposed evaluation assessment models can be used to: evaluate the relative success of projects; indicate the areas of strengths and weaknesses in performance and establish company benchmarking. Three projects were selected to validate the performance evaluation models.

Originality/value – This paper provides mathematical assessment models for evaluating overall construction project performance in Egypt, taking into consideration organization size and project type. Previously published research works on the subject matter are quite limited, and frequently deal with only one or two selected aspects.

Keywords Key performance indicators (KPIs), Exploratory factor analysis (EFA), Overall performance, Performance assessment model, Principal component analysis (PCA)

Paper type Research paper

1. Introduction
The construction sector is comparatively a significant sector in both developed and developing economies. In developed economies, gross domestic products (GDP) contribution of the construction sector amounts to almost 10 percent, and more than 4 percent for developing economies (Lowe, 2003). Commonly, the construction sector is the catalyst which motivates the growth of a nation’s economy, especially during recession. The construction sector is a major contributor to the Egyptian economy and is considered as one of its fastest-growing sectors. During several past decades of the twentieth century, the construction sector in Egypt achieved a considerable growth rate at an average of 9.7 percent per year (Egypt Economic Report, 2015). It similarly made a significant contribution to the economy, averaging 4.8 percent of the GDP in 2015. Small Egyptian construction firms have contributed a generous share of this amount. In addition to this economical contribution, it also plays a significant role in providing employment to the Egyptian people working in both private and public domains. The construction sector post
2011 Egyptian uprising plays a very important role in economic activities resulting from the construction of many new “Mega Projects” such as the Suez Canal tunnels, high capacity power plants, various renewable energy power plants, upgrade of the roads and metro infrastructure, building the new administrative capital and north coast holiday and tourism cities. Evaluating performance on such large-scale projects can no longer continue with the previous insufficient methods, and an improvement in this regards is required. Another challenge to the Egyptian construction sector results from an increase in private investment in infrastructure projects through many financial arrangements such as EPC finance and private public partnership projects. Key performance indicators (KPIs) represent one of the best methods to measure and improve performance evaluation and consequently improvement of the industry as a whole.

Slow economic growth, strong competition, local and international construction sector restructuring have placed high pressure on construction firms to continually improve their productivity and performance to either survive or be profitable. Several developed countries have recognized the urgent need and importance of having a measurement system to improve the performance of the industry. Since the construction sector is very dynamic in nature, being affected by many different uncertainties such as the application of new technologies, budgets shortcomings, differing site conditions, political instabilities, execution techniques, etc. To deal with such uncertainties, different inter-related components that influence the performance must be considered.

Project performance evaluation has recorded noticeable attention in the last three decades. The UK was the pioneer of construction performance measurement revaluation by using KPIs. In the 1990s, the UK Government commissioned two significant reports “Constructing the Team” and “Rethinking Construction.” Both the reports contributed to acknowledging that improvements and change were required within the construction industry. This resulted in developing the indicators and criteria in which performance could be evaluated, improved, defined certain performance targets and start applying a wide range of KPIs.

Traditional methods of performance measurement are exclusively the use of financial measurement figures such as profitability, liquidity, revenues to net working capital, return on total assets, return on investment and discounted cash flow which focused on one perspective only and rely on historical information (Luu et al., 2008) without any consideration to the other project performance indicators such as safety, quality, productivity and customer satisfaction. Organizations top management needs an integrated control system that has two facets: strategy oriented and operation oriented. This research proposes a methodology for measuring the performance from three different aspects:

1. Location – Egypt as the domain for the research.
2. Project types – institutional and commercial building construction; industrial construction; and infrastructure and heavy construction.
3. Organization size – large, medium and small.

The proposed research is considered novel since it utilizes exploratory factor analysis (EFA) method to analyze and determine the inter-relationships between the KPIs. The research also proposes the use of project overall performance indicator (POPI) assessment model to measure the performance of construction projects based on multiple aspects.

2. Literature review
Performance measurement and benchmarking are the cornerstones of challenging any business to become world class (Beatham et al., 2004). Oke et al. (2016) stated that the failure of any construction project is mostly related to the performance problems. Neely et al. (2002)
stated that if there is no measure, there will be no improvement. Performance measurements are the quantitative indicators that are chosen to track the specific states of performance such as productivity, effectiveness, efficiency, customer satisfaction, quality and cost. They are historic in nature and do not provide the opportunity to improve. De Wit (1988) stated that it is still not clear how to measure project success since project participants identify success or failure in different ways. Lim and Mohamed (1999) indicated that projects success should be observed from different perspectives, whether from the individual client, developer, contractor, end user or the general public.

Performance measurement, as a management tool that determines success or failure of performance, whether organizationally or functionally, can give answers to three crucial questions: “How well is an organization is performing? Is the organization achieving its objectives? How much has the organization improved since the last period of measurement?” (Phusavat et al., 2009).

Throughout the twentieth century, managers have strictly used some financial metrics for planning and cost control purposes which were not able to predict the future performance or analyze the performance from an engineering point of views. Many efforts, such as reducing defects in products and training employees, are difficult to evaluate financially but they will play important roles in achieving success. Bourne et al. (2003a, b) claimed that in the late 1970s and till the early 1990s, managers and academic researchers started expressing their dissatisfaction with traditional financial based performance measurement systems. Beatham et al. (2004) stated that the problem was that financial results were usually lagging in describing the results of actions or decisions after its implementation by at least a reporting period (Ali et al., 2013), which did not give the chance to take any mitigation plans or corrective actions in the right time for improvement. Moreover, construction managers would request the prediction measurement of performance based on the available data and expected improvement.

Kaplan and Norton (1992) first announced the concept of balanced scorecard (BSC). BSC is a strategic measurement system (Kaplan and Norton, 1996). BSC revolutionized conventional thinking about performance metrics and presented the evaluation of performance framework from four perspectives, which are financial, customer, internal business process and learning and growth. The new concepts allowed organizations to track financial outcomes while monitoring improvement in building the capabilities required for growth (Propa et al., 2015). BSC provides the facility to companies to improve their vision and strategy and then turn them into action; it translates the strategic objectives of an organization into a consistent set of performance measures. KPIs are one of the tools used for the performance measurement and provide results of the projects and organization for the BSC. The previous research works study the required KPIs for only one or two selected aspects such as industrial or building projects in any organization or for certain size of organization without considering the other different aspects’ parameters or provide overall assessment model based on the selected KPIs.

Neely and Adams (2000) identified seven main reasons why performance measurement is now a management priority. These reasons are: dynamic nature of work, strong competition, specific improvement initiatives, more enlightened consumers, local and international quality awards, changing organizational roles and information technology. Neely et al. (2002) suggested definitions for performance measurement, a performance measure and a performance measurement system and they were defined as follows:

1. Performance measurement: the process of quantifying the efficiency and effectiveness of action.

2. Performance measure: a metric used to quantify the efficiency and/or effectiveness of an action.
Performance measurement system: the set of metrics used to quantify both the efficiency and effectiveness of an action.

As a reaction to increased client’s criticism about the construction companies’ performance, UK Government established a “Construction Task Force” to assess and report on the efficiency and quality of the UK construction industry from a customer’s perspective. In 1998, Sir John Egan issued the committee finding, titled “Rethinking construction.” A main result of that task force was the urgent need to set objectives for improving UK construction industry performance. Declaring that “To drive dramatic performance improvement the Task Force believes that the construction industry should set itself clear measurable objectives, and then give them focuses by adopting quantified targets, milestones and performance indicators” (page 14). In order to match with the previous statement, KPIs are chosen as a simple and effective measurement system that helps an organization to explain and manage progress for a customer’s line up with the vision and mission of the organization (MAMPU, 2005).

Effective performance management system (PMS) needs to include KPIs (preferable to be limited), capable of providing an integrated and comprehensive view of a company’s performance. This is important to minimize information overload, to avoid confusion for their potential users and to give a clear picture of the critical organizational competitive factors.

Measuring performance is not an easy task and typically faces resistance from different operation levels within an organization. Pollanen (2005) specified the following four areas of difficulties which preclude PMS acceptance and implementation:

(1) institutional – resistance to transparency;
(2) technical – absence of specification and standard;
(3) financial – significant investment in resource and time; and
(4) pragmatic – inadequate convenience and reliability.

In recent years, strong competitive and significant changes in the construction industry are pushing executive managers to continuously improve the performance of their companies. There is no standard method that has been implemented for measuring performance that is also capable of comparing and evaluating the relative efficiency and effectiveness between competitors. Many research efforts have concentrated on measuring the performance of a project. The uniqueness of each construction project creates difficulty to propose a standard framework to measure the performance of different projects.

3. Research methodology

Many construction companies in Egypt whether large or small are lacking an effective method to evaluate their strengths and weaknesses measure their performance and propose future improvements. Due to the missing standard method of assessing project performance in Egypt and lack of historical performance data, there is a major lack of benchmark for these measurements. This research aims at identifying and prioritizing the main KPIs for the construction projects in Egypt, taking into consideration the other two aspects which are the size of the organization and the type of project. Based on these results, the research introduces different mathematical models to measure the overall project performance indicators. This paper considers three aspects in measuring project performance (location, size and type) which has not been addressed before in literature. The research also proposes the use of POPI assessment models to measure the performance of different construction projects. To achieve the objective of this research, the following methodology was followed:

- Investigated how Egyptian construction companies measured and assessed their performance and what adopted methods were being used by other countries.
Developed a group of KPIs to assess the performance of construction companies in Egypt. These sets of indicators were customized to suit the different types of construction projects.

Conducted a questionnaire and interviewed many experts in order to identify the indicators and their relative weights to rank the indicators according to the project type and size.

Conducted EFA using principal component analysis (PCA) through apply five tests including correlation matrix scan, multicollinearity and singularity check, anti-image correlation matrix scan, Kaiser–Meyer–Olken (KMO) measure of sampling adequacy (MSA) and Bartlett’s test in order to restructure the data by reducing the number of variables that can represent the linear relationships between sets of indicators and finally grouping the sets of performance indicators to principal factors for measuring construction project’s performance of different perspectives.

Used the questionnaire results after conducting EFA on it to rank the Egyptian construction project KPIs.

Use the first seven KPIs for every type and size of the projects based on relative importance index (RII) to generate POPI.

4. Identification of indicators
The application of KPIs has received noticeable attention after the issuance of “Rethinking Construction” report in 1998. Chan and Chan (2004) proposed using eight KPIs in measuring the construction projects performance. Westerveld (2003) used only four KPIs for measuring the performance. According to Beatham et al. (2004), 15 KPIs should be used. Costa et al. (2006) proposed that the measure of the construction project performance should be using up to 30 KPIs. Leu and Lin (2008) used only eight KPIs. Yeung et al. (2009) presented using 17 KPIs to measure the performance of the construction project. According to all these previous studies, the total different KPIs used before were 56 indicators. Yeung et al. (2012) studied these 56 indicators and found that many of them have the same technical meaning or were used to measure the same type of performance. His findings summarized the previous work in KPIs in using 28 different indicators. In this research, most of these 28 KPIs were used, plus the new KPIs proposed from construction experts based on the interviews with them. In total, 35 indicators were proposed and presented for evaluation through the survey.

5. Design of the questionnaires
The survey questionnaire was titled “Questionnaire to develop a series of KPIs to measure performance of construction projects in Egypt.” The questionnaire included questions that required the respondent to elicit thoughts of the importance of KPIs and to rank them after considering the size of the organization and type of the projects. The questionnaire started with an introduction that gave a summary of the purpose of the survey and provided definitions of terminology used inside the questionnaire. The main structure of the questionnaire contained three sections; the first section asked about the respondent’s personal information. The second section collected information about the organization and the current status of KPIs’ implementation in the organization. Finally, the third section contained 35 KPIs which were required to be evaluated and identified for their importance index based upon the respondents’ previous experience.

A pilot questionnaire is the most recommended method to ensure that the questionnaire is clear and its terminologies are correctly understood by all participants.
It can also be used to clarify ambiguities in the meaning of questions, confirm the length of time required to complete the questionnaire and to exclude any question that does not give any added value data. The pilot study was carried out amongst six of construction experts. Individual meetings were held with all of the experts to measure the clarity and validity of the proposed KPIs to measure construction companies’ performance in Egypt and all similar developing countries. Also, to obtain their recommendation regarding the suitable numbers of KPIs to create the overall performance assessment models. The results of these meetings were considered and the final copy of the questionnaire was issued.

6. Questionnaire findings and results analysis
The 230 questionnaires received from 71 Egyptian construction companies represent 74 percent of the total copies distributed. The respondents were from companies of different scales and working in different types of projects. The results obtained from the questionnaire analysis were as follows.

6.1 For positions of the respondent
The respondents were of several job positions and each contributed in a percentage of the total feedback obtained from the data analysis as follows: directors/general managers, senior managers/area managers, manager/team leader, seniors and engineers were 10.9, 17, 32.6, 21.6 and 17.8 percent, respectively. More than 60 percent of the respondents held a position higher than senior engineer.

6.2 Respondents years of experience
Figure 1 illustrates that more than 66 percent of the respondents have experience of more than 10 years which reflect the high level of construction experiences involved and consequently the level of trust in these respondents’ feedbacks.

6.3 Organization role in field of construction
In total, 83 percent of the respondents were working for contractors, 8.7 percent for owners and nearly the same percentage, 8.3 percent for consultants or project management organization.

Table I demonstrates that about 75 percent of the respondents implement or will implement KPIs in their organization. This percentage seems very high and not possible in Egypt with the reason being the misuse of the term “KPI” in different ways within a departmental level or a project level. It was found that the equipment department in one of the construction companies in Egypt utilized some measurements of the equipment under the title “Equipment KPIs.” Another company at a project level measures the project’s...
productivity for different activities under the title “Construction KPIs.” The use of the term “KPI” in different ways as mentioned earlier affects the accuracy of responses to the question about the number of years, an organization has used KPIs. Table II illustrates that 81.6 percent of the companies which use KPIs have used them for a period of more than three years.

Table III indicates that 66 percent of the respondents come from large organizations. This is likely the result of mainly two factors which are: the respondents that have experience in using KPIs were mainly from large organizations and most of the questionnaire’s non-respondents were from small and medium size organizations due to non-use or lack of the experience in it.

Based on the respondent’s feedback, the ranking of construction KPIs in Egypt was estimated by assessing the mean score (MS) and the RII. A five-point scale was used in calculating the MS and it was converted to RII using the relative index ranking method. Table IV summarizes the results of the third part of the questionnaire and lists the ranking of all KPIs without splitting results based on the organization size and project types. Another series of tables were created to get the KPIs ranking while considering the other two parameters (project types and organization size).

Similar rankings were generated for every type of projects while considering the third aspect of the measurement which is the organization size. The first ten indicators for every ranking are presented in Tables V–VII.

The size of organizations were found in commercial and residential projects while only the medium and large organizations were found in industrial, infrastructure, and heavy civil projects. Due to the nature of these types of projects, they require at least medium size subcontractors. But still, small subcontractors can be found working in these types of projects under the full supervision and control of the main contractors, but they are generally neither able nor interested in implementing KPIs.

<table>
<thead>
<tr>
<th>Usage of KPI</th>
<th>No. of organization</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Yes</td>
<td>163</td>
<td>70.9</td>
</tr>
<tr>
<td>2 No</td>
<td>36</td>
<td>15.7</td>
</tr>
<tr>
<td>3 About to use</td>
<td>10</td>
<td>4.3</td>
</tr>
<tr>
<td>4 Not sure</td>
<td>21</td>
<td>9.1</td>
</tr>
</tbody>
</table>

Table I. Use of KPIs in the Organization

<table>
<thead>
<tr>
<th>No. of years-use KPI</th>
<th>No.</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Less than 2 years</td>
<td>30</td>
<td>18.4</td>
</tr>
<tr>
<td>2 3–5 years</td>
<td>80</td>
<td>49.1</td>
</tr>
<tr>
<td>3 More than 5 years</td>
<td>53</td>
<td>32.5</td>
</tr>
</tbody>
</table>

Table II. No. of years organizations use KPIs

<table>
<thead>
<tr>
<th>Yearly turnover ($ m)</th>
<th>No.</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Less than 10</td>
<td>15</td>
<td>6.5</td>
</tr>
<tr>
<td>2 10–24</td>
<td>15</td>
<td>6.5</td>
</tr>
<tr>
<td>3 25–50</td>
<td>17</td>
<td>7.4</td>
</tr>
<tr>
<td>4 51–100</td>
<td>31</td>
<td>13.5</td>
</tr>
<tr>
<td>5 101–250</td>
<td>30</td>
<td>13.0</td>
</tr>
<tr>
<td>6 More than 250</td>
<td>122</td>
<td>53.0</td>
</tr>
</tbody>
</table>

Table III. Average yearly organization turnover
7. Exploratory factor analysis
PCA is a mathematical procedure that transforms a large number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. It combines input variables in a specific way after dropping the least important variables.
while still retaining the most valuable parts of all of the variables. Dogbegah et al. (2011) used PCA in the analysis of project management competencies for the Ghanaian construction industry. PCA is used by Francom and El Asmar (2014) to investigate the effect of using building information modeling on project changes.

To determine whether a set of data are appropriate for PCA, EFA method is recommended for identifying the inter-relationships between sets of variables. EFA is performed to restructure the data by reducing the number of variables that can be used to represent linear relationships between groups of perceived indicators (Nimalathasan, 2009). This approach is often called “data reduction” or “dimension reduction” technique.

To determine whether a set of data are appropriate for PCA, two items should be checked. The first being sample size and the second is the strength of the inter-correlation between the independent indicators. Correlation coefficients are fluctuant from sample to sample. Factors resulting from small data sets do not generalize as well as those from larger samples. Guadagnoli and Velicer (1988) stated that sample size is a critical factor to get reliable factor solutions. In total, 100–200 responses were the minimum sample size according to their studies. Tabachnick and Fidell (2007) proposed sample sizes as follows: 50 observations are very poor, 100 are poor, 200 are fair, 300 are good and 500 are very good. Tinsley and Kass (1979) recommended having five to ten observations per variable up to 300. The results of the survey which are 230 respondents could be considered as a good sample and having 6.6 observations per indicator. The second issue in EFA related to the strength of the inter-correlation between the variables by a correlation matrix (R-matrix) represents Pearson correlation coefficient between all pairs of indicators.

At the first iteration, factor analysis was conducted for all the 35 KPIs. The correlation matrix was scanned to test the relationships between performance indicators. It was

<table>
<thead>
<tr>
<th>KPI/Medium organization</th>
<th>KPI/Large organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost performance</td>
<td>Cost performance</td>
</tr>
<tr>
<td>Quality performance</td>
<td>Quality performance</td>
</tr>
<tr>
<td>Time performance</td>
<td>Time performance</td>
</tr>
<tr>
<td>Client’s satisfaction</td>
<td>Client’s satisfaction</td>
</tr>
<tr>
<td>Safety performance</td>
<td>Safety performance</td>
</tr>
<tr>
<td>Productivity</td>
<td>Productivity</td>
</tr>
<tr>
<td>Effiiciency of cash management</td>
<td>Effectiveness of cost control</td>
</tr>
<tr>
<td>Effectiveness of cost control</td>
<td>Procurement management</td>
</tr>
<tr>
<td>Billing management</td>
<td>Top management commitment</td>
</tr>
<tr>
<td>Effectiveness of material management</td>
<td>Effectiveness of material management</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>KPI/Industrial projects performance</th>
<th>KPI/Large organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost performance</td>
<td>Cost performance</td>
</tr>
<tr>
<td>Quality performance</td>
<td>Quality performance</td>
</tr>
<tr>
<td>Time performance</td>
<td>Time performance</td>
</tr>
<tr>
<td>Client’s satisfaction</td>
<td>Client’s satisfaction</td>
</tr>
<tr>
<td>Safety performance</td>
<td>Safety performance</td>
</tr>
<tr>
<td>Productivity</td>
<td>Productivity</td>
</tr>
<tr>
<td>Effectiveness of cost control</td>
<td>Effectiveness Of cost control</td>
</tr>
<tr>
<td>Procurement management</td>
<td>Procurement management</td>
</tr>
<tr>
<td>Top management commitment</td>
<td>Top management commitment</td>
</tr>
<tr>
<td>Billing management</td>
<td>Effectiveness of material management</td>
</tr>
</tbody>
</table>

Table VI. First ten KPIs for industrial projects

Table VII. First ten KPIs for heavy civil projects
scanned to look for correlation coefficients that are less than 0.3 for all indicators. There were four indicators that have a correlation coefficient of less than 0.3 within all indicators. These indicators were: scope of rework, long-term business relationships, litigation occurrence and magnitude and change management. There was no correlation coefficient that is greater than 0.9. Six iterations were required to ensure that no indicators had a correlation coefficient less than 0.3 within all indicators and no correlation coefficient with a value greater than 0.9 was found. Table VIII summarizes the EFA six iterations’ results and indicates which indicators were removed after every iteration. Figure 2 shows the sample results of correlation matrix, iteration No. 5 in which the reduction of paperwork performance indicator was removed based on the test result and the test limits.

### 7.1 Multicollinearity and singularity

Multicollinearity and singularity are the types of mathematical problems arising when variables are extremely correlated. Multicollinearity is a case that represents when two or more indicators that are related very closely or mathematically, and with a correlation coefficient equal to or greater than 0.90. Singularity occurs when one variable is actually a grouping of two other independent variables. Mathematically, correlation coefficients are equals to 1 or \(-1\). Variables which have a correlation coefficient equal to 0.9 or more will

<table>
<thead>
<tr>
<th>Iteration No.</th>
<th>Input KPIs</th>
<th>Output KPIs</th>
<th>Indicators removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35</td>
<td>31</td>
<td>SR – LTR – LOM – CM</td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>28</td>
<td>HWR – DOM – CA</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>26</td>
<td>TR – CAM</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
<td>24</td>
<td>EC – ST</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>23</td>
<td>RP</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

Table VIII. Summary of the EFA iteration results

Figure 2. Sample of results: correlation matrix for iteration No. 5
cause inaccuracy in factor analysis results. Two methods are used for evaluating multicollinearity or singularity:

1. Scanning the correlation matrix for correlation coefficients equals to or greater than 0.90. These strong correlated variables should be eliminated. This test was conducted at the first step in preliminary analysis.

2. Checking the determinant of the correlation matrix. They may exist if the determinant of the correlation matrix is less than $1 \times 10^{-5}$.

At the sixth iteration, factor analysis was conducted on the 23 remaining KPIs. The correlation matrix was scanned very well; no correlation coefficient was greater than 0.90 within any another indicator. The determinant of the matrix was equal to $1.48 \times 10^{-4}$ which is greater than $1 \times 10^{-5}$. These results indicate that there are quite well relationships among all remaining indicators.

7.2 Anti-image correlation matrix scan
Another test was conducted to examine the anti-image correlation matrix that contains measures of sampling adequacy (MSA) for each variable along the diagonal and partial correlation on the off-diagonals (Beaumont, 2012). The anti-image correlation matrix for the sixth run conducted for 23 indicators showed diagonal elements greater than 0.5 for all indicators and all off-diagonal elements were close to 0. These results confirm that EFA is feasible.

7.3 Kaiser–Meyer–Olkin
The KMO is another method to evaluate the degree of inter-correlations between indicators (Metts et al., 2006). The resulting value close to 1 indicates that patterns of correlations are relatively dense and factor analysis will result in discrete and reliable factors. Kaiser suggests accepting values greater than 0.5. KMO tests for the data after considering all the previous tests till the sixth iteration are equal to 0.828. It falls into the great range according to test results classification. So, there is confidence that all KMO values of individual variables in anti-image correlation matrix for the sixth iteration are better than the acceptable limit of 0.5.

7.4 Bartlett’s test
Bartlett’s test is usually conducted to check the suitability, all the diagonal terms should be equal to 1 and all off-diagonal terms are equal to 0. If the test result has a significance value less than 0.05, it indicates that the test is significant which means that the correlation matrix is not an identity matrix. For the indicators set in the sixth iteration, the Bartlett’s test of degree of freedom (df = 253) is equal to 1,944 and significant value ($p$-value) = 0.001 which is less than 0.001.

7.5 Tests results and discussion
Previous tests’ results showed that the correlation between indicators is sufficiently large and suitable for EFA. Table IX summarizes the results obtained through different EFA tests.

<table>
<thead>
<tr>
<th>Kaiser–Meyer–Olkin measure of sampling adequacy</th>
<th>0.828</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bartlett’s test</strong></td>
<td></td>
</tr>
<tr>
<td>Approx. $\chi^2$</td>
<td>1,944</td>
</tr>
<tr>
<td>df</td>
<td>253</td>
</tr>
<tr>
<td>Sig.</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table IX. Results of EFA tests
Factor extraction includes determining the minimum number of factors able to present the inter-relations among the set of indicators. Kaiser’s criterion or “eigenvalue rule” is a common technique for determining the number of retaining factors. Kaiser advised retaining all factors of eigenvalues greater than 1 based on that the eigenvalues represent the amount of variation explained by a factor and that an eigenvalue of 1 represents a substantial amount of variation. The first six factors as shown in Table X satisfy the previous criteria and represent the 61.19 percent of the total variance.

8. Overall assessment models
Generating the POPI is the main target for this research. Cheung et al. (2004) introduced a web-based project performance follow up and control system. Eight performances measuring categories were defined in this system. Since 2007, Scottish construction center has chosen nine KPIs to measure the performance. Starting in 2005, Danish contractors bidding for jobs were requested to present and compare competence in 14 KPIs suggested by the benchmark center for the Danish construction sector. Starting from 2010, a new group of only seven KPIs was proposed to measure and evaluate performance for Danish contractors. Yeung et al. (2007) used the Delphi survey technique to develop a model to measure the performance of partnering projects in Hong Kong by using seven KPIs. Yeung et al. (2009) applied the same technique again to formulate a model to assess the success of relationship-based construction projects in Australia by using eight KPIs selected according to survey results. Khooravi and Afshari (2011) proposed project success index based on measuring five KPIs. Chan and Hiap (2012) generated overall performance equations for target cost contracts by using seven KPIs. Ten KPIs were used by Yeung et al. (2012) to develop a benchmarking model for construction projects in Hong Kong. El Asmar and Hanna (2013) developed a framework to measure and evaluate construction project performance by using five main KPIs. Ofori-Kuragu et al. (2016) proposed using nine KPIs for measuring the performance of Ghanaian contractors.

Based on previous works presented above, five to nine KPIs were normally used in generating the project performance index. Five KPIs seemed low and insufficient to measure the project performance from different perspectives. Six to eight KPIs were more suitable. In this research, seven KPIs were used to generate performance evaluation models based on the results of the survey while considering the project type and size of organization. The proposed models can be used to evaluate the relative success of projects throughout the organization, indicate areas of strengths and weaknesses in performance, and finally, establish the company’s benchmarking. Table XI represents a sample of calculations for

<table>
<thead>
<tr>
<th>FAC 1</th>
<th>FAC 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time performance</td>
<td>Innovation and improvement</td>
</tr>
<tr>
<td>Cost performance</td>
<td>Top management commitment</td>
</tr>
<tr>
<td>Quality performance</td>
<td>Effectiveness of waste management</td>
</tr>
<tr>
<td>Profit and financial objectives</td>
<td>FAC 4</td>
</tr>
<tr>
<td>Productivity</td>
<td>Effectiveness of planning</td>
</tr>
<tr>
<td>Client’s satisfaction</td>
<td>Effectiveness of cost control</td>
</tr>
<tr>
<td>Billing management</td>
<td>Efficiency of cash management</td>
</tr>
<tr>
<td>FAC 2</td>
<td>FAC 5</td>
</tr>
<tr>
<td>Safety performance</td>
<td>Effectiveness of material management</td>
</tr>
<tr>
<td>Environmental issues</td>
<td>Supplier performance</td>
</tr>
<tr>
<td>Pollution occurrence</td>
<td>Procurement management</td>
</tr>
<tr>
<td>Effectiveness of risk management</td>
<td>FAC 6</td>
</tr>
<tr>
<td>Defects</td>
<td>Defects cost</td>
</tr>
<tr>
<td>Defects time</td>
<td></td>
</tr>
</tbody>
</table>

Table X.
Principal factors to measure the construction project’s performance
indicators relative weights that are used to formulate the general performance equation for commercial and residential projects in medium size organizations.

Based on the above, the overall performance indicator for commercial and residential projects in a medium size organization is:

Project overall performance indicator (POPI) = 16.4% CP + 15.1% TP + 14.3% PR + 14% QP + 13.6% SP + 13.4% CS + 13.2% ECC

Similarly, the following equations were formed for POPI:

1. Commercial and residential projects, small size organization:
   \[ \text{POPI} = 16.9\% \text{ CP} + 14.9\% \text{ PR} + 14.5\% \text{ CS} + 14.2\% \text{ QP} + 13.7\% \text{ TP} + 13.4\% \text{ LTR} + 13.4\% \text{ HTR} \]

2. Commercial and residential projects, medium size organization:
   \[ \text{POPI} = 16.4\% \text{ CP} + 15.1\% \text{ TP} + 14.3\% \text{ PR} + 14\% \text{ QP} + 13.6\% \text{ SP} + 13.4\% \text{ CS} + 13.2\% \text{ ECC} \]

3. Commercial and residential projects, large size organization:
   \[ \text{POPI} = 15.8\% \text{ CP} + 14.7\% \text{ QP} + 14.4\% \text{ TP} + 14.1\% \text{ SP} + 13.8\% \text{ PR} + 13.6\% \text{ CS} + 13.5\% \text{ ECC} \]

4. Industrial projects, medium size organization:
   \[ \text{POPI} = 16.2\% \text{ CP} + 15.3\% \text{ QP} + 14.4\% \text{ TP} + 14.0\% \text{ SP} + 13.5\% \text{ CS} + 13.3\% \text{ PR} + 13.3\% \text{ ECC} \]

5. Industrial projects, large size organization:
   \[ \text{POPI} = 15.5\% \text{ CP} + 14.8\% \text{ QP} + 14.6\% \text{ SP} + 14.2\% \text{ TP} + 13.9\% \text{ PR} + 13.6\% \text{ CS} + 13.4\% \text{ ECC} \]

6. Heavy civil projects, medium size organization:
   \[ \text{POPI} = 15.4\% \text{ CP} + 14.8\% \text{ QP} + 14.6\% \text{ TP} + 14.4\% \text{ PR} + 14.1\% \text{ CS} + 13.6\% \text{ SP} + 13.1\% \text{ ECC} \]

<table>
<thead>
<tr>
<th>Sr.</th>
<th>Indicators</th>
<th>Av.</th>
<th>II (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cost performance</td>
<td>4.41</td>
<td>16.4</td>
</tr>
<tr>
<td>2</td>
<td>Time performance</td>
<td>4.06</td>
<td>15.1</td>
</tr>
<tr>
<td>3</td>
<td>Productivity</td>
<td>3.82</td>
<td>14.3</td>
</tr>
<tr>
<td>4</td>
<td>Quality performance</td>
<td>3.77</td>
<td>14.0</td>
</tr>
<tr>
<td>5</td>
<td>Safety performance</td>
<td>3.65</td>
<td>13.6</td>
</tr>
<tr>
<td>6</td>
<td>Client’s satisfaction</td>
<td>3.59</td>
<td>13.4</td>
</tr>
<tr>
<td>7</td>
<td>Effectiveness of cost control</td>
<td>3.53</td>
<td>13.2</td>
</tr>
</tbody>
</table>

Table XI: KPIs relative index for medium size, commercial projects
(7) Heavy civil projects, large size organization:

\[
\text{POPI} = 15.7\% \text{CP} + 15.0\% \text{QP} + 14.6\% \text{TP} + 14.1\% \text{CS} + 13.8\% \text{SP} + 13.4\% \text{PR} + 13.3\% \text{ECC}
\]

Based on the above equations and previous results, the following can be concluded:

- Five common performance indicators are represented in the equations of overall performance indicators for all types of projects and all sizes of organizations. They are: cost performance, quality, productivity, client satisfaction and time performance but with different ranks and weights.
- The weights for the first seven indicators are very close and the maximum difference in weight is about 4 percent which reflects nearly the same importance for these indicators. The small difference in results comes from the subjectivity of the evaluation.

9. Validation of the proposed models

A set of KPIs models have been developed in the previous section to measure the overall performance of different types and sizes of construction projects in Egypt. Three case studies were chosen to check and validate these models. The data were collected and adjusted to be in the same financial manner by using net present value technique and average yearly interest rate in Egypt up to 2016.

The scope of analysis for each case covers the KPIs in the project overall performance model according to the project’s category. The datum of KPI levels to measure the status of the project needs to be defined by the companies’ top management based on the minimum acceptable levels for profitability, safety, quality, etc. Table XII shows the summary information and the results of KPIs of these case studies.

The first project was classified as a heavy civil project starting in 2011. The project was executed by a large company with the following evaluation: under budget, behind schedule, average measurement of KPIs for safety, quality, productivity and client satisfaction. By implementing model No. 7, the project gets POPI equals to 82.5 percent which reflect a project with “very good” performance.

<table>
<thead>
<tr>
<th>Case study</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of the project</td>
<td>Heavy civil</td>
<td>Building</td>
<td>Industrial</td>
</tr>
<tr>
<td>Company size</td>
<td>Large</td>
<td>Large</td>
<td>Large</td>
</tr>
<tr>
<td>Contract type</td>
<td>Unit rate</td>
<td>Unit rate</td>
<td>Unit rate</td>
</tr>
<tr>
<td>Total project value</td>
<td>EGP205m</td>
<td>EGP338m</td>
<td>EGP258m</td>
</tr>
<tr>
<td>Total project duration</td>
<td>25</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>POPI No.</td>
<td>7</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>KPIs results (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost performance</td>
<td>100</td>
<td>36</td>
<td>85</td>
</tr>
<tr>
<td>Time performance</td>
<td>60</td>
<td>75</td>
<td>71</td>
</tr>
<tr>
<td>Productivity</td>
<td>75</td>
<td>60</td>
<td>88</td>
</tr>
<tr>
<td>Quality performance</td>
<td>83</td>
<td>80</td>
<td>82</td>
</tr>
<tr>
<td>Safety performance</td>
<td>85</td>
<td>87</td>
<td>97</td>
</tr>
<tr>
<td>Effectiveness of Cost Control</td>
<td>83</td>
<td>90</td>
<td>96</td>
</tr>
<tr>
<td>Client’s satisfaction</td>
<td>90</td>
<td>65</td>
<td>85</td>
</tr>
<tr>
<td>Overall performance indicators</td>
<td>82.5</td>
<td>70.1</td>
<td>86.2</td>
</tr>
</tbody>
</table>

Table XII. Validation of KPIs models for different type of projects
The second project was classified as a building project starting in 2011. The project was executed over budget, behind schedule, with low productivity rates, average KPIs for safety, quality and client satisfaction. By implementing model No. 5, the project gets POPI equals to 70.1 percent which reflect "medium" performance project (Table XII).

The third project was an industrial project. It had a better performance regarding safety quality, client’s satisfaction and productivity. By using model No. 3 for evaluating project performance, the project achieves 86.2 percent which could be considered as an "excellent" project.

10. Summary and conclusion

Measuring the performance of construction projects during the entire project life cycles is a very important issue. This evaluation is required to identify and measure the project success in a quantitative manner from many perspectives, and not only from a financial and profitability point of views. Performance measurement is a dynamic knowledge area. Indeed, it is considered as the force that drives project management improvement. Also, measuring performance helps an organization to identify the areas requiring attention and improvement, given the chance to take the required corrective actions or mitigation plans. The measuring, evaluation and action are critical for controlling the project.

This paper provides a mathematical assessment model for measuring construction performance within Egypt while considering project type and size of an organization. Three types of projects were considered which were commercial and residential projects, industrial projects and heavy civil projects. The size of an organization was classified as small, medium and large according to the yearly turnover and the number of employees.

A questionnaire survey was conducted to identify the importance index for the chosen 35 KPIs for the construction projects. A ranking system of importance was determined. EFA was used to remove the inter-relationships between a set of indicators and to restructure the data. After implementing all of the steps of the EFA techniques, the number of indicators was reduced to be 23 indicators and grouped in 6 main factors.

Seven indicators with the highest importance indicator were chosen to generate POPI. Seven models were the resultant findings of the survey. The following are the findings and conclusions of this research:

- The cost performance indicator was the most important indicator in all types of construction projects and for all organization sizes.
- Industrial projects record a high importance index for safety, quality and productivity.
- With an increase in the size of an organization, the impotency index for quality, safety and time performance increased which reflected the important role of functional departments such as project control, safety and quality inside the large organization.
- Top management commitment indicator appears in the first ten indicators for large organizations only, in all types of projects which reflect the need in this size of organization to the support and commitment of company’s top management.
- Client satisfaction index is a major indicator in measuring overall projects’ performance.
- Results of factor analysis, six perspectives present the major perspective for evaluation of the project performance with its sub-indicators.
- KPIs could be used by the Egyptian Federation for Construction & Building Contractors to rank the Egyptian subcontractors similar to many developed countries which consider the performance as the main basis for a contractor’s evaluation. It will give a positive impact on using KPIs in all of the Egyptian companies. The implementation of this concept needs a strong system to collect real performance data from all the companies to issue contractors’ classification manual.
The future work in KPIs is based on the previous works for evaluating the performance of the project and creating prediction models to predict the KPIs for the construction project based on the results during the project life cycle and consequently, predicting the project overall performance.

References


MAMPU (2005), “Guidelines for establishing key performance indicators (KPIs) and executing performance measurement in government agencies”, Appendix to the Public Administration Progress Circular No. 2, Malaysian Administrative Modernization and Management Planning Unit (MAMPU), Putrajaya.


Further reading


Corresponding author
Emad Fayez Gaid can be contacted at: Emad.fayez@orascom.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
The impacts of leadership support and coworker support on employee creative behavior

Michel Zaitouni and Mohamed Laid Ouakouak
Department of Business Administration,
Gulf University for Science and Technology, Kuwait City, Kuwait

Abstract

Purpose – Previous research has emphasized how leadership style and collegial relationships can foster creativity in the workplace; in a related sense, understanding how support from leaders and coworkers affect the creativity process is critical too. Therefore, the purpose of this paper is to seek a deeper understanding of how leadership support and coworker support influence employee creativity.

Design/methodology/approach – A quantitative study was conducted among 299 employees working at eight organizations in a non-western country.

Findings – The results of this research revealed that both leadership support and coworker support exert positive influences on employee creativity, moderated by several additional variables. Specifically, employee engagement and tenure both strengthen the relationship of leadership support with employee creativity. Furthermore, individual creativity has a positive impact on job performance.

Research limitations/implications – These findings offer theoretical and practical implications, as well as directions for further research.

Originality/value – This research examined the impacts of leadership support and coworker support on employee creative behavior and individual performance. To better understand the nature of these impacts, the authors introduced two moderators such as employee engagement and tenure. To the best of our knowledge, there are no empirical evidences regarding whether and how leadership support and coworker support interact with employee engagement and tenure to influence employee creative behavior and individual performance.

Keywords Innovation, Creativity, Employee productivity, Leadership

Paper type Research paper

1. Introduction

Employees’ creative behavior is critical for modern organizations (Lin and Wong, 2014; Wong and Ladkin, 2008). Individual creativity constitutes a foundation for innovation, effective work practices and organizational competitiveness (Campo et al., 2014; Tajedini and Trueman, 2012). Generally defined as the generation of novel, useful ideas (Beth and Amabile, 2010), creativity also entails an ability to produce work that is both novel and useful (McLean, 2005). Because creativity drives innovation, growth and competitiveness (Chang et al., 2011; Gong, Kim, Zhu and Lee, 2013; Gong, Zhou and Chang, 2013), most organizations invest heavily in and search for effective ways to encourage employees’ creativity (Liu et al., 2012), especially when they confront competitive, dynamic business environments. For example, service providers seek to foster employee creativity as a means to deliver high-quality services (Kim and Lee, 2013; Wang et al., 2014; Hon and Lui, 2016). In such service industries, employees can add substantial value if they are able to generate, exploit and implement creative ideas (Lepak and Snell, 2002). Accordingly, creativity constitutes a key driver of service companies’ success and competitiveness (e.g. Campo et al., 2014; Lim, 2016; Mittal and Dhar, 2016).

In extensive research dedicated to understanding the dynamics of creativity, by identifying its drivers (Alencar and Bruno-Faria, 1997; Zhou and George, 2001), various personal and contextual factors have emerged as beneficial for employee creative behavior, including motivation (Zhang and Bartol, 2010), a learning orientation (Gong et al., 2009),
the organizational culture (Lin et al., 2013), transformational leadership (Shin and Zhou, 2003), coworker support (Zhou and George, 2001) and experience (Tripsas, 1997). According to Mumford (2000), an appropriate context that supports creativity effectively (Amabile and Conti, 1999) enables organizations to generate, exploit, renew and implement new and valuable ideas to improve their performance. Other researchers concur that the workplace atmosphere influences employees’ creative behavior (Wang et al., 2014). Therefore, successful, innovative organizations likely are those that support initiative taking, reward creative employees, encourage cooperative work, do not blame individuals for mistakes and encourage members to take risks.

As part of an effective, creative and innovative context, organizations should build both leadership support and coworker support. We emphasize these two factors in this study for several reasons. A growing body of research suggests that leaders are key to supporting creativity, because by displaying positive, supportive leadership behaviors, they encourage followers to try being more creative. They can motivate employees to generate new ideas by rewarding, appreciating and recognizing their efforts. Rafferty and Griffin (2004) have defined supportive leadership as leaders’ concerns for their followers’ needs and well-being. Therefore, leadership behaviors likely are key antecedents of creative behavior (Bledow et al., 2009; Shalley and Gilson, 2004). In addition, coworker support or “the extent to which employees believe their coworkers are willing to provide them with work-related assistance to aid in the execution of their service-based duties” (Susskind et al., 2003, p. 181), enhances organizational engagement, job satisfaction, worker well-being and retention (e.g. Sloan, 2012; Chiaburu and Harrison, 2008; Iverson, 1999). All these benefits in turn should enable employees to become more productive, creative and innovative. According to De Dreu (2006), poor coworker relationships, such as those characterized by conflict, frustration or anger, instead hinder individual creativity.

Although various studies have examined the influences of leadership and coworker support on employee creativity, no empirical, detailed insights are available regarding whether and how they do so. Further studies are needed to uncover the nature of these relationships; in particular, we seek to understand how leadership support and coworker support influence employee creativity by introducing two moderating variables – employee engagement and tenure – that may help clarify these links. In so doing we aim to contribute to extant literature in three main ways. First, we extend debates about whether and how leadership support and coworker support foster employee creativity in the workplace. Second, our study introduces two moderating variables, employee engagement and tenure, in an initial attempt to explain the relationships of leadership support, coworker support and employee creativity according to a moderation approach. Third, though some studies consider the association between creativity and firm performance, few examine this direct relationship (Gong et al., 2009; Zhang and Bartol, 2010). As Chang and Teng (2017) argue, the limited research that has tested this relationship is focused on the hospitality industry, and empirical evidence is still lacking. We therefore propose and test the notion that when people can generate new, useful ideas, they achieve better work performance, such that employee creativity should have a positive impact on job performance.

In our theoretical model, we elaborate on the influences of leadership support and coworker support on employee creativity, as well as the influence of employee creativity on job performance. Moreover, our model includes employee engagement and tenure as moderating variables, so we can test their influences on the central relationships, as we detail in Figure 1.

2. Theory and hypothesis development

2.1 Coworker support and individual creativity

Creating perceived support in today’s dynamic business environment is key driver to foster creativity in many organizations (Ibrahim et al., 2016). Therefore, coworker support can
encourage knowledge sharing and the generation of new ideas, suggesting its relevance for creativity and innovative behavior (Lee and Choi, 2003; Hayton, 2005; Hsu et al., 2007). When an employee encounters a difficult situation or a novel task, coworker support represents a critical type of assistance, in the form of shared knowledge and expertise (Scott and Bruce, 1994). Interactions among coworkers facilitate the creation of new ways of doing things; coworkers also can function as sounding boards for creative ideas (Farr and Ford, 1990; Woodman et al., 1993; Zhou and George, 2001). Because employees work within the same environment, they also may be expected to share creative ideas, which requires sufficient mutual trust to allow the parties to feel psychologically safe with regard to exchanging knowledge and ideas openly. Such a setting creates a good opportunity for novel ideas to be generated (Madjar, 2005).

Research on individual creativity supports arguments derived from studies of coworker support, in that a supportive work environment encourages employees to exchange knowledge and expertise, actively and constructively, which grants them the freedom to operate creatively (Prieto and Perez-Santana, 2014). In other words, strong support from coworkers increases employees’ sense that creativity is an effective option, so coworker support gets channeled into creative activities (Zhou and George, 2001). We thus propose:

\textbf{H1.} Coworker support relates positively to individual creativity.

Employee engagement also should strengthen this effect of coworker support on employee creativity. According to Kahn (1990), who defines employee engagement as “harnessing of organizational members’ selves to their work roles,” highly engaged employees are more likely to generate positive work-related outcomes. Alfes \textit{et al.} (2013) contend that employees who are affectively and socially engaged are willing to discuss work improvements with coworkers, and Madjar (2005) suggests that coworker support for creative ideas results in higher levels of engagement among employees. Engaged employees feel safe in their work environment, which is characterized by trust-based interpersonal relationships and supportiveness, so they can generate new ideas (Kahn, 1990). Accordingly, we propose and test the following prediction:

\textbf{H2.} The positive impact of coworker support on individual creativity increases with the degree of employee engagement.

Tenure also might strengthen this relationship. Weisberg (1999) posits that tenure is a key element for creative success, because employees must be familiar enough with their work to be able to perform creatively. That is, for employees to be creative, they must have some

![Figure 1. Research model](image-url)
experience and knowledge about the tasks they perform. Branzei and Vertinsky (2006) assert that senior employees can use their extensive knowledge to engage in creative processes. In this sense, employees who acquire task-relevant knowledge and expertise from senior coworkers may be better able to find new ways to do things (Woodman et al., 1993). When senior coworkers share their expertise and provide encouragement to colleagues, those employees can more readily exert effort to generate new ideas and successfully implement them (Farr and Ford, 1990; Zhou and George, 2001). Researchers investigating tenure-creativity relationships have clearly emphasized the importance of organizational tenure as a key facet of experience that has an impact on individual creativity (Uppal et al., 2014; Uppal and Mishra, 2014). Thus, a work environment characterized by the existence of helpful senior coworkers should increase the extent to which employees perceive their organization as a supportive environment that allows and accepts new ways of doing things. Formally:

$H3$. The positive impact of coworker support on individual creativity increases with the degree of tenure.

2.2 Leadership support and individual creativity

To signal the value they place on creativity, organizations might rely on supervisor encouragement or leadership support, which refers to the extent to which employees perceive that their leaders are actively engaged in the creative process (Zhang and Bartol, 2010). When employees perceive that their contributions are valued, supported, and rewarded by supervisors, they likely exhibit creative behavior (Scott and Bruce, 1994; Zhou and Shalley, 2003; Kossek et al., 2011). For example, Amabile and Gryskiewicz (1989) find that employees believe that their creative behavior is effective only if their supervisors welcome feedback and accept suggestions for improvement. Similarly, several researchers suggest that creative supervisors can foster employees’ creativity through their novel and new ways in defining problems and in setting creative goals for them (Huang et al., 2016; Mainemelis et al., 2015).

Similarly, Cheung and Wong (2011) determine that supervisors who encourage and maintain cooperative interpersonal relationships can influence employees’ engagement in the creative process, by enhancing their self-efficacy. Tierney and Farmer (2004) contend that employees with higher levels of self-efficacy tend to be more creative if they are aware of their supervisors’ expectations of their creative behaviors. Employees thus engage in creative behavior as long as supervisors’ expectations for creativity are clear (Carmeli and Schaubroeck, 2007). Thus, leadership support should relate significantly to employees’ creative behavior (Shalley and Gilson, 2004).

The outcomes of leadership support also may vary with individual personality traits though. For example, George and Zhou (2001) contend that conscientiousness, one of the big five personality traits (Costa and McCrae, 1992), involves the extent to which people are creative and able to interact with close monitoring, so conscientious employees who are closely controlled and not supported by their coworkers likely exhibit low levels of creativity. Having proposed these assumptions, we posit:

$H4$. Leadership support relates positively to individual creativity.

Employee engagement, in turn, might strengthen the relationship between leadership support for creativity and employee creativity. Eisenberger et al. (2002) find that employees who perceive their supervisor as supportive and caring about their well-being are more likely to increase their level of engagement with the organization. This relationship between supervisor support and employee engagement can be viewed in light of social exchange theory (Shore et al., 2006): Employees who perceive their supervisor as supportive
reciprocate by generating a higher level of engagement (Blau, 1964). From a different perspective, Maslach et al. (2001) posit that a lack of support from supervisors might lead to employee disengagement instead of employee engagement in the creative process. Strong supervisor encouragement for creativity should direct employees’ attention toward the generation of novel ideas and thus increase their engagement in the creative process. In other words, employees’ perceptions of useful supervisor support for creativity increases their level of engagement, thereby channeling their attention toward creative activities. We propose in turn:

H5. The positive impact of leadership support on individual creativity increases with the degree of employee engagement.

Part of leaders’ success in encouraging creative ideas is rooted in their ability to share their experience and knowledge with employees, so seniority may moderate the relationship of leadership support with employee creativity. According to Vroom and Jago (2007), creative supervisors interact with their employees and motivate them to work collaboratively to achieve new, useful outcomes. De Jong and Den Hartog (2007) find that the leader’s role is essential for creating a climate for creativity by stimulating knowledge diffusion. Similarly, strong ties between experienced supervisors and employees facilitate exchanges of creative information and support, increasing the likelihood that supervisors and employees think in broader ways and channel their thoughts toward common creative goals (Erdogan et al., 2004). Thus, we expect leadership support and tenure to interact to influence employees’ tendencies to generate new and novel ideas, and we propose:

H6. The positive impact of leadership support on individual creativity increases with the degree of tenure.

2.3 Employee creativity and job performance

Employee creativity drives individual performance (Woodman et al., 1993; Mumford, 2000; Gilson et al., 2005); as an integral part of an overall organizational strategy, it also should relate closely to job performance (Amabile et al., 2004; Gong et al., 2009; Hon, 2013) and thus to organizational performance (e.g. Jiménez-Jiménez and Sanz-Valle, 2011; Camisón and Villar-López, 2014). For example, creativity enhances the delivery of high-quality services in the hospitality industry (e.g. Amabile et al., 2005; Oldham and Cummings, 1996). Grewal et al. (2009) accordingly posit that employees who solve problems effectively and creatively exhibit superior performance, and Tierney and Farmer (2004) contend that employees’ use of creative task information increases their creative capacity, with positive attributions for job performance. Drawing on these arguments, we anticipate that employees achieve superior performance as their engagement in the creative process increases. Employee creativity also should enable the organization to be more competitive and build sustainable competitive advantages (Baer and Oldham, 2006; Coelho et al., 2011). In their research, Agnihotri et al. (2013) uncover a positive association between creativity and sales professionals’ job performance. Thus, we hypothesize:

H7. Individual creativity relates positively to job performance.

We present the relationships among the studied constructs in Figure 1.

3. Methodology
3.1 Sample and data

To collect the data for this study, we surveyed 299 employees working in eight different organizations in Kuwait. The theoretical population of our research consists of public and private organizations in Kuwait. In selecting our target population, we used a convenience
sampling approach by focusing on one criterion, which is the willingness of target employees to participate to our study. The questionnaire was originally developed in English, then translated into Arabic to ensure the clarity of questions and avoid interpretation problems. Participants had adequate time to complete the questionnaire, and they were guaranteed anonymity (i.e. only aggregated data would be used for the study). In the first section of the questionnaire, participants rated, on a five-point Likert scale, their perceptions of coworker support, supervisors’ encouragement of creativity and level of engagement with the organization. They also reported their perceptions of their creative capabilities and individual performance. Then the second section of the questionnaire included the demographic questions, such that participants noted their position in the organization, years of experience and the type of industry in which their organization functioned.

Of 507 distributed questionnaires, we obtained 299 usable ones, for a response rate of 58.97 percent. Among this segment, 41.39 percent of the respondents worked in non-managerial positions, 29.47 percent were in supervisory positions, 20.20 percent were middle-level managers and 8.94 percent were executives. Furthermore, 88.74 percent of the participants worked in for-profit organizations and their tenure ranged from 1 to 10 years. For the type of industry, we found that 40.73 percent worked in the banking sector, 30.79 percent in the oil sector and 28.48 percent in the healthcare sector.

3.2 Measures
The constructs in this research were operationalized using reliable and valid scales previously published in high-quality peer-reviewed journals. All responses were assessed on a five-point Likert-type scale, ranging from “1: strongly disagree” to “5: strongly agree.”

3.2.1 Coworker support. Coworker support was assessed with a scale of four items developed by Podsakoff et al. (1997). The initial version includes high items, but a short version with four items was tested and adopted in some studies (i.e. Hon, 2013). We used the same short version which includes the following items, “Help each other out if someone falls behind in his/her work,” “Willingly share their expertise with other members of the crew,” “Willingly give of their time to help crew members who have work-related problems,” and “Encourage each other when someone is down.”

3.2.2 Leadership support. We measured leadership support variable through five items that have been developed by Scott and Bruce (1994). These items were “My supervisor encourages and emphasizes or reinforces creativity by learners,” “My supervisor respects learners’ ability to function creatively,” “My supervisor allows learners to try to solve the same problems in different ways,” “My supervisor will reward learners who are creative in doing their job,” and “My supervisor will publicly recognize those who are creative.”

3.2.3 Employee creativity. We used five items on a five-point scale to measure employee creativity (Scott and Bruce, 1994). The items are as follows “I always suggest new ways to achieve goals or objectives,” “I come up with new and practical ideas to improve performance,” “I search out new technologies, processes, and techniques,” “I suggest new ways to increase quality,” and “I come up with creative solutions to problems.”

3.2.4 Job performance. We measured job performance with four items developed by Podsakoff and Mackenzie’s (1989). These items were “I always complete the duties specified in my job description,” “I meet all the formal performance requirements of the job,” “I fulfill all responsibilities required by my job” and “I never neglect aspects of the job that I am obligated to perform.”

3.2.5 Employee engagement. The scale used to measure job performance was a five-item measure developed by (Soanea et al., 2012). The items were, “I feel positive about my work,” “I feel energetic in my work,” and “I am enthusiastic in my work.”
3.3 Common method bias
Since the data were collected using a single survey and on the same time, there is a potential risk of common method biases (Podsakoff et al., 2003). We used two ways to check and avoid this concern. First, to be concise and avoid any ambiguities regarding the items composing the survey, all the items were borrowed from prior studies published in good journals. Second, a Harman’s one-factor test was conducted using SPSS. All the items of the four variables used in our research were entered into a single-factor analysis. The results indicated that the single factor explains only 35.98 percent of the total variance between the variables, which was below the recommended maximum value of 50 percent. Thus, the common method variance seems to be negligible.

3.4 Data analysis methods
We used various methods to analyze our data. First, we used SPSS software to compute the descriptive statistics (means, standard deviations), the correlations between variables, as well as the Cronbach’s alpha to assess the reliability of the measures. Second, we used SmartPLS in addition to SPSS to calculate some reliability and validity indexes such as CR and AVE. Third, we used AMOS software to test the research hypotheses using the structural equation modeling techniques.

4. Results
The data analysis involved two steps. First, we estimate a measurement model to confirm the reliability and validity of the constructs. Second, we derive the structural model to test the relationships predicted in our theoretical model in Figure 1.

4.1 Measurement model
The assessment of the measurement model is determined by evaluating various statistical indicators of reliability, convergent validity and discriminant validity.

We first conducted the confirmatory factor analysis (CFA) to check whether the items effectively represent the latent variables. We applied the following indices and standards to assess the model fit of the CFA measurement model: goodness-of-fit index (GFI) higher than 0.80, comparative fit index (CFI) greater than 0.90, ratio of the $\chi^2$ to the degrees of freedom ($\chi^2$/df) less than 3.0, and root mean square of approximation (RMSEA) inferior to 0.08 (Hair et al., 2006). The estimation of the CFA measurement model provided the following results (GFI = 0.895; CFI = 0.947 ($\chi^2$/df = 2.063; RMSEA = 0.060), indicating that there is a satisfactory fit to the data. The CFA results revealed that all standardized factor loadings (see Table I) were larger than 0.60 and statistically significant ($p < 0.001$).

The reliability of the variables were assessed using two tools. First, all Cronbach’s $\alpha$ values are above the minimum threshold of 0.70 (see Table I). Second, we ran a principal components analysis with varimax rotation and extracted five factors. The five factors together explained 70.614 percent of the total variance. As detailed in Table II, all the items load significantly on their corresponding constructs, in support of unidimensionality or the degree to which items load only on their respective constructs. Based on this, the constructs meet acceptable reliability.

To assess the convergent validity of the scales, we checked three criteria (Hair et al., 2006). First, all indicator loadings are greater than 0.70. Second, the composite reliabilities (CR) exceed 0.8. Third, the average variance extracted (AVE) for each construct is greater than 0.5. Thus, the results in Table I confirm adequate convergent validity for the research constructs.

Finally, to assess discriminant validity, or the degree to which measures of two variables are empirically distinct, Bagozzi et al. (1991) suggest that the AVE of each latent construct...
must be greater than the squared correlations between that construct and the rest of the latent constructs. As Table III shows these conditions for discriminant validity are fulfilled.

4.2 Descriptive statistics and correlations
Table IV provides an overview of the descriptive statistics and correlations between the variables used in our study. As can be expected, coworker support was positively correlated to employee creativity and job performance \((r = 0.499, p < 0.01; r = 0.513, p < 0.01,\) respectively). Similarly, leadership support was positively correlated to employee creativity and job performance \((r = 0.266, p < 0.01; r = 0.303, p < 0.01,\) respectively). Furthermore, employee creativity was positively correlated to job performance \((r = 0.403, p < 0.01)\).

4.3 Structural model
To test the direct causal relationships inspired by our research model (Figure 1), we ran three structural equation models in AMOS 20.0.

Model 1: the first model tests for the direct relationships among the variables. Its estimation produced the following indicators of fit \((\text{GFI} = 0.900; \text{CFI} = 0.940; \chi^2/df = 2.234; \text{RMSEA} = 0.064)\). Because these fit indices all are satisfactory, the model is consistent with the data (see Appendix 1 for the recommended thresholds). We present the results of the first model estimation in Figure 2.
### Table II.
Results of principal component analysis with varimax rotation

<table>
<thead>
<tr>
<th>Constructs</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coworker support</td>
<td>0.692</td>
<td>0.132</td>
<td>0.057</td>
<td>0.145</td>
<td>0.814</td>
</tr>
<tr>
<td>Leadership support</td>
<td>0.132</td>
<td>0.669</td>
<td>0.072</td>
<td>0.145</td>
<td>0.814</td>
</tr>
<tr>
<td>Employee creativity</td>
<td>0.057</td>
<td>0.072</td>
<td>0.783</td>
<td>0.145</td>
<td>0.814</td>
</tr>
<tr>
<td>Job performance</td>
<td>0.145</td>
<td>0.145</td>
<td>0.145</td>
<td>0.783</td>
<td>0.814</td>
</tr>
<tr>
<td>Employee engagement</td>
<td>0.814</td>
<td>0.814</td>
<td>0.814</td>
<td>0.814</td>
<td>0.814</td>
</tr>
</tbody>
</table>

Notes: Values in italics on the diagonal represent the AVE; the other values are the squares of the interconstruct correlations.

### Table III.
Discriminant validity results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coworker support</td>
<td>4.174</td>
<td>0.608</td>
<td>1</td>
<td>0.307**</td>
<td>0.266**</td>
<td>1</td>
</tr>
<tr>
<td>Leadership support</td>
<td>3.690</td>
<td>0.812</td>
<td>0.307**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee creativity</td>
<td>4.015</td>
<td>0.672</td>
<td>0.499**</td>
<td>0.266**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Job performance</td>
<td>4.242</td>
<td>0.656</td>
<td>0.513**</td>
<td>0.303**</td>
<td>0.405**</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: **Correlation is significant at the 0.01 level (two tailed)

### Table IV.
Descriptive statistics and correlations

| Coworker support | 0.570 (8.113) |
| Leadership support | 0.179 (3.136) |

---

Employee creative behavior

![Figure 2. Results of direct hypothesized relationships (Model 1)](image)
In *H1*, we predicted that coworker support positively influences employee creativity. The results in Figure 2 confirm this positive relationship ($\beta = 0.570; t$-value $= 8.113; p = 0.000$) and support *H1*. In line with *H4*, we find a positive, significant relationship between leadership support and employee creativity ($\beta = 0.179; t$-value $= 3.136; p = 0.002$). For *H7*, the results confirm the predicted positive link between employee creativity and job performance ($\beta = 0.517; t$-value $= 7.429; p = 0.000$), as Figure 2 reveals.

Model 2: with a multigroup model, we examine the moderating effect of employee engagement on these direct relationships. We first dichotomized the sample into two groups, using the median value for employee engagement (4.00), such that the low engagement group included 105 responses and the high engagement group featured 194 responses. The estimation of the model produced the following fit indexes: GFI $= 0.855$; CFI $= 0.913$; $\chi^2$/df $= 1.803$; and RMSEA $= 0.052$. Thus, the indexes met the recommended thresholds (see the Appendix), suggesting a good model fit.

As Table V shows, employee engagement does not moderate the relationship between coworker support and creativity. Regardless of the degree of employee engagement, this relationship is always significant, across both low ($\beta = 0.607; t$-value $= 4.257; p = 0.000$) and high ($\beta = 0.498; t$-value $= 5.774; p = 0.000$) levels. Thus, we cannot confirm *H2*. However, in the relationship between leadership support and creativity, employee engagement has a moderating effect, such that the influence is stronger at high ($\beta = 0.175; t$-value $= 2.310; p = 0.021$) as opposed to low ($\beta = 0.063; t$-value $= 0.667; p = 0.505$) levels of employee engagement, in support of *H5*.

Model 3: with another multigroup moderation model, we examine the influence of tenure on the relationships of coworker support and leadership support with employee creativity. For this analysis, we again divided the sample into two groups: those with 1–5 years of work experience (132 responses) and those with more than five years of experience (164 responses). Three participants did not provide any response regarding their tenure. The estimation of this model produced the following fit indexes: GFI $= 0.853$, CFI $= 0.923$, $\chi^2$/df $= 1.797$ and RMSEA $= 0.052$, confirming the good fit of the model with the data (see the Appendix for the recommended thresholds).

As detailed in Table VI, tenure does not have any moderating effect on the relationship between coworker support and creativity. Regardless of the degree of tenure, this relationship remains positive and significant, for both low tenure ($\beta = 0.516; t$-value $= 4.678; p = 0.000$) and high tenure ($\beta = 0.589; t$-value $= 6.551; p = 0.000$), so we cannot...
confirm H3. Regarding the relationship between leadership support and creativity, tenure moderates it. As shown in Table VI, for low tenure, leadership support no longer exerts any impact on creativity ($\beta = 0.107; t$-value $= 1.222; p = 0.222$), whereas for high tenure, leadership support has a positive and significant effect on creativity ($\beta = 0.252; t$-value $= 3.301; p = 0.000$), in support of H6.

5. Discussion and implications

5.1 Main findings
This research has sought to determine whether and how leadership support and coworker support influence employee creativity. Using structural equation modeling, we find that coworker support has a direct positive influence on employee creativity (Woodman et al., 1993; Zhou and George, 2001; Cheung and Wong, 2011). Coworker support can foster creativity directly by providing employees with new ideas and knowledge derived from their experience, especially when they face difficult and unclear tasks (Madjar, 2005). This finding also depends on how employees perceive, comprehend and understand the context in which they work. For example, coworker support has a stronger effect on employee creativity in a collectivist culture than in an individualist culture (Markus and Kitayama, 1991). Kuwaiti employees tend to interact intensively, in concordance with their cultural context’s requirements; Kuwait is a collectivist society (Hofstede, 1980). In general, this study highlights the important influence of social forces such as coworker support on employee creativity, especially when they involve cultural norms (Farmer et al., 2003).

The results also show that leadership support has a direct positive influence on employee creativity, in line with previous studies by Zhou and Shalley (2003) and Kossek et al. (2011). By examining the impact of leadership support on employee creativity, we help explain the role of leaders for enhancing employee creativity, including in terms of providing an appropriate environment (e.g. George, 2007; Reuvers et al., 2008; Zhang and Bartol, 2010). Leadership and creativity thus seem inextricably associated (Castro et al., 2012).

We included two moderating variables (employee engagement and tenure) to test whether the degree of employee engagement and tenure strengthen the effects of coworker and leadership support on employee creativity. To the best of our knowledge, this research is the first to rely on employee engagement and tenure as moderating variables, in an attempt to explain the influences of coworker and leadership support on employee creativity. In contrast with our expectations, employee engagement and tenure do not moderate the relationship between coworker support and employee creativity (see Tables IV and V). Employees who intend to support their coworkers are willing to do so, regardless of their engagement or tenure. These findings may arise because, in non-western countries (Kuwait, in our case), the culture encourages coworkers to help one another. They do not rely on personal or contextual factors to decide whether to provide help. As an Arabic country, Kuwait’s collectivist culture also encourages citizens to help one another without looking for any return. This religious orientation recommends that people support one another. This research contributes to organizational behavior literature by offering new insights into whether and how people support one another in the workplace.

However, we find that both employee engagement and tenure moderate the relationship between leadership support and employee creativity. Leaders can more easily influence and motivate committed employees to generate new ideas. Similarly, senior employees are more likely to be encouraged by their leaders to generate new ideas. Over time, people can build a serious, trusting relationship with their leaders, as they become more motivated and creative. These results support previous research that indicates that employees tend to be more creative when they realize their efforts will be valued, but they also demonstrate the important role of employee engagement and seniority in promoting and supporting effective creative processes within organizations. These findings contribute to leadership literature,
by demonstrating in particular that the relationship between leaders and employees is deeply complex. The moderation approach we have used in this research helps explain this relationship by showing that the relationship between leaders and employees is affected by personal and contextual factors, such as employee engagement and tenure.

Finally, we show that employee creativity has a positive impact on job performance. As we hypothesized, employees who enjoy being creative tend to internalize the belief that creative behavior will benefit their work and thus create superior performance. This result enriches extant literature by confirming that people who exert efforts to create new ways to solve problems and perform their tasks are more likely to attain superior performance (Grewal et al., 2009; Tierney and Farmer, 2004). We build theoretical arguments for and demonstrate a connection between employee creativity and job performance. Creativity researchers have suggested the need for a greater understanding of the processes that employees follow to produce better performance outcomes. To the best of our knowledge, no previous research has assessed the potential connections between employee creativity and job performance in non-western cultures.

5.2 Implications for theory

The findings of this study contribute to the creativity literature by highlighting the importance of leadership support and coworker support on influencing innovative behavior. We contribute to extant literature by affirming the important role of leadership in enhancing employee creativity, as well as by revealing the importance of a high-quality relationship between leaders and followers. Such a supportive relationship with supervisors can increase employees’ beliefs that their performance will be valued and rewarded, which then can drive their creative behavior.

Furthermore, our results clarify the relative effect of coworker support on employee creativity. These findings contribute to this stream of research based on the literature we developed concerning employee perceptions of coworker support. Factors such as trust, shared knowledge and expertise and useful feedback from coworkers play an important role in shaping perceptions of the effectiveness of creativity. Such feedback may enhance employee beliefs that their new ideas have a great opportunity to be encouraged by coworkers and thus successfully implemented.

In addition, there has been a growing interest in examining outcomes of a creative process. Our findings show that creative employees are more likely to generate high level of job performance. By this study, we demonstrated that leader’s expectations of creativity could increase the likelihood that employees exhibit higher levels of job performance over time.

A final potential theoretical contribution of this study is the development of new theoretical perspective to examine new mechanisms that shape creativity. There is no or few evidence in the literature focusing on job tenure or employee engagement as moderators of the relationship between leadership support and coworker support on employee creativity. Nevertheless, our examination of the moderating effect of employee engagement and job tenure can enrich our understanding of the creativity theory. The results of our study indicate that in a supportive organizational context, employees will engage more in enhancing their creativity. Furthermore, job tenure may play a significant role in facilitating the creative enactment of employees. More specifically, employees with pertinent knowledge and understanding of the task they behave in a way that supports new ideas and triggers individual creativity.

5.3 Implications for practice

These findings have a few empirical implications for leaders. First, our study reveals that leadership support is critical for employee creativity. Literature pertaining to transformational leadership confirms the key role of leaders in influencing and reshaping
employees’ behaviors. To establish creative, innovative organizations, leaders must be conscientious and recognize their role and contributions to creative, innovation processes. As Bledow et al. (2009) and Shalley and Gilson (2004) note, leadership behavior is a critical antecedent of employee creativity. Second, the role of leadership in supporting creativity may vary, depending on personal and contextual factors. The current study shows that employee engagement and tenure both influence the impact of leadership on employee creativity. Thus, leaders should take care to behave appropriately with different members of their organization, to influence their creative behavior. As revealed by our study, committed employees are more likely to be affected by leadership support and thus become more creative. Leaders must also ensure the engagement of their organizations’ members. Similarly, our results show that the more time people work for a company, the more they are influenced by their leader. Leaders accordingly must behave differently with new employees relative to those who have been with the company for a longer time. Third, coworker support is associated with employee creativity. Employees are a main source of organizational success (Barney, 1991); as Ouakouak and Ouedraogo (2013, p. 323) explain, “To remain competitive in an increasingly dynamic and complex environment, organizations need to optimize and wisely use their human resources capacities and get strong ‘buy-in’ from all their employees.” To use their human capital effectively, organizations should facilitate employee collaboration. Modern successful organizations require more than talented, skilled people; they demand employee collaboration. Therefore, organizations should find ways to ensure such collaboration, as a critical driver of employee creativity. Fourth, we uncover a positive association between employee creativity and job performance. To remain competitive in dynamic business environments, organizations must enhance creativity among their employees, which provides a route to remain competitive, as noted by Baer and Oldham (2006, p. 963), who recognize that “employee creativity can make a substantial contribution to an organization’s […] competitiveness.”

A final implication of this research provides crucial insight into the leadership literature since the model to predict employee creativity was developed and tested in a non-western context like Kuwait. To the best of our knowledge, our research sheds the first light on employee creativity as outcome of the impact of leadership support and coworker support on employee creativity and job performance in Kuwait. In this regard, it is important to mention that Kuwaiti organizations are becoming more westernized with the way of doing business. Since the Kuwaiti collective culture is built on a foundation that consists of friendship, sympathy and care, therefore, leaders and coworkers should carry out these characteristics when promoting creative behavior.

6. Limitations and further research
As does all research, our study has some limitations that point to potential directions for research. First, we relied solely on quantitative methods for this study. It would be beneficial for further research to combine quantitative and qualitative methods and thereby gain richer insights into these studied relationships. Second, we conducted this research in Kuwait. Accordingly, it would be interesting to replicate our research with comparative studies conducted across various countries, especially nearby places such as Qatar, Saudi Arabia, UAE, Oman and Bahrain. Third, the current research offers new opportunities to examine and exploit other variables that have the potential to moderate or mediate the relationships of leadership and coworker support with employee creativity. Fourth, we found a positive relationship between employee creativity and job performance, whereas Gong, Kim, Zhu and Lee (2013), Gong, Zhou and Chang (2013) argue that employee creativity does not necessarily have a strong, direct impact on firm performance. Therefore, it would be helpful to adopt moderating and mediating approaches, in attempts to explain and understand the nature of the relationship between employee creativity and job performance in more depth.
1. A relationship is significant if the $t$-value is greater than 1.96.

References


Costa, P.T. and McCrae, R.R. (1992), Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual, Psychological Assessment Resources, Odessa, FL.


Further reading


Appendix

<table>
<thead>
<tr>
<th>Index Name</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of the $\chi^2$ to the degrees of freedom ($\chi^2$/df)</td>
<td>$&lt; 3.00$</td>
</tr>
<tr>
<td>Comparative fit index (CFI)</td>
<td>$&gt; 0.90$</td>
</tr>
<tr>
<td>Goodness-of-fit index (GFI)</td>
<td>$&gt; 0.80$</td>
</tr>
<tr>
<td>Root mean square of approximation (RMSEA)</td>
<td>$&lt; 0.08$</td>
</tr>
</tbody>
</table>

Notes: “Suggested values are based on Gefen et al. (2000); Etezadi-Amoli and Farhoomand (1996) and Hair et al. (2006).

About the authors
Dr Michel Zaitouni is Assistant Professor of Business Administration at Gulf University for Science and Technology in the College of Business Administration, Kuwait, since September 2008. Research Interests Dr Zaitouni’s fields of interest are human resource management, leadership, employee training and development and learning styles in organizations. Using matched employer employee data, his research in industrial organization examines the effect of human resource practices on firm and worker performance. His research in learning styles focuses on the link between organizational learning and strategy formulation.

Dr Mohamed Laid Ouakouak is Assistant Professor of management at Gulf University for Science and Technology, Kuwait, since September 2014. He graduated from the University of Reims Champagne-Ardenne (France) with a Master of Science in management in 2007 and a PhD in management in 2012. Dr Ouakouak has served as a full time lecturer in management at University of Reims Champagne-Ardenne and subsequently, as an assistant professor at IDRAC International School of Management in France. He has published in several international peer-reviewed journals. Dr Ouakouak main research interests include creativity, innovation, employee engagement, leadership, strategy. Dr Mohamed Laid Ouakouak is the corresponding author and can be contacted at: mohamed.ouakouak@gmail.com

Employee creative behavior

1763

Table AI.
Recommended thresholds

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
A valid and applicable measurement method for knowledge worker productivity

Jalil Heidary Dahooie
School of Management,
University of Tehran, Tehran, Islamic Republic of Iran

Mohammad Reza Ghezel Arsalan
Department of Industrial Engineering,
University of Tehran, Tehran, Islamic Republic of Iran, and

Ali Zolghadr Shojai
Department of Industrial Engineering,
Amirkabir University of Technology, Tehran,
Islamic Republic of Iran

Abstract

Purpose – The purpose of this paper is to propose a new method for knowledge worker productivity measurement which is based on valid principles and appropriate viewpoints.

Design/methodology/approach – Based on an extensive and thorough literature review the elements that need to be taken into consideration, while designing a method for knowledge worker productivity measurement, are determined and divided into principles and viewpoints. These elements must be incorporated into the design of knowledge worker productivity measurement methods so that the correctness and accuracy of these methods can be verified. The proposed model, which is based on appropriate principles and viewpoints, determines the outputs of knowledge work with respect to the tasks that a worker’s job includes. Considering nine measures, these outputs are evaluated using fuzzy numbers and, then, quantified. The inputs of knowledge work are knowledge, skills and abilities (KSAs) required to do the job. These inputs are identified and quantified using Job Element Method. Furthermore, fuzzy Data Envelopment Analysis is employed to model the productivity.

Findings – In this paper, the proposed method for knowledge worker productivity measurement follows both appropriate principles and viewpoints, simultaneously. In order to validate the obtained results and explore the applicability of the proposed method, a case study was carried out at an Iranian organization in electric power industry. Statistical analyses are employed to prove the validity of the results. Based on the obtained results, the productivity of a knowledge worker is said to be high when he/she delivers the expected amount of job outputs considering the values of his/her inputs (KSAs).

Originality/value – The originality of this paper is twofold. First, the extracted principles and viewpoints can serve as a guideline for the development of similar methods. Second, the proposed model offers an effective and efficient tool that can serve as the basis for the comparison among relative productivity of knowledge workers. Furthermore, the obtained results could form a basis to examine the productivity trend of each knowledge worker over different periods of time.

Keywords Viewpoint, Productivity measurement, Knowledge worker, Principle, Fuzzy DEA

Paper type Research paper

1. Introduction

Currently, the number of knowledge workers and their rate of growth are larger than those of manual workers (Drucker, 1999; Xiao and Nembhard, 2014). Consequently, knowledge worker management and particularly knowledge worker productivity measurement are frequently faced by academics and practitioners to the extent that Drucker believed knowledge worker...
productivity measurement and knowledge worker productivity improvement to be the most challenging managerial issues of the twenty-first century (Drucker, 1991, 1999).

Drucker (1959) and Machlup (1962) first defined knowledge work as a specific type of work which needs more thinking rather than performing physical labor (Cortada, 1998; Ökkonen, 2003). The proposed definitions of knowledge work/knowledge worker are not homogeneous which, in turn, has resulted in ambiguity (Guns and Valikangas, 1998; Shi-you, 2008). Most of the definitions are also conceptual and difficult to apply to real-life cases (Shi-you, 2008). Ramirez and Steudel (2008) took the first step to establish the exact structured definition of knowledge work. Afterwards, Heidary Dahooie et al. (2011) tried to take the matter several steps further and complete their work. Principles and concepts developed by Heidary Dahooie et al. (2011) are followed in this paper and knowledge work is defined as a job in which workers are involved with knowledge; and knowledge activities (such as data gathering, analyzing information and decision making) related to the main activities of the job.

Considering the characteristics of knowledge work (such as non-repetitiveness, creativity and intangibility), the managerial tools and models developed for manual workers cannot be useful for knowledge workers (Xiao et al., 2012). Thus, it is necessary to take the initiative and design new methods for knowledge worker productivity measurement (Ramirez and Nembhard, 2004; Palvalin et al., 2015).

Moreover, when it comes to knowledge worker productivity, both knowledge workers and managers are fully aware of the challenges and difficulties that arise while measuring the knowledge worker productivity. However, their knowledge regarding methods and solutions for knowledge worker productivity management is not sufficient to resolve the issues (Shi-you, 2008; Xiao and Nembhard, 2014). Nevertheless, Drucker believed that although managers do not know how to measure the productivity of knowledge workers, they have been able to improve it (Tankoonsombut, 1998).

The importance of knowledge worker productivity measurement is self-evident for many researchers (Drucker, 1999); however, a number of reasons can be found in the literature which are listed below:

- the complexity of new businesses environment and considerable development of knowledge intensive and service sectors compared to the other sectors (Ray and Sahu, 1989);
- the higher growth rate of knowledge workers’ population than manual workers’ and the requirement of using appropriate management methods for knowledge workers (Tankoonsombut, 1998; Ray and Sahu, 1989; Thomas and Baron, 1994; Davis, 2002; Mundbrod et al., 2013);
- the shortcomings of traditional productivity management methods used for manual workers considering the specific characteristics of knowledge workers (such as cleverness, creativity, academic and practical training) (Xiao et al., 2012; Xiao and Nembhard, 2014; Dan, 2014; Davenport, 2015);
- despite the low rate of productivity improvement regarding this type of workers, there are ample opportunities for productivity improvement (Thomas, 2003; Palvalin et al., 2015); and
- the interdependence of success factors of organization (such as realization of goals, growth and profitability) and knowledge worker productivity (Pan et al., 2008; Ghezel Arsalan et al., 2014).

In this paper, we aim to develop an appropriate method for knowledge worker productivity measurement. The desired method must be designed based on the extracted principles. It is also necessary to determine the right and applicable viewpoints while devising the method.
The rest of the paper is organized as follows: the literature review is presented and analyzed in Section 2 in order to identify research gaps and an appropriate approach is then discussed based on the identified research gaps to resolve them. Section 3 gives insight into the principles and viewpoints which must be followed while designing the proposed method. Section 4 presents the proposed method in detail. A case study is explained in Section 5. Finally, Section 6 provides conclusion and outlines future research trends.

2. Literature review

Study of the approaches presented to measure the knowledge worker productivity shows that there are two distinct types of approaches: conceptual models and methods for knowledge worker productivity measurement (Ramirez and Nembhard, 2004).

It must be noted that conceptual models do not represent a specific method; rather, they are broad categories of measurement approaches and mark several significant points in the field of knowledge worker productivity measurement. Consequently, these types of models will be used as the main source to extract the principles of knowledge worker productivity measurement. Six conceptual models identified in the literature are mentioned below:

(1) Mundel’s Model (Mundel, 1975; Ramirez and Nembhard, 2004).
(2) Bumbarger’s Model (Bumbarger, 1984; Ramirez and Nembhard, 2004).
(3) Gordon’s Model (Gordon, 1997; Ramirez and Nembhard, 2004).
(4) Drucker’s Model (Drucker, 1999; Okkonen, 2003; Ramirez and Nembhard, 2004).
(5) Schroeder et al.’s Model (Schroeder et al., 1982).
(6) Schainblatt’s Model (Schainblatt, 1982; Tankoonsombut, 1998).

These models were carefully studied and it became evident that the principles used by each of them to measure knowledge worker productivity have much in common. Table I illustrates seven extracted principles along with corresponding references for each principle.

In addition to these conceptual models, other different methods have also been developed by researchers to measure the productivity of knowledge workers. These methods provide a framework for measuring the productivity of knowledge workers. Ramirez and Nembhard (2004) described 21 different methods designed to measure the productivity of knowledge workers. Few other relevant methods have also been identified and are listed below:

- Modified Morgeson and Humphrey’s Work Design Questionnaire (Moussa et al., 2017).
- Fuzzy Data Envelopment Analysis (DEA) (Abdoli et al., 2011).

<table>
<thead>
<tr>
<th>No.</th>
<th>Features</th>
<th>Identified conceptual models</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The first step is to determine the goals and tasks</td>
<td>Mundel, Bumbarger, Drucker, Schainblatt</td>
</tr>
<tr>
<td>2</td>
<td>The focus is on effectiveness</td>
<td>Gordon, Drucker, Schroeder’s et al., Schainblatt</td>
</tr>
<tr>
<td>3</td>
<td>The focus is on efficiency</td>
<td>Mundel, Gordon, Schroeder’s et al., Schainblatt</td>
</tr>
<tr>
<td>4</td>
<td>Innovation is considered</td>
<td>Bumbarger, Drucker, Schroeder’s et al.</td>
</tr>
<tr>
<td>5</td>
<td>The autonomy of knowledge workers is considered</td>
<td>Bumbarger, Drucker, Schroeder’s et al.</td>
</tr>
<tr>
<td>6</td>
<td>The focus is on the intangible factors and the knowledge required to carry out the job</td>
<td>Mundel, Drucker</td>
</tr>
<tr>
<td>7</td>
<td>The focus is on quality</td>
<td>Gordon, Drucker, Schainblatt</td>
</tr>
</tbody>
</table>

Table I.

Seven extracted principles
Productivity Payback (Kaczmarczyk and Murtough, 2002).

Cognitive Turnover Methodology (Jones and Chung, 2006).

The detailed investigation and comparison of these models led to the identification of similarities and differences among them (see Table AI). It was also determined that to what extent each model would follow the principles mentioned in Table I. It can be inferred from Table AI that not all the models were initially designed to measure the productivity. Some of these models aim to analyze the effects of different factors on knowledge worker productivity.

These methods use subjective productivity measurement (SPM) which is based on collecting subjective data from target group and performing the productivity analysis accordingly (Antikainen and Lönqvist, 2006; Laihonen et al., 2012). Possible problems affecting the factors of productivity as well as targets for HR development are identified through soliciting employees' opinion (via interviews or questionnaires) and analyzing the data.

Moussa et al. (2017) discussed that human resource management's Job/Work Design Theory can be used to develop a method which can be applied to all knowledge workers and knowledge-based companies. Then, a new method is proposed based on the Modified Morgeson and Humphrey’s Work Design Questionnaire.

Several models take efficiency criteria into account, while others focus on effectiveness. Only few models have been proposed to incorporate both indices. Moreover, only a small number of models address and quantify intangible factors that play an important role in the knowledge work; and productivity measurement models that were examined focus on individual and team levels. Models that measure productivity at individual level focus on activities, tasks and jobs. Finally, in some models, knowledge worker productivity measurement entails obtaining feedback and opinion from the knowledge worker himself/herself.

Besides, Ramirez and Nembhard (2004) recognized and analyzed some variables and factors used in the methods. These variables and factors include the following:

- quantity;
- cost or profitability;
- timeliness;
- autonomy and independency;
- efficiency;
- quality;
- effectiveness;
- customer satisfaction;
- creativity/innovation;
- project success (successful accomplishment of projects);
- responsibility/importance of work;
- Knowledge workers’ perception of productivity; and
- absenteeism.

2.1 Research gap

It is evident that various researches and studies have been conducted to tackle the subject of knowledge worker productivity measurement, but there is no guideline on designing productivity measurement models that are efficient and valid. Although some efforts were
made by Ramirez and Nembhard (2004) to take the matter several steps further, this paper specifically addresses design principles for knowledge worker productivity measurement methods. In addition to these principles, it became clear from reviewing the existing literature that productivity measurement models are designed and implemented with respect to multiple requirements and concerns such as the target application, nature of activities and culture, which has caused some components of the models to be different from each other. These differences are referred to as “viewpoints” in this paper. The viewpoints given in this paper have been extracted from the models available in the literature.

Based on the design principles and in accordance with appropriate viewpoints, a new model is proposed. What differentiates this model from all the others is complete adherence to the design principles. Furthermore, the research works reviewed in the literature suffer from another shortcoming. They do not provide necessary details related to the implementation of the designed models in organizations; so, the applicability of the models is in question. For example, this issue can be detected in “Efficiency, standard times and operating efficiency” model and “Operation-based productivity measure” model. This paper provides full description of the steps needed to be taken to implement the model along with the required details.

3. Principles and viewpoints

Based on the analysis of the models and methods, it was concluded that there are two main issues that must be considered while designing a method for knowledge worker productivity measurement: principles and viewpoints.

“Principle” is a statement (in the form of law or axiom) accepted by researchers only because it seems logical, and is not necessarily proven to be true (Rafipoor, 2002). It is essential for the developed measurement methods to follow the extracted principles in order to be valid. The six conceptual models mentioned earlier are used as the main source to extract the principles of knowledge worker productivity measurement.

In addition to these principles, the proposed measurement methods share another difference that originates from a notion called “viewpoint.” As a matter of fact, the design of measurement methods requires the consideration of various subjects and researchers take different approaches to deal with them. This indicates that researchers have different viewpoints. In order to develop measurement methods, researchers choose viewpoints based on the circumstances that they face such as goals or intended applications of the methods.

It can be helpful to mention some main applications of the productivity measurement regarding knowledge workers. In what follows, several instances are presented:

1. Sardana and Prem (1987) considered three objectives for productivity measurement:
   - identifying the potential improvements;
   - making decisions regarding resource allocation; and
   - determining how well previously established goals have been met.

2. Schroeder et al. (1982) considered three general applications for productivity measurement:
   - self-development;
   - performance appraisal, salary and job promotion; and
   - feedback and communication.

Ramirez and Nembhard have also mentioned some applications and advantages of productivity measurement. Table II shows some of the applications along with advantages (Ramirez and Nembhard, 2004).
The type of application intended for productivity measurement affects the measurement method. For example, when the application is personal development, it is interesting for each person to measure his/her own performance and use it to improve the productivity (Schroeder et al., 1982). Choosing the right viewpoints plays an important role in the applicability of the developed method.

Table AI forms the basis for the extracted viewpoints. At first, the table had three columns (Identified method, Evaluated variable(s) and Overall approach). Afterwards, it was completed gradually by comparing various models and identifying the similarities and differences. Finally, four viewpoints were extracted.

In what follows, the principles and viewpoints are introduced.

3.1 Extracted principles

3.1.1 Identification of the job’s tasks and goals as the first step. The first step in measuring the productivity of knowledge workers is the identification of the tasks and goals related to their jobs. Considering the nature of the knowledge work, it is likely that knowledge workers, who have the same job, do not share the same understanding of its tasks and goals (Drucker, 1999). If knowledge workers hold different opinions on their tasks, the result of productivity measurement will not be reliable. So, the first step would be to ask “what are the tasks?” (Drucker, 1999).

3.1.2 Paying attention to productivity, efficiency and effectiveness at the same time. As mentioned before, based on the presented conceptual models and opinions shared by most of the researchers, both “efficiency” and “effectiveness” must be taken into account while developing the concept of knowledge worker productivity (Schainblatt, 1982; Klassen et al., 1998; Thomas, 2003), but some researchers believe that the productivity regarding knowledge work is completely independent of efficiency and effectiveness. They integrate these three concepts into a more general concept called “performance” (Sink, 1983; Phusavat et al., 2009). Moreover, some researchers use mathematical functions such as division and multiplication to calculate the productivity (Chase and Aquilano, 1992; Starr, 1996; Klassen et al., 1998).

3.1.3 Considering the intangible factors. Intangibility could be the most important characteristic of knowledge work which makes it distinct from manual work (Moussa et al., 2017). Inputs, outputs and even processes of knowledge work have some degrees of intangibility (Guthrie and Petty, 2000; Lim and Dallimore, 2004; Taylor, 2006; Bontempi and Mairesse, 2008; Nesta, 2008). Furthermore, the main source of intangibility is the knowledge that is involved with the inputs, outputs and processes. It can be inferred from the studied methods that some researchers have totally ignored this principle. In other words, although

<table>
<thead>
<tr>
<th>Application</th>
<th>Potential benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring the performance of individuals, systems and the firm in order to identify unusual patterns of productivity</td>
<td>Monitoring the knowledge workers</td>
</tr>
<tr>
<td>Determining the capacity of knowledge workers, performance prediction</td>
<td>Capacity planning</td>
</tr>
<tr>
<td>Better work assignment, better selection of individuals, defining requirements, job assignment decisions</td>
<td>Strategic planning</td>
</tr>
<tr>
<td>Displaying the changes in the current system and simulating the changes before they occur</td>
<td>Simulating the performance of knowledge workers</td>
</tr>
<tr>
<td>Comparing the performance of individuals, divisions, teams, systems, companies, etc.</td>
<td>Establishing new benchmarks</td>
</tr>
<tr>
<td>Reducing subjectivity associated with evaluations</td>
<td>Consistent evaluation method</td>
</tr>
<tr>
<td><strong>Source:</strong> Ramirez and Nemhard (2004)</td>
<td></td>
</tr>
</tbody>
</table>

**Table II.** Applications of productivity measurement along with advantages
the “intangibles” play a vital role in the determination of productivity improvement, the link between these two is poorly understood (Bontempi and Mairese, 2008).

3.1.4 Considering the quality and innovation. Another principle used by many researchers and conceptual models is the need for the consideration of quality and innovation in measuring the productivity of knowledge workers (Thomas and Baron, 1994; Takala et al., 2006; Haas and Hansen, 2007; Phusavat et al., 2009; Liao, 2012). Quality and innovation are so significant that Drucker (1999) believed them to be more important than the quantity of knowledge worker’s outputs (Drucker, 1999; Kelloway and Barling, 2000). However, the review of methods for knowledge worker productivity measurement reveals that researchers have not paid enough attention to this principle.

3.1.5 Considering the autonomy of knowledge workers. One of the main subjects in productivity measurement is the person or department responsible for the evaluation of the knowledge worker productivity. Most of the researchers advise that knowledge workers themselves should participate in the development of indices and processes related to productivity measurement (Schroeder et al., 1982; Coates, 1986; Thomas and Baron, 1994; Drucker, 1999; Lind and Sulek, 2000; Holtshouse, 2010; Vuolle et al., 2014; Davenport, 2015). However, due to cultural issues and traditional approach toward the management of knowledge workers, a large number of models are designed based on the suggestions provided by managers and productivity evaluation is also carried out only by them. In other words, some managers have not acknowledged the autonomy of knowledge workers and its importance.

3.2 Identified viewpoints
Identified viewpoints fall into four categories. In what follows, we present the description of each category.

3.2.1 The level of job hierarchy for which the productivity is measured. One of the differences among viewpoints employed by productivity measurement models is the level of job hierarchy for which the productivity is measured. As the literature review implies, developed models usually focus on job, task and activity levels. Some researchers such as Drucker believed that the best level for measuring the productivity of knowledge workers is the task level (Drucker, 1999).

3.2.2 The individual level and the team level. Researchers have not reached consensus on the individual level or the team level. Some claim that unlike manual workers, measuring the productivity of knowledge workers must be focused on the individual level (Greenberg, 1986). On the other hand, a number of researchers believe that the team level is the most appropriate level for measuring and analyzing the productivity of knowledge workers (Sassone, 1996; Lind and Sulek, 2000; Thomas, 2003; Liao, 2012). Furthermore, a considerable number of researchers consider both levels to be proper for measuring the productivity of knowledge workers. At the same time, they insist that productivity measurement must be focused on the team level if meaningful indices cannot be defined for the individual level with respect to the organizational goals (Schroeder et al., 1982; Klassen et al., 1998; Wake, 2015; Palvalin et al., 2015). Another group of researchers state that choosing the appropriate level is affected by some factors such as cultural issues surrounding organizations or countries. For example, performing productivity measurement at the team level is preferred for an organization in Japan, whereas carrying out productivity measurement at the individual level for a firm in a Latin country could be more beneficial (Takala et al., 2006).

3.2.3 The process and the output. Another difference among viewpoints revolves around focusing on the process or the output. Most of the researchers find it challenging to focus on process (Schainblatt, 1982; Ray and Sahu, 1989); thus, they prefer to deal with output
The main reason is the intangibility faced by researchers while analyzing most of the knowledge activities. These activities are done through mind abilities of knowledge workers and hence cannot be seen (Kelloway and Barling, 2000).

On the other hand, some researchers believe that considering process is as important as considering output (Scott, 2003; Scarbrough, 1999; Palvalin et al., 2015). Moreover, some methods focus only on process and do not take output into account due to several reasons such as measurement difficulties.

3.2.4 The overall approach toward productivity modeling. Another issue that needs to be mentioned is the overall approach of current methods in productivity modeling. Generally, there are two types of viewpoints discussing productivity modeling: classic and SPM (Najafi, 2011; Antikainen and Lönnqvist, 2006). The detailed information on these two viewpoints along with several examples is given in Table III.

The review of the methods shows that these two viewpoints have also been considered for the knowledge worker productivity measurement.

Consequently, it can be concluded that methods in the field of knowledge worker productivity measurement do not benefit from all the identified principles at the same time. The scientific basis and applicability of these methods suffer from this fact.

Ramirez and Nembhard (2004) and Schroeder et al. (1982) drew a similar conclusion. They believed that several efforts have been carried out to measure the productivity of knowledge workers and few approaches have been developed, but none of them has been accepted by the majority of researchers. Furthermore, none of the approaches has been properly implemented due to the lack of complete features expected from an appropriate measurement method.

4. The proposed method

The proposed principles were followed to design a new method. Moreover, since the main purpose of designing the proposed method is to monitor knowledge workers and establish benchmarks for them, appropriate viewpoints were also selected. The viewpoints considered in the design of the proposed method are demonstrated in Table IV.

So, in this paper, the exact definition of knowledge worker productivity is as follows: "Knowledge worker productivity is a concept that determines to what extent a knowledge..."
worker delivers outputs or achieves the intended goals of his/her job in a creative, efficient and effective way within a specific time period considering his/her own competencies (knowledge, skills and standard abilities required for the job). With respect to this definition, the designed model and the proposed method have the following features:

- the model is designed in accordance with the extracted principles of knowledge worker productivity measurement;
- each knowledge worker is evaluated based on the job to which he/she is assigned; and
- the proposed method is capable of comparing the productivity of different groups of knowledge workers within a specific time period and the productivity of a specific individual within consecutive time periods.

Considering the given definition of the knowledge worker productivity, its associated index is defined as the ratio of quantified job outputs to quantified personal competencies.

Therefore, three stages are required for the measurement process. The first stage is focused on the quantification of the job outputs delivered by knowledge worker. The quantification of job inputs (the extent to which competencies of knowledge worker match job requirements) is done in the second stage. In the third stage, a DEA-based model is used to calculate the productivity index of knowledge worker (see Figure 1).

4.1 The quantification of job outputs

A four-step process is employed to quantify the job outputs. The process is demonstrated in Figure 2. In what follows, each step of the process is explained.

4.1.1 Identification of the job’s tasks and evaluation of their importance. In order to quantify job outputs, we first need to identify the job’s tasks. In most of the cases, the tasks related to a specific job are documented. When faced with lack of proper documentation, we have to identify the job’s tasks through appropriate methods of job analysis such as

<table>
<thead>
<tr>
<th>Available categories to be considered for the selection of viewpoints</th>
<th>Selected viewpoint (in this paper)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The level of job hierarchy for which the productivity is measured</td>
<td>The task level</td>
</tr>
<tr>
<td>The individual level and the team level</td>
<td>The individual level (the focus is on knowledge workers)</td>
</tr>
<tr>
<td>The process and the output</td>
<td>The output</td>
</tr>
<tr>
<td>The overall approach toward productivity modeling</td>
<td>Classic</td>
</tr>
</tbody>
</table>

**Table IV.**

**Selected viewpoints**

4.1.2 Quantification of knowledge work outputs.

4.1.3 Calculating productivity index (DEA model).

**Figure 1.**

The proposed model for knowledge worker productivity measurement.
Functional Job Analysis and Hay method. Then, the duration and relative importance of each task is calculated to determine its weight. In order to do this, an appropriate framework introduced by Heidary Dahooie et al. (2011) is employed.

First, the duration of each task is determined using self-reporting technique. $TW_{NT}^a$ represents the normalized time proportion of task $a$ and is defined as the ratio of the time required to accomplish task $a$ to the total time needed to accomplish the job. $TW_{NT}^a$ is given by the following equation:

$$TW_{NT}^a = \frac{TW_T^a}{\sum_{k=1}^{m} TW_T^k},$$

(1)

where $TW_T^a$ and $m$ are the duration of task $a$ and the total number of the tasks of the job, respectively.

In order to obtain the importance of task $a$ ($TW_{NI}^a$), the job incumbent and other subject matter experts (SMEs) are asked to assign a number from the set $\{1, 2, 3, \ldots, 7\}$ to each task. “1” is assigned to the least important task and “7” is assigned to the most important task. Next, the normalized importance weight of task $a$ ($TW_{NI}^a$) is calculated using the following equation:

$$TW_{NI}^a = \frac{TW_I^a}{\sum_{i=1}^{m} TW_I^i},$$

(2)

Finally, the total weight of task $a$ ($W_a$) is calculated by the following equation:

$$W_a = \frac{TW_{NI}^a + TW_{NT}^a}{2}.$$ 

(3)

4.1.2 Identification of the job outputs and evaluation of their importance. The worker is asked to determine the job outputs with respect to the identified tasks. Then, the early list of these outputs is presented to main stakeholders (such as immediate manager, head of the department, colleagues, etc.) to prepare the final list of job outputs. The weight of each output is calculated based on the total weight of each task that leads to the output.

We consider Table V in order to calculate the weight of each output. Rows and columns of Table V represent tasks and outputs, respectively. $b_{ij}$ is a real number that shows the impact of task $i$ on the output $j$ and $n$ is the number of job outputs. These parameters are determined based on the worker opinion and in a way that the following equation is fulfilled:

$$\sum_{i=1}^{m} b_{ij} = 1 \quad j = 1, 2, \ldots, n.$$ 

(4)

The weight of each output ($w_j$) is calculated according to the following equation:

$$w_j = \sum_{i=1}^{m} b_{ij} \cdot W_i^j \quad j = 1, 2, \ldots, n.$$ 

(5)
4.1.3 Assessment of job outputs. After the determination of job outputs and their weight, we must assess them according to appropriate measures. Two groups of measures introduced by Ramirez and Nembhard (2004) and Abdoli et al. (2011) were integrated:

1. quantity;
2. costs and/or profitability;
3. timeliness;
4. efficiency;
5. quality;
6. effectiveness;
7. customer satisfaction;
8. innovation/creativity; and
9. learning/process improvement.

To assess the outputs of each job, it is necessary to select an appropriate set of the mentioned measures and then, corresponding to each measure, appropriate indices are determined based on the opinions received from the workers and stakeholders (especially, the immediate managers).

It is obvious that all outputs cannot be assessed precisely in a quantitative form but maybe in a qualitative one. A more realistic approach would be the use of linguistic assessments instead of numerical values by means of linguistic variables, that is, variables whose values are not numbers but words or sentences in a natural or artificial language (Xu, 2008). Thus, the outputs are assessed using the linguistic variables and then they are quantified using the fuzzy numbers corresponding to the linguistic variables. These linguistic variables and their corresponding fuzzy numbers are shown in Table VI.

4.1.4 Quantification of job outputs. In this step, the outputs of knowledge work are quantified according to Table VII. Rows and columns of Table VII represent job outputs and measures, respectively. First, the weight of each output is normalized using the following equation:

$$w_a^N = \frac{w_a}{\sum_{j=1}^{n} w_j}, \quad (6)$$

where $w_a^N$ is the normalized weight of the output $a$. 

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Total Weight</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$w_1$</td>
<td>$b_{11}$</td>
</tr>
<tr>
<td>2</td>
<td>$w_2$</td>
<td>$b_{12}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$b_n$</td>
</tr>
</tbody>
</table>

Table V. The required information to calculate the weight of each output.
\(a_{jk}\) is the fuzzy number obtained from assessing the output \(j\) according to the measure \(k\). 
\(R_k\) is the quantified value of the measure \(k\) which is given by the following equation:

\[
R_k = \sum_j a_{jk} \cdot w_j^N.
\]  

(7)

Since there are nine measures for the assessment process, we have \(1 \leq k \leq 9\) in Equation (7). Equation (7) is valid as long as all outputs are assessed according to all measures. However, in the real situations, considering the nature of the job, we may not be able to assess some job outputs according to all measures. So, we define \(E = \{j \mid a_{jk} \text{ can be defined}\}\) and \(R_k\) is calculated by the following equation:

\[
R_k = \frac{\sum_{j \in E} a_{jk} \cdot w_j}{\sum_{j \in E} w_j} \quad 1 \leq k \leq 9.
\]  

(8)

4.2 The quantification of job inputs

As was mentioned before, the inputs of the model for knowledge worker productivity are competencies of knowledge workers including knowledge, skills and abilities (KSAs). These inputs are in line with the Personal performance model proposed by Shields (2007).

<table>
<thead>
<tr>
<th>Fuzzy number</th>
<th>Linguistic variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.75, 1, 1)</td>
<td>Very good</td>
</tr>
<tr>
<td>(0.5, 0.75, 1)</td>
<td>Good</td>
</tr>
<tr>
<td>(0.25, 0.5, 0.75)</td>
<td>Average</td>
</tr>
<tr>
<td>(0, 0.25, 0.5)</td>
<td>Poor</td>
</tr>
<tr>
<td>(0, 0, 0.25)</td>
<td>Very poor</td>
</tr>
</tbody>
</table>

Table VI. Linguistic variables and their corresponding fuzzy numbers

<table>
<thead>
<tr>
<th>Outputs</th>
<th>Weight</th>
<th>Quantity (1)</th>
<th>Timeliness (2)</th>
<th>Quality (3)</th>
<th>Efficiency (4)</th>
<th>Costs and/or profitability (5)</th>
<th>Effectiveness (6)</th>
<th>Customer satisfaction (7)</th>
<th>Innovation/creativity (8)</th>
<th>Learning/Process improvement (9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(w_n)</td>
<td>(a_{n1})</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
<tr>
<td></td>
<td>(m)</td>
<td>(w_m)</td>
<td>(a_{n1})</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
</tbody>
</table>

Table VII. Quantified values of knowledge work measures
In the literature related to the job analysis, these competencies are categorized into three groups of KSAs. To calculate knowledge worker productivity, KSAs should be quantified. Throughout the quantification process, the worker’s KSAs are assessed and quantified with respect to the KSAs needed for the job. As shown in Figure 3, four steps are taken in order to quantify the job inputs.

4.2.1 Identification of the required competencies for the job. A worker-oriented job analysis revolves around the characteristics of the worker who performs the work. These qualifications fall into three general groups: KSAs.

Knowledge is the existence in memory of a retrievable set of technical facts, concepts, language and procedures directly relevant to job performance. Skill is the developed or trained capacity to perform tasks that require the use of tools, equipment or machinery. Abilities involve the relatively enduring capacity to obtain skills or knowledge, and to carry out tasks at an acceptable level of proficiency (Primoff, 1975; Madani, 2007; Khorasani and Hasanzadeh, 2008).

Worker-oriented analysis can be performed by several methods such as Job Element Method (JEM) and PAQ (Madani, 2007; Khorasani and Aidy, 2010). In this paper, we use JEM. This method focuses on work behaviors and the results of these behaviors rather than more abstract characteristics. An element in JEM is a combination of behaviors and associated evidence. Elements are named through terms commonly used in the workplace rather than terms developed by psychologists. For example, “the behavior of acting in a dependable fashion, evidenced by punctuality, commendations for dependability and a record of doing exactly what is required by the job, is an element termed Reliability” (Primoff and Eyde, 1988, p. 807). These job elements are actually KSAs. Elements include a wide range of behaviors such as cognition, skills and mental issues (Primoff, 1975; Madani, 2007; Khorasani and Aidy, 2010).

JEM is usually conducted by a professional analyst, who serves as the project leader, and a team of six SMEs, who are usually job incumbents and supervisors. There are usually two sessions that take 3–5 h each. During the first session, SMEs perform brainstorming and rate a list of elements that the analyst collects. The outcome of the session is a list of job elements and ratings of them. The analyst, then, leaves the session to analyze the ratings and obtains preliminary results. Then, the analyst returns to the SMEs for the second session, in which the results of the first session are put to some particular use, such as developing a test, creating a performance measure or designing a training program (Primoff, 1975; Madani, 2007).

4.2.2 Determining the importance of required competencies for the job. In JEM, we have to perform analysis on KSA. Each of the elements should be assessed considering the following measures and the status described in Table VIII (Primoff, 1975; Madani, 2007):

1. B – barely acceptable workers: this measure requires a decision on whether barely acceptable workers must possess an element to do the job. Barely acceptable workers are those who are just scraping by. If they were any worse, they would not be qualified to do the job.

2. S – superior workers to be picked out: this measure asks SMEs to describe how well the element can distinguish the superior worker. For example, it might be helpful to think about things that might earn a worker some kind of reward from his or
her employer. The main point is not merely that the superior workers possess the
class characteristic but that, in addition, the characteristic distinguishes the superior
worker from other workers.

(3) T – trouble likely if not considered: this measure requires SMEs to determine if it is
going to cause any troubles when the element is not considered as a requirement for
new employees.

(4) P – practical: asks SMEs whether job applicants are likely to possess a specific
element. How many job applicants will qualify considering the chosen element?

TV is an index which is calculated for each element based on this information. 0, ✓
and + equal 0, 1 and 2, respectively. TV was developed to compare the abilities of job
applicants. It can be used to determine the importance of KSAs. TV is obtained using the
following equation:

$$TV = (S - B - P) + ((S \times P) + T).$$

The value of TV represents the importance of the KSA related to the job.

4.2.3 Determining the competency level of knowledge workers. KSAs are determined
through the first step of the input quantification process. It is evident that a knowledge
worker does not have the same amount of proficiency when it comes to each of the KSA.
Thus, it is necessary to determine the knowledge worker’s level of proficiency in each of
the KSA in order to quantify them. The main subject areas are determined for each job element
(KSA). For example, as shown in Table IX, applied statistics as a job element can be divided

<table>
<thead>
<tr>
<th>Element No.</th>
<th>Job element</th>
<th>Barely acceptable workers (B)</th>
<th>Superior workers to be picked out (S)</th>
<th>Trouble likely if not considered (T)</th>
<th>Practical (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0: almost none have</td>
<td>✓: some have</td>
<td>+: all have</td>
<td>✓: very important</td>
<td>✓: valuable</td>
<td>✓: some trouble</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>Skills</th>
<th>Ability</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Matter No.</th>
<th>Subject area</th>
<th>Level 1: to some extent not familiar with the subject</th>
<th>Level 2: to some extent familiar with the subject</th>
<th>Level 3: to completely familiar with the subject</th>
<th>Level 4: to completely familiar with the subject and able to use it to accomplish the assigned tasks</th>
<th>Level 5: not only able to use the subject in order to accomplish the assigned task, but also, able to pass the concept of the subject on to others and train them</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ANOVA</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Multidimensional analysis</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Time series</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *Cells that correspond to the specified level of each subject area

Table VIII. Assessment worksheet based on measures of JEM

Table IX. The table for the evaluation of subject areas related to knowledge, skills and abilities
into ANOVA, multidimensional analysis and time series and the knowledge worker’s level of proficiency in each of these subject areas is evaluated. Since knowledge and skills are different in nature, they are evaluated using different methods. In order to evaluate the subject areas related to knowledge and skills, five levels are considered (Khorasani and Hasanzadeh, 2008; Khorasani and Aidy, 2010):

1. Level 1: not familiar with the subject;
2. Level 2: to some extent familiar with the subject;
3. Level 3: completely familiar with the subject;
4. Level 4: completely familiar with the subject and able to use it to accomplish the assigned tasks; and
5. Level 5: not only able to use the subject in order to accomplish the assigned task, but also able to pass the concept of the subject on to others and train them.

Based on the first three levels, subject areas related to knowledge are evaluated; and subject areas associated with skills are evaluated with respect to all five levels. Moreover, subject areas related to ability are evaluated based on five levels covering the range from too weak (1) to very well (5) shown by $a_i$ for the $i$th job element.

The evaluation of subject areas related to KSAs is carried out based on the knowledge worker’s opinion and under the supervision of his/her manager(s). Knowledge worker is provided with a table similar to Table IX and he/she determines his/her level of proficiency in each subject area while considering an expert opinion given by his/her manager(s).

4.2.4 Quantification of job inputs. The two previous steps gave insight into the importance of each job element and knowledge worker’s level of proficiency in them. Based on this information, the quantification is performed in this step. First, knowledge worker’s score in each job element is determined. To do this, the following variables are defined:

- $d_{ij}$: the corresponding score of the $j$th subject area related to the $i$th job element;
- $d_i$: knowledge worker’s score in the $i$th job element; and
- $n_i$: the number of subject areas related to the $i$th job element.

Knowledge worker’s score is calculated using the following equation:

$$d_i = \frac{\sum d_{ij}}{n_i}.$$  \hspace{1cm} (10)

Then, the corresponding weight of each job element is determined based on the following equation:

$$TV_{Ni} = TV_i + \frac{2}{6}.$$  \hspace{1cm} (11)

where $TV_i$ and $TV_{Ni}$ are the weight of the $i$th job element and normalized weight of the $i$th job element, respectively.
Finally, the knowledge worker’s level of proficiency in KSAs can be calculated using the following equations while considering maximum values for the scores related to KSAs:

\[
S_K = \frac{\sum_{i \in K} TV_i^N \cdot \min(d_i, 3)}{\sum_{i \in K} TV_i^N}, \tag{12}
\]

\[
S_S = \frac{\sum_{i \in S} TV_i^N \cdot \min(d_i, 4)}{\sum_{i \in S} TV_i^N}, \tag{13}
\]

\[
S_A = \frac{\sum_{i \in A} TV_i^N \cdot a_i}{\sum_{i \in A} TV_i^N}, \tag{14}
\]

where \(K\), \(S\) and \(A\) are the set of job elements related to KSAs, respectively. Moreover, \(S_K\), \(S_S\) and \(S_A\) are the knowledge worker’s level of proficiency in KSAs, respectively; and can be considered as the quantified job inputs.

4.3 Productivity measurement using DEA

DEA is a non-parametric technique used to measure the relative efficiencies of a set that consists of decision-making units (DMUs) which generate multiple outputs with respect to several inputs. DEA was first introduced by Charnes et al. (CCR model) (1978). Afterwards, other researchers, such as Banker et al. (1984) (BCC model), made improvements to this technique. Nowadays, DEA is applied to various problems (Singh et al., 2000; Yu and Lee, 2013; Bulak and Turkyilmaz, 2014). Several approaches have been proposed to formulate the linear programming problem for DEA. One of the common ways is to solve a sequence of the following linear programming problems:

\[
\text{min } E_n, \tag{15}
\]

Subject to:

\[
\sum_{j=1}^{N} w_{j}x_{kj} - E_{n}x_{kn} \leq 0 \quad k = 1, \ldots, K, \tag{16}
\]

\[
\sum_{j=1}^{N} w_{j}y_{ij} - y_{ii} \geq 0 \quad i = 1, \ldots, I, \tag{17}
\]

\[
w_{j} \geq 0 \quad j = 1, \ldots, N. \tag{18}
\]

Considering this formulation, there are \(N\) DMUs in the sample that produce \(I\) different outputs (\(y_{in}\) shows the observed amount of output \(i\) for DMU \(n\)) with respect to \(K\) different inputs (\(x_{kn}\) is the observed amount of input \(k\) for DMU \(n\)). The \(w_{j}\)s are weights applied to the \(N\) DMUs. Solving the \(n\)th linear program results in weights which determine the most efficient method to produce outputs for the \(n\)th DMU. \(E_{n}^*\) is the efficiency score for the \(n\)th DMU. It is the smallest number \(E_{n}\) which fulfills the three sets of constraints mentioned above.
In order to obtain a complete set of efficiency scores, this problem needs to be solved \( N \) times (once for each DMU in the sample).

In this research, the application of DEA technique could be either input oriented or output oriented. It means that, considering the managerial policies, output-oriented approach can be taken when we aim to improve performance indices of knowledge workers (maximum output), while the current level of proficiency in KSAs remains unchanged (constant input). On the other hand, in order to obtain the same amount of output (constant output) when the least level of proficiency in KSAs is provided (minimum input), we can adopt the input-oriented approach.

It must be mentioned that the relationship between input and output in knowledge work is not constant returns to scale; so, we consider a variable-returns-to-scale (VRS) model where any multiples of inputs could produce the same, smaller or larger multiples of outputs. In this paper, we use the BCC model due to its ability to deal with the VRS case.

It is sensible to keep the cumulative number of inputs and outputs as low as possible, in order to enhance the discriminatory power of the model (Pesenti and Walter, 1996). Considering the nature of the measures introduced in Section 4.1.3, we can classify them into four main groups in this paper:

1. Group 1: efficiency, quantity, timeliness, costs and/or profitability;
2. Group 2: effectiveness;
3. Group 3: quality, customer satisfaction; and
4. Group 4: innovation/creativity, learning/process improvement.

Based on Groups 1–4, outputs of the model are defined as follows:

- \( y_1^j \): the corresponding output value of Group 1 for knowledge worker \( j \);
- \( y_2^j \): the corresponding output value of Group 2 for knowledge worker \( j \);
- \( y_3^j \): the corresponding output value of Group 3 for knowledge worker \( j \); and
- \( y_4^j \): the corresponding output value of Group 4 for knowledge worker \( j \).

Considering Section 4.1.4, these outputs are calculated using the following equations:

\[
y_1^j = \frac{R_1 + R_2 + R_5 + R_6}{4}, \quad (19)
\]
\[
y_2^j = R_4, \quad (20)
\]
\[
y_3^j = \frac{R_3 + R_7}{2}, \quad (21)
\]
\[
y_4^j = \frac{R_8 + R_9}{2}, \quad (22)
\]

Inputs of the model are defined as follows:

- \( x_1^j \): the level of proficiency of knowledge worker \( j \) in knowledge (SK);
- \( x_2^j \): the level of proficiency of knowledge worker \( j \) in skills (SS); and
- \( x_3^j \): the level of proficiency of knowledge worker \( j \) in abilities (SA).

In the standard DEA model, all inputs and outputs are crisp. As it was mentioned before, outputs related to knowledge worker productivity are deemed to follow fuzzy logic.
Therefore, it is necessary to devise an appropriate fuzzy DEA model in order to obtain productivity index and ranking for knowledge workers.

Several approaches have been made to solve fuzzy DEA. In this paper, we use the method developed by Soleimani-damaneh et al. (2006). In their research, they discussed shortcomings of various methods in computational and theoretical aspects; and proposed a method that does not have any of the mentioned weaknesses.

First, LR-fuzzy numbers must be defined. A LR-fuzzy number \( \tilde{a} \) is defined using the following membership function:

\[
\tilde{a} = \begin{cases} 
L\left(\frac{m-x}{\beta}\right) & \frac{m-\beta}{\gamma} \leq x \leq m \\
1 & m \leq x \leq m + \gamma \\
R\left(\frac{m-x}{\gamma}\right) & m \leq x \leq \bar{m} + \gamma 
\end{cases}
\]  

(23)

where \( L \) and \( R \): \([0, 1] \rightarrow [0, 1] \), with \( L(0) = R(0) = 1 \) and \( L(1) = R(1) = 0 \), are non-increasing, continuous shape functions. The LR-fuzzy number can be shown by \( \tilde{a} = (m, \bar{m}, \beta, \gamma) \), and \( m, \bar{m} \) is the peak of \( \tilde{a} \).

In this method, we assume that for each input and output, \( x_{ij} \) and \( y_{rj} \) are LR-fuzzy numbers and share the same shape function. This means that:

\[
\begin{align*}
x_{ij} &= \left(\tilde{x}_{ij}, L_{ij}, R_{ij}\right) \\
y_{rj} &= \left(\tilde{y}_{rj}, L_{rj}, R_{rj}\right)
\end{align*}
\]

(24)

This assumption is not highly restrictive. If all the inputs and outputs of DMUs (knowledge workers) are trapezoidal or triangular fuzzy numbers, this assumption can be fulfilled. In order to solve the model, the following change of variables is performed:

\[
\begin{align*}
\psi_i &= \int_0^1 L_i^{-1}(x)dx, & k_i &= \int_0^1 R_i^{-1}(x)dx \\
\phi'_r &= \int_0^1 L'_r^{-1}(x)dx, & k'_r &= \int_0^1 R'_r^{-1}(x)dx \\
\hat{x}_{ij} &= \tilde{x}_{ij} - \beta_{ij}\psi_i + \bar{x}_{ij} + \gamma_{ij}k_i \\
\hat{y}_{rj} &= \tilde{y}_{rj} - \beta_{rj}'\phi'_r + \bar{y}_{rj} + \gamma_{rj}'k'_r
\end{align*}
\]

(25)

\( \hat{x}_{ij} \) is indeed the sum of the corresponding lower and upper bounds of all \( \alpha \)-cuts of \( \tilde{x}_{ij} \) obtained by integrating over \( \alpha \). Similarly, \( \hat{y}_{rj} \) is the sum of the corresponding lower and upper bounds of all \( \alpha \)-cuts of \( \tilde{y}_{rj} \) given by integrating over \( \alpha \).

Considering the standard DEA model and the change of variables, we can obtain the following model:

\[
\begin{align*}
\text{min} & \quad E_k = \theta_k, \\
\text{s.t.:} & \quad \sum_{i=1}^m \lambda_{ik} \hat{x}_{ij} \leq \theta_k \hat{x}_{ik}, & i = 1, 2, 3,
\end{align*}
\]

(27)
It can be seen that, the defuzzified model is similar to the standard DEA model and can be solved using common commercial softwares.

5. Case study

In order to demonstrate the applicability of the proposed method, a case study was carried out at an Iranian organization in electrical power industry. Seven knowledge workers were chosen (as shown in Table X) and their productivity was measured.

Seven job descriptions were first provided by the organization. The job description standard model in this company includes tasks and requirements regarding KSA. The available information on tasks of each knowledge worker was used to quantify the output. Furthermore, a questionnaire was prepared for each knowledge worker based on KSAs extracted from job descriptions and the five levels of KSA. The questionnaires were filled by knowledge workers. Then, we consulted the supervisors and confirmed the stated levels. The process of reviewing the job and obtaining the required data was rather insignificant (it took nearly an hour to complete the process). On the other hand, the process of contacting knowledge workers, filling and reviewing the questionnaires, was much more time consuming (it took a day to complete this process). As a part of the validation process, 13 managers were provided with the obtained results.

First, the level of proficiency in KSAs is calculated for each knowledge worker. Table XI shows the obtained values.

Then, each output of the knowledge worker productivity model was assessed according to nine measures introduced in Section 4.1.3. The obtained results were categorized into four groups, introduced in Section 4.3, and the corresponding output values were calculated using Equations (19)–(22). These values can be seen in Table XII.
Finally, the productivity index and ranking for all seven knowledge workers were determined by solving the DEA model. Table XIII provides the obtained values. It can be inferred from Table XIII that:

- Two of the knowledge workers have the highest productivity among others. These two people have (almost) lower level of proficiency in KSAs compared to other knowledge workers. Thus, one of the main reasons behind their high productivity is their large output indices despite their relatively low input values. In other words, they made best of their low KSAs and provided (almost) the same amount of output compared to other knowledge workers. This has resulted in their high productivity.

- The two knowledge workers, who ranked 5th in productivity, have (almost) the highest input values. Their job outputs are not high enough compared to the input values. Afrazeh has introduced three groups of factors that have impact on productivity: Will, Can and May (Afrazeh et al., 2003). So, in order to analyze the circumstances that these two knowledge workers have experienced, we need to consider the factors in all three groups and do not rely only on individual aspects (Will and Can).

5.1 Results validation
Considering the significance of results validation, it must be noted that inaccuracies in productivity measurement are acceptable if the level of inaccuracy remains constant over time. The measures are mostly important for tracking the trends, not quantifying empirical data (Thomas and Baron, 1994). Therefore, it is recommended to focus on the trends and comparison of knowledge workers with respect to the obtained results. In this research, results validation is based on the comparison of knowledge workers’ rankings. First, 13 managers were asked to rank the knowledge workers under study. These managers were completely familiar with the knowledge workers and had enough ability to compare the productivity of knowledge workers. Then, the rankings provided by the managers were

<table>
<thead>
<tr>
<th>Row</th>
<th>Job</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>0.55</td>
<td>0.8</td>
<td>0.93</td>
<td>0.44</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>0.63</td>
<td>0.88</td>
<td>0.99</td>
<td>0.51</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>0.24</td>
<td>0.49</td>
<td>0.71</td>
<td>0.18</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>0.46</td>
<td>0.71</td>
<td>0.89</td>
<td>0.52</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
<td>0.7</td>
<td>0.95</td>
<td>0.99</td>
<td>0.49</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>0.42</td>
<td>0.67</td>
<td>0.92</td>
<td>0.29</td>
</tr>
<tr>
<td>7</td>
<td>G</td>
<td>0.45</td>
<td>0.7</td>
<td>0.9</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Table XII. The corresponding output values of Groups 1–4 for each knowledge worker.

<table>
<thead>
<tr>
<th>Row</th>
<th>Job</th>
<th>DEA index</th>
<th>Final rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>E</td>
<td>0.88</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>0.81</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>G</td>
<td>0.77</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>0.72</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>D</td>
<td>0.72</td>
<td>5</td>
</tr>
</tbody>
</table>

Table XIII. The productivity index and ranking for all seven knowledge workers.
compared to the ones given by the proposed method. In order to carry out model validation, two statistical hypotheses were considered as follows (the level of significance is 0.05 for all statistical tests).

5.1.1 The average value of rankings provided by managers for each knowledge worker equals the ranking given by the proposed method. In this case, the average value of rankings for each of seven knowledge workers provided by 13 managers (decision makers of the organization) is calculated and, then, the obtained values are compared with the ones given by the proposed method. Since we are dealing with few ranking data which are not continuous in nature, a non-parametric test such as Mann–Whitney U test can be used. The null hypothesis of this test is that both groups share the same distribution, i.e., the corresponding rankings of knowledge workers productivity given by the proposed method share the same distribution with average rankings provided by the managers. The results shown in Tables XIV and XV were obtained using IBM SPSS software.

Considering Table XII, the p-value is 0.710 and since it is greater than the level of significance the alternative hypothesis is rejected. Thus, both groups come from the same population.

5.1.2 The corresponding rank of each knowledge worker provided by each manager equals the corresponding rank given by the proposed method. In this case, we aim to know whether all the decision makers of the organization and our proposed method yield results from the same distribution when it comes to the ranking of knowledge workers productivity. Following the approach taken by Pasha Sharifi and Najafi Zand (1991), we employed Kruskal–Wallis test which is an extension of the Mann–Whitney U test to three or more samples. The results depicted in Tables XVI and XVII were obtained using IBM SPSS software.

It can be seen in Table XVII that p-value is 1. It is greater than the level of significance; so, we cannot reject the null hypothesis. Consequently, the samples originate from the same distribution.

### 6. Conclusion

In this paper, we proposed a new method to measure the productivity of knowledge workers based on the general concepts and characteristics of knowledge work. In order to develop this method, we first determined the appropriate principles and viewpoints. Then, inputs and outputs of knowledge work were identified. Moreover, we described the quantification

<table>
<thead>
<tr>
<th>No.</th>
<th>n</th>
<th>Mean rank</th>
<th>Sum of ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>KWP</td>
<td>1</td>
<td>7</td>
<td>7.07</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7</td>
<td>7.93</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>KWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mann–Whitney U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilcoxon W</td>
</tr>
<tr>
<td>Z</td>
</tr>
<tr>
<td>Asymp. sig. (two-tailed)</td>
</tr>
<tr>
<td>Exact sig. (2 x (one-tailed sig.))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mann–Whitney U</th>
<th>21.500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilcoxon W</td>
<td>49.500</td>
</tr>
<tr>
<td>Z</td>
<td>−0.386</td>
</tr>
<tr>
<td>Asymp. sig. (two-tailed)</td>
<td>0.700</td>
</tr>
<tr>
<td>Exact sig. (2 x (one-tailed sig.))</td>
<td>0.710</td>
</tr>
</tbody>
</table>
processes of job inputs and outputs. Since knowledge work is involved with several inputs and outputs and due to the fuzzy nature of the outputs, we employed fuzzy DEA to calculate productivity. It can serve as the basis for the comparison among relative productivity of knowledge workers. Furthermore, the obtained results could form a basis to examine the productivity trend of each knowledge worker over different periods of time. A case study was discussed to demonstrate the applicability of the method and the validity of the obtained results.

Moreover, the developed method can serve other purposes such as determining important job’s tasks and job outputs; identifying main job requirements (main KSAs needed to perform the job); and determining initial training programs for each job. These could help managers of organizations design jobs and identify necessary training courses better than before.

Future research in this area includes:

- Other viewpoints and approaches concerned with knowledge worker productivity could be explored to develop new quantitative models. In this case, the most important initiatives include designing models for productivity measurement at the team level and organizational department level.

- In addition to productivity measurement, other steps related to management of knowledge worker productivity (productivity analysis, productivity improvement planning and implementation of productivity improvement solutions) could be designed based on the proposed method in this research.

- It is essential to develop measurement methods that are in line with other applications of knowledge worker productivity measurement such as capacity planning and simulation of knowledge worker’s performance.
References


Primoff, E.S. (1975), “How to prepare and conduct job element examinations”, Personnel Research and Development Center (PRDC), United States Civil Service Commission (Bureau of Policies and Standards), Washington, DC.


Further reading


### Appendix

**Corresponding author**

Jalil Heidary Dahooie can be contacted at: heidaryd@ut.ac.ir

---

#### Table AI.

The analysis of the developed methods

<table>
<thead>
<tr>
<th>Cell</th>
<th>Knowledge worker productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Cells correspond to the fulfilled criteria were shadowed in black
Abstract

Purpose – The purpose of this paper is to investigate the productivity of rice production by decomposing the growth of total factor productivity (TFP) into four components: technological change, scale effects, technical and allocative efficiencies.

Design/methodology/approach – This study employed an econometric approach to decompose TFP growth into four components: technological change, technical efficiency, allocative efficiency and scale effect. Unbalanced panel data used in this study were surveyed in 1994, 2004 and 2014 from 360 rice farming operations. The model used the stochastic frontier transcendental logarithm production technology to estimate the technology parameters.

Findings – The results indicate that the primary sources of TFP growth were technological change and allocative efficiency effects. The contribution of technical efficiency was low because it grew sluggishly.

Research limitations/implications – This study has several shortcomings, such as very low $R^2$ and the insignificant elasticity of labour presented in the findings. Another limitation is the limited time period panel covering long interval, which resulted in unbalanced data.

Practical implications – The government should improve productivity growth by allocating more areas for rice production, which enhances the scale and efficiency effects and adjusting the use of capital and material inputs. Extension services should be strengthened to provide farmers with training on improved agronomic technologies. This action will enhance technical efficiency performance and lead to technological progress.

Social implications – As Indonesian population is still growing at a significant rate and the fact that rice is the primary staple food for Indonesian people, the productivity of rice production should increase continually to ensure social security at a national level.

Originality/value – The productivity growth is decomposed into four components using the transcendental logarithm production technology based on farm-level data. The measure has not been conducted previously in Indonesia, even in rice-producing countries.

Keywords Agriculture, Technological change, Decomposition of productivity, Econometric approach, Efficiency performance, Scale effect

1. Introduction

Indonesian rice agriculture still faces a challenge of increasing demand for food as a result of population growth. Some analysis (e.g. Timmer et al., 2010; Bourgeois and Kusumaningrum, 2008) shows that the total rice demand in Indonesia will decline in coming decades as a consequence of per capita consumption that falls faster than the rate of population growth.

Rice production in Indonesia has slightly increased, as shown in Figure 1. The increase in production is mostly due to the moderate rise in productivity as the sluggish increase in total land cultivation. Since 2004, the cultivated land has slightly increased because of agricultural revitalisation (Mariyono, 2009b).

Relying on imported rice to support food security is one of the alternatives. With a population of more than 250 million (BPS, 2016), this choice is too risky, however. The International market for rice is narrow, which is supplied by only a few rice-producing...
countries. Once Indonesia will start relying on imported rice to guarantee national food security, it will substantially influence international price and stock of rice. In case there is a shortage in the global market, it has a potential turmoil in Indonesia since Indonesian people currently rely on rice as the primary staple food. Meeting this challenge will require continually increasing productivity. Yet, evidence suggests that productivity growth is slowing and the availability of land for future expansion is limited. Thus, the improvement of productivity growth is a wise alternative for fulfilling the demand for food.

This paper aims to estimate the productivity growth of rice agriculture to determine what factors drive it and formulate suitable policies related to the findings. The next part of this section reviews the importance of agricultural productivity growth in the economic development. The literature review section provides brief results of recent studies of productivity growth and their policy implications. The methodology section proposes a method which relaxes some assumptions and anticipates to offer a complete decomposition of productivity growth. Result and conclusion follow.

The productivity of outputs including agriculture is essential in the economy. It has been stated that the importance of productivity growth is that “Productivity isn’t everything, but in the long run it is almost everything. A country’s ability to improve its standard of living over time depends almost entirely on its ability to raise its output per worker (Krugman, 1992, p. 9)”. Regarding agricultural sector, there are linkages between increases in farm productivity and poverty reduction. There are multiple pathways due to which the increase in agricultural productivity is able to alleviate poverty, which includes real income changes, job creation, rural non-farm multiplier effects and food prices effects (Schneider and Gugerty, 2011).

When agriculture retains a significant share of employment and gross domestic product (GDP), then improving agricultural productivity is likely to provide a more substantial influence on poverty than in economies where it has dropped to a smaller portion of employment (Martin, 2016). Gollin and Rogerson (2014) showed that a 10 per cent increase in agricultural TFP leads to a 20 per cent decrease in the fraction of the population engaged in subsistence agriculture; moreover, it leads to a considerable welfare improvement, comparable to the effects of increasing consumption levels for everyone in the economy by 62 per cent.

Agricultural growth reduces poverty and could benefit the poor. The poorer households do not only retain fewer farm assets but also have less productive farm implements and face other socioeconomic constraints in factor-product markets, and thus will be benefitted less from growth. The increase in the productivity of farmers is beneficial in improving the
The agricultural sector still serve as a crucial source of livelihood for a significant proportion of people in the rural areas. Classical economic development theories and recent empirical evidence suggest that agricultural productivity growth is a key condition for rural economic transformation and poverty alleviation. The agricultural productivity is the main factor behind much of the agricultural expansion (Wickramasinghe, 2017).

The productivity gains in agriculture are commonly more effective in the global poverty alleviation than equivalent-sized productivity gains in industry or services. This means that accumulative GDP by identical amounts through the improvement of productivity in different sectors would slash poverty most if the productivity achievement is in agriculture. Even when the size of the productivity gain in agriculture is not adjusted for the lower share of agriculture in the majority of developing countries, poverty alleviation through improvements in agriculture is frequently on equivalence with the gains from equally sized productivity gains in other sectors (Ivanic and Martin, 2018).

In the long run, it can be stated that improving the rice productivity through many efforts leads to the development of the respective countries. Rakotoarisoa (2011) discovered that in high-income countries, the levels of rice productivity have been significantly higher than in the rest of the world. The average rate of productivity growth in low-income rice-producing and importing countries has been the lowest. For the groups of rich countries, developing countries exporting rice, and developing nations importing rice, the productivity growth rates were 1.7, 1.1 and 0.8 per cent, respectively.

The world starts with over-population, and the presence of climate change has exaggerated such condition. The world faces a serious over-population problem, and the effects exist throughout the globe, especially in Africa and Asia (Pimentel, 2012). The current rate of increase in population is 1.2 per cent which means a doubling time of 58 years (Population Reference Bureau, 2011). Enhancement of agricultural productivity has been the world’s important precaution against a repetitive Malthusian crisis – where the needs of a growing population exceed the ability of human and resources to supply food. Over the past 50 years, global gross agricultural output has more than tripled in volume, and productivity growth in agriculture has enabled food to become more abundant and cheaper. Nevertheless, food prices have been rising since around 2001. This has renewed concerns about the speed of agricultural productivity growth. If productivity growth slows, then more resources – land, labour, energy, fertilisers and other inputs – would be required to fulfil the rising demand and the cost of foods (Fuglie and Wang, 2012).

An implication of slowing productivity growth in agriculture would likely have consequences for global food and environmental security (Baldos and Hertel, 2014). Agricultural productivity is indeed significant; it ensures the food security of a nation (Costa et al., 2013). Productivity growth is important for sustainable rice production. If the goal is to have highly productive, efficient and sustainable agricultural systems, we must manage them precisely with a lot of knowledge embedded in every decision to make along each growth cycle (Dobermann, 2017).

A substantial role for agricultural productivity as a driver of structural change is the evidence (McArthur and McCord, 2017). The increase in agricultural productivity is remarkable because it has many substantial effects. The World Bank (2008) reports that agrarian sector is three times more efficient in reducing poverty than other major economic areas. Based on this background, it is clear that agricultural productivity measures provide a vital indicator of the performance of a country’s agricultural sector. It has long been, and now there exists a lot of literature on agricultural productivity measurement. The aims of most productivity studies have been to monitor the performance of the agricultural sector, including its efficiency, in order to make comparisons across industries and countries and to help policymakers formulate policies to enhance productivity.
2. Literature review

Considering the importance of productivity, many recent studies have been conducted using various models. Regardless the models used, the main objective of the studies is to investigate the sources of and factors determining the productivity of various sectors such as agriculture (Fuglie et al., 2017), banking (Jreisat et al., 2017, 2018; Sharma and Dalip, 2014), energy (Li et al., 2017), industry (Benzaquen, 2018) and tourism (Assaf and Tsonias, 2018), as well as multi sectors (Ivanic and Martin, 2017). The significant findings have been used to propose appropriate formulations of policies, and eventually, the improvement of the productivity is expected to provide the projected impact on the economy. In the agricultural sector, many studies of productivity growth also have been conducted all over the world. The studies investigate at various levels including inter-country, regional, national and provincial levels. Based on the data and commodity, the studies also have been conducted using aggregate- and firm-level data, and based on multi products and single particular products.

This section provides a brief review of the recent studies focussing on the productivity of agricultural sector, both in developed and developing countries. At a macro level, Trindade and Fulginiti (2015) paid attention to the agricultural productivity growth related to innovation. Their study shows that spending on education and research and development, as well as irrigation infrastructures, is important in understanding differential performance across regions of South America. For the last period 2000–2009, there is no evidence of a slowdown in agricultural productivity growth in South America. During the period, agricultural production increased annually by 3.91 per cent, and the productivity grew from 2.16 per cent during the 1990s to 2.29 per cent in the 2000s. The positive productivity growth was identified as the advantage of innovations in progressive economies in addition to those resulting from the investments in research and development programs. Anik et al. (2017) focussed on agricultural productivity in South Asia. Their study finds that the primary drivers of agricultural productivity growth were the endowments resulting from improvement in natural, human and technological capitals. But financial capital and crop diversification provide opposite effects. A suggested policy should include land and tenure reforms intended at amalgamating farm operation size and smooth operation of the land rental market to improve natural and human capital. Investments in education and agricultural research and development to enhance human capital and technological capitals boost agricultural productivity growth in the regions. Fuglie et al. (2017) also showed that the growth in agricultural productivity in the USA was identified from the investment in research and development capital, especially by public institutions, as a prime driver. However, the agricultural innovation system in the USA is experiencing an insightful change. The three studies have mentioned that investments in research and development to find innovation in agriculture play a significant role in boosting agricultural productivity. Chen et al. (2008) also showed that the primary determinants of technical progress, which was the main source of productivity, were the agricultural tax cut, public investment in research and development, infrastructure, as well as mechanisation while market reform, education and disaster mitigations are associated with efficiency improvement. De los Santos-Montero and Bravo-Ureta (2017) showed that an intervention programme on natural resource management positively impacted the two dimensions of productivity, i.e., technical change and technical efficiency. This finding contributed to the understanding of how an environmental intervention programme can also increase the income of poor farm households through increases in productivity.

Productivity is also associated with trade policies. Rakotoarisoa (2011) showed that the gap in the level of supports between rich and emerging countries affected the variation of rice productivity across countries. The high levels of protection and subsidy for rice agriculture in rich countries and taxation on rice producers in developing countries amplified the gap of rice productivity. But reducing the support gap following the
implementation of the agreement has narrowed the productivity gap. Differences in agriculture infrastructure such as irrigation and input delivery also influenced the rice-productivity gap. Different level of openness and high import barriers also widened the rice-productivity gap between developed and developing countries. Chen et al. (2008) found that agricultural tax-cut was identified as one of the main determinants of technical progress that contributed to agricultural productivity.

Many studies try to find the sources of productivity growth in agricultural sector. Hoang (2011) analysed the agricultural productivity in several countries and showed that production technology progressed, but there were reductions in technical efficiency, scale efficiency and residual mix efficiency levels. This suggests that different combinations of inputs and outputs drove productivity change. In the United States, technical change was the main driver (Ang and Kerstens, 2017). In Nigeria, improvement in technical efficiency rather than technical progress had contributed to the observed acceleration of the growth rate (Adedeji et al., 2016). Moreira and Bravo-Ureta (2016) provided evidence that productivity of dairy farming gained through technical efficiency improvements was limited. Conversely, the average technological progress was rather high. This study suggests that additional investments in research and the subsequent adoption of improved technologies would have a positive effect on the productivity growth. The dairy farming exhibits decreasing returns to size, implying that costs of production rise as farm size increases. The findings also revealed that farm size is not associated with productivity growth for the dairy farms.

Ogundele and Okoruwa (2014) found that scale efficiency was the only source that made a significant contribution to the agricultural productivity growth in Nigeria. The result was constrained by the inefficient use of resources and inability of farmers to minimise the cost of production which resulted in low technical efficiency. Consequently, achieving agricultural transformation in Nigeria will require more efforts to increase the technical efficiency of the farmers and this can only be achieved through the efficient use of productive inputs. Chen et al. (2008) also realised that the source of productivity growth is technical progress and that the regional disparities in productivity growth worsen over time. A study by Rodriguez and Elasraag (2015) shows that allocative efficiency component is the greatest source in determining the negative performance of the productivity. The inputs were not allocated in the correct proportions, i.e., the input combination that minimises the cost of production was not chosen. This relates to the characteristics of the misuses of inputs and the workers. Should the farms are small, and the workers are inadequately trained, this might prevent the inputs from being used adequately and in the correct proportions.

Wondemu (2016) examined the significance of farm efficiency improvement in enhancing the total factor productivity (TFP) of small-scale farmers in Ethiopia and found that small-scale farming was characterised by increasing returns to scale and economies of scope. Even though farmers exhibited efficiency improvement, significant technical, scale, scope and mix inefficiencies persisted. There exists a potential for increasing farm output by improving the economic efficiency alone. Improvement in technical efficiency and economies of scope provide a substantial increase in productivity in short to medium terms. In Africa, a study shows a result of the productivity components, that is, only technical change is positive. This study proposes that credit should be made available to farmers in Egypt, Morocco, Algeria and Tunisia. Furthermore, agricultural development programs should be implemented in South Africa and Senegal to improve the productivity in these countries (Mohamed et al., 2016).

Agronomic technologies are identified as one of the indirect sources of agricultural productivity through technical efficiency improvement and scale effects. Suphannachart (2013) revealed that one of the positive determinants of agricultural productivity in Thailand was the adoption of high-yielding rice varieties. It is found by Kondo (2015) that the direction of rice varietal improvement has shifted from high-yield to high-quality...
orientations in Korea as well as Japan. Rice productivity stagnated in both countries, whereas income disparity between agricultural and non-agricultural sectors in Korea has been growing in contrast to Japan given the excess supply of rice due to a decrease in rice consumption per capita with economic growth. Large-scale production of rice per household was barely observed in either Japan or Korea (Kondo, 2015). Applying balanced inorganic and organic fertilisers (Amanullah et al., 2016) and hybrid varieties (Peykani et al., 2008) could improve rice productivity.

Concerning the effect of climate change, Chun et al. (2016) suggested that combinations of adaptation strategies such as irrigation, fertiliser management practices and adjustment of planting dates may be able to reduce the negative impacts of climate change on rice substantially. Li et al. (2017) suggested that the government should continue to encourage and support technical invention. The reason is that the innovation of renewable and clean energy increases the speed of technological progress that is faster than its efficiency. The change makes a significant contribution to TFP.

There is a study that evaluates components of agricultural productivity and associates with appropriate policies. According to the findings, the growth of TFP in agricultural production had specific distinguishing features. For instance, it was susceptible to agricultural policy, technical progress coincided with the loss of technical efficiency, it strongly relied on technical progress, and advanced technologies were inadequately disseminated. Agricultural production should focus on the dissemination and innovation of advanced technology to make full usage of the input elements. Nevertheless, although eliminating the external environmental factors and random factors can clearly illustrate the changing trends and characteristics of productivity, in practice it is not an easy task to remove the influence of these variables. Thus, to improve the efficiency of the agricultural production, some measures are required. These include promoting the standardisation of land transfer, operating the farmland to a proper scale, strengthening the construction of the rural infrastructure and increasing investment in human capital (Song et al., 2016).

Kijima et al. (2012) suggested a hypothesis that one of the major constraints on the growth in productivity of rice farming in sub-Saharan Africa is the absence of effective extension systems. Wondemu (2016) recommended that policy measures on encouraging more farmers to participate in extension services and improve their schooling will have a significant impact on the technical efficiency of the farmers. Enhancing the competitiveness of the markets for goods and other factors as well as reducing transaction costs, which influence farm-gate relative prices, promote a more optimal cropping and input use pattern and enable farmers to realise economies of scope. Identifying the factors that give rise to the suboptimal application of fertiliser in particular and taking the necessary measures to address them will have significant and immediate impacts on farm productivity. Since farmers currently operate at a level of scale that is below the most productive scale, promoting a competitive land market and further land consolidation would have a significant impact on farm productivity. Combined improvements in mix and scale efficiencies would have a cumulative effect on productivity.

The key issue in the studies on agricultural productivity is the identification and assessment of sources of productivity growth. Knowing whether productivity is arising from technical and allocative efficiency gains, technological progress or returns to scale can fundamentally influence investment, performance management and policy assessment. There are various methods used in assessing the total productivity growth. Every method holds assumptions such as full efficiency in terms of technical and allocative, Hicks-neutral technological change, and constant returns to scale. In fact, such assumptions are likely to be unrealistic.

There is a long debate about the option of the most suitable methods surrounding the literature on the decomposition of the technology-based productivity measures (Felipe and
McCombie, 2003). Every technique has advantage and disadvantage in modelling fields of efficiency and productivity analysis, which depends on the specific objectives (Karagiannis and Lovell, 2016). One should pick the right productivity index or indicator depending on their ultimate goals; either measuring TFP or measuring the local change of the technology (Kerstens et al., 2018). Despite the debate, the productivity analysis in recent years has received increased recognition in several fields of performance evaluation, and more applications are to continue. As long as the objectives are clear and the method is theoretically justifiable, there is no constraint to carry out studies on the subject. This present paper provides a rigorous theoretical framework for defining the core components of productivity growth while ensuring that the concepts are well defined using relatively unrestricted assumptions. Thus, it is different in the following respects. First, this study relaxes the assumptions that producers are allocative and technically efficient. Second, this study allows non-neutral technological change and non-constant returns to scale. Third, the data used in this study are expressed in physical quantities.

3. Research methodology
3.1 Theoretical framework
TFP provides a wider theory of agricultural productivity than methods that compare the output to just one input, like land or labour. TFP is estimated by totalling the land, labour, capital and material inputs used in agriculture and compare growth in total inputs with growth in total output of agricultural products. TFP is said to be improved if the total output grows faster than total inputs; in other words, TFP encompasses the average productivity of all inputs employed in the production of all agricultural commodities (Fuglie and Wang, 2012).

This paper defines TFP as the portion of output unexplained by the quantity of inputs used in production. Its value represents a performance of how efficiently and intensely the inputs are utilised in production (Fulginiti et al., 2004). When the economy is competitive, i.e. production is economically efficient and exhibit constant returns to scale, then factors’ income share will be equal to the corresponding output elasticities and the TFP growth is equal to technical progress. Considering that Indonesian rice producers were inefficient (Mariyono, 2014), the production technology was not Hicks-neutral (Mariyono, 2009a; 2015; Mariyono et al., 2010) and might not exhibit constant returns to scale.

Following the primal method proposed by Kumbhakar and Lovell (2000), this study decomposes TFP growth into technological change, changes in technical and allocative efficiency and scale effects. A stochastic production function is used to decompose productivity growth. The deterministic kernel production frontier with input $X$ and input $Z$, technology parameter vector $\beta$, time trend $t$ as a proxy for technological change, and output-oriented technical inefficiency $u \geq 0$ is represented as:

$$ Y = f(X, Z, t, \beta) \exp(-u). $$

(1)

Technical efficiency is expressed as $\psi = Y/(f(X, Z, t, \beta) = \exp(-u) \leq 1$, which allows it to vary over time. A primal measure of the rate of change in technical efficiency is given as:

$$ \dot{\psi} = \frac{\ln \exp(-u)}{\partial t} = \frac{\partial u}{\partial t} $$

(2)

$\dot{\psi}$ can be interpreted as the rate at which a producer shifts towards or away from the production frontier, keeping everything else constant. Taking the log and totally differentiating Equation (1) and then differentiating with respect to $t$ yields:

$$ \dot{Y} = \frac{\partial \ln f(\cdot)}{\partial t} + \frac{\partial f(\cdot)}{\partial X} \frac{\partial \ln X}{\partial t} + \frac{\partial f(\cdot)}{\partial Z} \frac{\partial \ln Z}{\partial t} + \frac{\partial \ln \exp(-u)}{\partial t}, $$

(3)
where $\dot{Y} = (\partial \ln Y / \partial t)$ is output growth, $f(\cdot) = f(X)$, $Z$, $t$, $\beta$ is the deterministic kernel of the stochastic production frontier, $(\partial \ln f(\cdot) / \partial t) = \dot{A}$ is the rate of technological change, $(\partial \ln X / \partial t) = \dot{X}$ is the growth rate of input $X$, $(\partial \ln Z / \partial t) = \dot{Z}$ is the growth rate of input $Z$, $(\partial f(\cdot) / \partial X)(X/f(\cdot)) = \theta_X$ is the output elasticity with respect to input $X$, $(\partial f(\cdot) / \partial Z)(Z/f(\cdot)) = \theta_Z$ is the output elasticity with respect to input $Z$, and $(\partial \ln \exp[-u] / \partial t) = -(\partial u / \partial t) = \dot{\psi}$ is the rate of change in technical efficiency.

The growth of TFP is defined as the growth of output which is unexplained by the growth of inputs (Fulginiti et al., 2004), that is, $\dot{t} = Y - S_X X - S_Z Z$. By substituting the expression for $\dot{Y}$ in Equation (3) into $\dot{t}$, then it yields:

$$\dot{t} = \dot{A} + (\theta - 1) \left( \frac{\theta_X}{\theta} \dot{X} + \frac{\theta_Z}{\theta} \dot{Z} \right) \left( \frac{\theta_X}{\theta} S_X \right) \dot{X} + \left( \frac{\theta_Z}{\theta} S_Z \right) \dot{Z} + \dot{\psi},$$

(4)

where $\theta = \theta_X + \theta_Z$ is the scale elasticity that provides a primal measure of return to scale of the production frontier; $S_X = (W_x X/W_x X + W_Z Z)$ and $S_Z = (W_Z Z/W_x X + W_Z Z)$ is input cost share of $X$ and $Z$, respectively; $W_x$ and $W_z$ are the prevailing price of input $X$ and $Z$, respectively. The notation of $(\theta_X / \theta)$ and $(\theta_Z / \theta)$ is called normalised output elasticity with respect to input $X$ and $Z$, respectively, because the sum of the normalised output elasticity with respect to both inputs will be exactly equal to unity for any returns to scale production technology. The effect of returns to scale is represented by $(\theta - 1)$, which will be positive, negative, or zero if the production technology exhibits increasing, decreasing or constant returns to scale, respectively. It is also reasonable to say that allocative efficiency of input use will be reached if normalised output elasticity with respect to all inputs is equal to the share of the cost of the respective inputs. This is equivalent to a condition where the marginal rate of technical substitution is equal to the price ratio of inputs.

Equation (4) shows that the TFP growth is decomposed into four components: technological change, scale effect, allocative efficiency and technical efficiency. If there is no technological change or change in the production frontier over time, the component of technological change will be zero. If the technical efficiency is time-invariant, the decomposition implies that the change in technical efficiency does not affect the TFP growth. If the production technology exhibits constant returns to scale over time, the scale effect is zero. Allocative inefficiency is represented by the deviations between normalised output elasticity and share of input cost. When all gaps equal zero, the use of inputs is allocative efficient and there is no effect on the TFP growth. Finally, if all inputs are constant or grow at a zero rate, the scale and allocative efficiency component will be zero, and the TFP growth will be driven solely by technological change and technical efficiency. Therefore, when agricultural industry is always allocative and technically efficient and the production technology exhibits constant returns to scale, the TFP growth is equal to the rate of improvement in technology or technological change.

3.2 Data and variables
This study used a database established from a longitudinal survey of 360 farm operations in 1994, 2004 and 2014. The samples were proportionately collected from five rice bowl regions across Indonesia: West and East Java, Lampung, Wes Nusa Tenggara, North and South Sulawesi.

Table I shows the description of observations related to the data. Observations were typically made in 1994 and 2004. One of major causes of the few observations in 2014 was some farmers were no longer cultivating rice as some farmers had passed away and the family members discontinued to cultivate rice. Farmers in North Sulawesi were no longer interviewed in 2014 because of various reasons. The data were an unbalanced panel because of this lack of continuity.
The number of variables observed in the data varies widely across survey years. This is because the survey pursued to accommodate any variations in farming which were spatially and temporally specific. For example, in some regions, voluntary labour was common during early planting and harvesting seasons, but it was unusual in other regions. Specific types of fertilisers were not adopted in some regions but were always implemented in others. Given the variations of samples, the data were aggregated to avoid a problem of missing data [1].

The description and measurement of aggregated inputs and output variables are given in Table II. Tables III and IV show the summary statistics for key variables across time and region, respectively. On average, production increased over time. Land cultivation and the use of materials and agrochemicals also increased over time. On the other hand,

<table>
<thead>
<tr>
<th>Year</th>
<th>Java</th>
<th>Lampung</th>
<th>South Sulawesi</th>
<th>West Nusa Tenggara</th>
<th>North Sulawesi</th>
<th>By year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>126</td>
<td>75</td>
<td>84</td>
<td>36</td>
<td>21</td>
<td>342</td>
</tr>
<tr>
<td>2004</td>
<td>122</td>
<td>79</td>
<td>51</td>
<td>34</td>
<td>21</td>
<td>307</td>
</tr>
<tr>
<td>2014</td>
<td>63</td>
<td>54</td>
<td>36</td>
<td>18</td>
<td>0</td>
<td>171</td>
</tr>
<tr>
<td>By region</td>
<td>311</td>
<td>208</td>
<td>171</td>
<td>88</td>
<td>42</td>
<td>820</td>
</tr>
</tbody>
</table>

**Source:** Authors’ analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production (Y)</td>
<td>Total production of rice in a growing season</td>
<td>kg</td>
</tr>
<tr>
<td>Land (L)</td>
<td>Total land cultivated to rice in a growing season</td>
<td>ha</td>
</tr>
<tr>
<td>Labour (N)</td>
<td>Total labour (family, voluntary and hired labour) used for all farming stages</td>
<td>person-working day</td>
</tr>
<tr>
<td>Capital (K)</td>
<td>Capital (large tractors and cows) mainly used in land preparation in a growing season</td>
<td>cow-equivalent working day</td>
</tr>
<tr>
<td>Materials (M)</td>
<td>Total materials (seed, water irrigation and green manure) used in a season of rice farming</td>
<td>monetary value (000 IDR)</td>
</tr>
<tr>
<td>Chemicals (C)</td>
<td>Fertilisers (Urea, Phosphate, Ammonium Sulphate and Potassium Chloride) and Pesticides (solid and liquid formulations) applied to rice farming</td>
<td>monetary value (000 IDR)</td>
</tr>
</tbody>
</table>

**Notes:** Materials and chemicals were measured in terms of monetary value because of variations in measurement unit and quality. The monetary values of materials and chemicals were based on 1993 constant price using agricultural price index. IDR is Indonesian Rupiah

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>Mean</td>
<td>1,856</td>
<td>0.552</td>
<td>41.69</td>
<td>8.19</td>
<td>35.50</td>
<td>52.41</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1,751</td>
<td>0.534</td>
<td>35.34</td>
<td>17.12</td>
<td>39.25</td>
<td>54.71</td>
</tr>
<tr>
<td>2004</td>
<td>Mean</td>
<td>2,121</td>
<td>0.621</td>
<td>78.99</td>
<td>1.45</td>
<td>58.58</td>
<td>64.90</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2,866</td>
<td>0.604</td>
<td>59.79</td>
<td>2.59</td>
<td>60.88</td>
<td>71.21</td>
</tr>
<tr>
<td>2014</td>
<td>Mean</td>
<td>3,445</td>
<td>0.882</td>
<td>57.77</td>
<td>0.44</td>
<td>81.32</td>
<td>254.9</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>3,972</td>
<td>0.957</td>
<td>70.29</td>
<td>2.26</td>
<td>109.76</td>
<td>2,257</td>
</tr>
</tbody>
</table>

**Note:** See Table II for units of measurement

**Source:** Author’s analysis
there was a considerable reduction in capital use over time. Labour increased almost
twofold in 2004 (from 1994), but then decreased in 2014. It is important to note that the
standard deviation of each variable in each region was relatively high, indicating a
considerable variation in these variables.

3.3 Empirical model
This study uses transcendental logarithmic (translog) production technology introduced by
Christensen et al. (1973)[2]. The stochastic frontier translog production technology is
specified as:

\[
\hat{Y} = \beta_0 + \beta_1 L + \beta_2 N + \beta_3 K + \beta_4 M + \beta_5 C \\
+ \beta_{11} LL + \beta_{12} LN + \beta_{13} LK + \beta_{14} LM + \beta_{15} LC \\
+ \beta_{22} NN + \beta_{23} NK + \beta_{24} NM + \beta_{25} NC \\
+ \beta_{33} KK + \beta_{34} KM + \beta_{35} KC \\
+ \beta_{44} MM + \beta_{45} MC \\
+ \beta_{55} CC + \beta_{12} tL + \beta_{13} tN + \beta_{14} tK + \beta_{15} tM + \beta_{15} tC \\
+ \beta_5 t + \beta_{11} t^2 + v + u,
\]

where \( Y, L, N, K, M \) and \( C \) is production, land, labour, capital, materials and chemicals,
respectively; \( t \) is time trend that captures technological progress; \( \beta_s \) are coefficients
of production technology to be estimated; \( v \) is error terms representing other factors beyond
this study; \( u \) is level of inefficiency; and tilde (\( \sim \)) over the variables represents logarithmic
operation. The full translog production technology captures more precise estimates of
technical efficiency, which are subsequently used for calculating the decomposition of
productivity growth of rice production. Given the estimated parameters in the production
function, the mean elasticities of output with respect to inputs are formulated as follows.

The elasticity of production with respect to land is expressed as:

\[
\theta_L = \frac{\partial \hat{Y}}{\partial L} = \beta_1 + 2\beta_{11} L + \beta_{12} N + \beta_{13} K + \beta_{14} M + \beta_{15} C + \beta_{11} t.
\]

The elasticity of production with respect to labour, capital, materials and chemicals can be
conducted by taking the same step as that of land specified in Equation (6). Note that the
output elasticities of all inputs are not constant and are dependent on the levels of all inputs.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Java</td>
<td>Mean</td>
<td>2,482</td>
<td>0.804</td>
<td>60.74</td>
<td>7.344</td>
<td>77.01</td>
<td>158.4</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2,965</td>
<td>0.802</td>
<td>47.79</td>
<td>17.010</td>
<td>76.00</td>
<td>168.1</td>
</tr>
<tr>
<td>Lampung</td>
<td>Mean</td>
<td>2,477</td>
<td>0.583</td>
<td>61.65</td>
<td>1.035</td>
<td>26.68</td>
<td>62.88</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>3,989</td>
<td>0.616</td>
<td>67.57</td>
<td>2.991</td>
<td>32.30</td>
<td>79.70</td>
</tr>
<tr>
<td>South Sulawesi</td>
<td>Mean</td>
<td>2,445</td>
<td>0.655</td>
<td>68.02</td>
<td>2.811</td>
<td>64.97</td>
<td>81.59</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2,574</td>
<td>0.530</td>
<td>62.59</td>
<td>9.060</td>
<td>90.55</td>
<td>84.62</td>
</tr>
<tr>
<td>W. Nusa Tenggara</td>
<td>Mean</td>
<td>1,341</td>
<td>0.265</td>
<td>42.39</td>
<td>2.669</td>
<td>23.71</td>
<td>44.29</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1,318</td>
<td>0.230</td>
<td>44.12</td>
<td>3.409</td>
<td>30.37</td>
<td>45.15</td>
</tr>
<tr>
<td>North Sulawesi</td>
<td>Mean</td>
<td>1,284</td>
<td>0.557</td>
<td>30.56</td>
<td>2.544</td>
<td>32.63</td>
<td>37.01</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1,645</td>
<td>0.657</td>
<td>24.45</td>
<td>4.399</td>
<td>36.16</td>
<td>43.61</td>
</tr>
</tbody>
</table>

Note: See Table II for units of measurement
Source: Author’s analysis
as well as the time trends. The mean output elasticities are then evaluated at the average level of each input and time period.

The rate of technological change is defined as the percentage change in output due to an increment of time in which all inputs are kept constant, that is:

$$\dot{Y} = \frac{\partial \tilde{Y}}{\partial t} = \beta_{11} \tilde{L} + \beta_{12} \tilde{N} + \beta_{13} \tilde{K} + \beta_{14} \tilde{M} + \beta_{15} \tilde{C} + \beta_t + 2\beta_{tt} t$$  \hspace{1cm} (7)

The rate of technological change consists of two components: biased and pure technological changes. The biased technological change is shown by $\beta_{11} \tilde{L} + \beta_{12} \tilde{N} + \beta_{13} \tilde{K} + \beta_{14} \tilde{M} + \beta_{15} \tilde{C}$. The biased technological change component is producer specific. The pure technological change is shown by $\beta_t + 2\beta_{tt} t$, which will be constant, increasing or decreasing at a constant rate according to whether $\beta_t$ is zero, positive or negative, respectively.

Following Cornwell et al. (1990), the temporal pattern of technical efficiency is modelled as a quadratic function of time, that is:

$$\dot{c} = a_0 + a_1 t + a_2 t^2$$  \hspace{1cm} (8)

where $\alpha$ is coefficients to be estimated and tilde ($\sim$) over the variables represents logarithmic operation. The rate of change in technical efficiency is specified as:

$$\dot{c} = \frac{\partial \tilde{c}}{\partial t} = a_1 + 2a_2 t.$$  \hspace{1cm} (9)

Input growth is considered to vary over time. The rate of land growth is estimated using the expression:

$$L = a_0 \exp \{ (\delta_1 + \delta_2 t) t \},$$  \hspace{1cm} (10)

where $a_0$ is a proxy for the initial level of land, $\delta$ is coefficients to be estimated, and $\delta_1 + \delta_2 t$ represents a non-constant rate of input growth. Taking logarithm of both right- and left-hand sides gives a log linear expression:

$$\tilde{L} = \tilde{a}_0 + \delta_1 t + \delta_2 t^2,$$  \hspace{1cm} (11)

where tilde ($\sim$) over the variables represents logarithmic operation. Equation (11) can be easily estimated using a method of ordinary least square regression. The rate of land growth is obtained as:

$$\frac{\partial \tilde{L}}{\partial t} = \delta_1 + 2\delta_2 t.$$  \hspace{1cm} (12)

The growth rates of labour, capital, materials and chemicals can be estimated using the same steps as that of land specified in Equations (10)–(12).

4. Results and discussion

The estimation of stochastic frontier production technology is presented in Table V. The estimated production technology shows that $\gamma$ was highly significant. This means that there was a significant deviation of actual outputs from the potential outputs caused by technical inefficiency. In other words, the average translog production technology was significantly less than that at the frontier.

By following the estimation of frontier production technology, the four components of TFP are calculated. The first component is technological change, which consists of non-neutral and pure effects. The second, third and fourth components are, respectively,
Table V. Estimated frontier production function of rice, specified in Equation (5)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>T-Value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y )</td>
<td>8.054 (7.11***), 1.228 (0.62***), 0.650 (0.62***), -0.060 (−0.06)</td>
<td>0.006, 0.006, 0.006, 0.006</td>
<td>1.37, 1.37, 1.37, 1.37</td>
<td>**, ***, ***, respectively</td>
</tr>
<tr>
<td>( L )</td>
<td>-1.00 (−0.56)</td>
<td>0.006</td>
<td>-17.3</td>
<td>**</td>
</tr>
<tr>
<td>( N )</td>
<td>0.006 (0.006)</td>
<td>0.006</td>
<td>1.00</td>
<td>ns</td>
</tr>
<tr>
<td>( K )</td>
<td>0.006 (0.006)</td>
<td>0.006</td>
<td>1.00</td>
<td>ns</td>
</tr>
<tr>
<td>( M )</td>
<td>0.006 (0.006)</td>
<td>0.006</td>
<td>1.00</td>
<td>ns</td>
</tr>
<tr>
<td>( C )</td>
<td>0.006 (0.006)</td>
<td>0.006</td>
<td>1.00</td>
<td>ns</td>
</tr>
<tr>
<td>( L/N )</td>
<td>-0.62 (−0.62)</td>
<td>0.006</td>
<td>-10.0</td>
<td>**</td>
</tr>
<tr>
<td>( L/K )</td>
<td>-1.21 (−1.21)</td>
<td>0.006</td>
<td>-20.0</td>
<td>**</td>
</tr>
<tr>
<td>( L/M )</td>
<td>-2.06 (−2.06)</td>
<td>0.006</td>
<td>-33.3</td>
<td>**</td>
</tr>
<tr>
<td>( L/C )</td>
<td>0.44 (0.44)</td>
<td>0.006</td>
<td>7.4</td>
<td>**</td>
</tr>
<tr>
<td>( N/K )</td>
<td>0.56 (0.56)</td>
<td>0.006</td>
<td>9.3</td>
<td>**</td>
</tr>
<tr>
<td>( N/M )</td>
<td>-0.88 (−0.88)</td>
<td>0.006</td>
<td>-15.0</td>
<td>**</td>
</tr>
<tr>
<td>( K/M )</td>
<td>-2.18 (−2.18)</td>
<td>0.006</td>
<td>-37.3</td>
<td>**</td>
</tr>
<tr>
<td>( K/C )</td>
<td>-0.08 (−0.08)</td>
<td>0.006</td>
<td>-1.3</td>
<td>ns</td>
</tr>
<tr>
<td>( M/C )</td>
<td>-2.12 (−2.12)</td>
<td>0.006</td>
<td>-37.3</td>
<td>**</td>
</tr>
<tr>
<td>( C/K )</td>
<td>4.81 (4.81)</td>
<td>0.006</td>
<td>81.3</td>
<td>***</td>
</tr>
<tr>
<td>( C/M )</td>
<td>7.31 (7.31)</td>
<td>0.006</td>
<td>122.2</td>
<td>***</td>
</tr>
</tbody>
</table>

Notes: Numbers in the parentheses following the coefficients represent the standard error. ns denotes insignificant. ***, **, ***Significant at 10, 5 and 1 per cent levels, respectively.

Source: Author's analysis
the rate of change in technical efficiency; scale effects, which involve output elasticity with respect to each input and the input growth of each respective input; and allocative efficiency, which involves shares in input costs.

The technological change relates to the time trend in the frontier production technology. A joint test for Hicks-neutral technological change and pure technological change was rejected. This indicates that there were movements in production frontiers across time, representing technological progress. The temporal pattern of estimated technical efficiency could be represented as \( \dot{\psi} = 0.616 + 0.044 \cdot t - 0.005 \cdot t^2 \). The joint test for time-invariant technical efficiency showed that \( F_{358}^2 = 3.85 \) and it rejected the constant technical efficiency over time at 5 per cent significance, meaning that technical efficiency was not time invariant. The technical efficiency increased at a decreasing rate, and the rate of change in technical efficiency was expressed as \( \dot{\psi} = (\partial \dot{\psi} / \partial t) = 0.044 - 0.010 \cdot t \). The rate of technological change and change in technical efficiency in each year is given in Table VI.

This finding suggests technological progress of rice agriculture in Indonesia across the periods of 1994, 2004 and 2014. In total, the technological change was positive and increasing. The impressive growth in technological change is an indication that farmers had adopted better technology for rice production, which also explains why the rate of change in technical efficiency was low[3].

The rate of change in technical efficiency in 1994, 2004 and 2014 was 0.033, 0.022 and 0.011, respectively. The rate of change in non-neutral technological change was positive and increasing, i.e. there was input-augmenting technological change. The implication is that technological change led to intensifications in input use over time. The rate of change in the pure technological change was positive and increasing[4]. This indicates that with given levels of input use, rice production increased over time.

Both scale effects and allocative efficiency are relevant to output elasticity with respect to each input. The output elasticity derived from translog production technology was not constant and dependent on the level of each input use. Mean output elasticities calculated at the average level of each input used are shown in Table VII.

Along with the factor input growth, the average output elasticity in each year was used to calculate the scale and allocative effects. Input growth was estimated using a regression of logged input on quadratic time trends. The result of the regression is given in Table VIII.

### Table VI.

<table>
<thead>
<tr>
<th>Year</th>
<th>Biased</th>
<th>Pure</th>
<th>Total</th>
<th>Technical efficiency (( \dot{\psi} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.388</td>
<td>0.591</td>
<td>0.989</td>
<td>0.033</td>
</tr>
<tr>
<td>2004</td>
<td>0.419</td>
<td>1.182</td>
<td>1.601</td>
<td>0.022</td>
</tr>
<tr>
<td>2014</td>
<td>0.438</td>
<td>1.773</td>
<td>2.211</td>
<td>0.011</td>
</tr>
</tbody>
</table>

**Source:** Authors’ analysis

### Table VII.

<table>
<thead>
<tr>
<th>Year</th>
<th>Land</th>
<th>Labour</th>
<th>Capital</th>
<th>Materials</th>
<th>Chemicals</th>
<th>Returns to scale (( \theta ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.721</td>
<td>Insignificant</td>
<td>0.034</td>
<td>0.104</td>
<td>0.001</td>
<td>0.860</td>
</tr>
<tr>
<td>2004</td>
<td>0.097</td>
<td>Insignificant</td>
<td>0.049</td>
<td>0.103</td>
<td>0.088</td>
<td>0.937</td>
</tr>
<tr>
<td>2014</td>
<td>0.743</td>
<td>Insignificant</td>
<td>0.032</td>
<td>0.123</td>
<td>0.192</td>
<td>1.090</td>
</tr>
</tbody>
</table>

**Notes:** The output elasticity was evaluated at the average of all input use (in logarithmic form) in 1994, 2004 and 2014. Labour was insignificant and the elasticity of output with respect to labour was considered to be zero

**Source:** Author’s analysis
As mentioned above, input growth was not expected to be constant over time. All regressions were highly significant in overall tests, despite the fact that some coefficients were individually insignificant. This is because the time series trend consists of just three sets of observations, and the data were unbalanced. These circumstances led to a strong correlation between linear and quadratic trends, resulting in a multicollinearity problem. Severe multicollinearity brings about individual effects that tend to be insignificant since the standard error of the coefficient is very high (Wooldridge, 2003). Since the joint tests showed high significance, these coefficients were then used for calculating the rate of input growth in each year. The rate of input growth of each input is given in Table IX.

On average, the use of inputs increased over time, except for capital which sharply decreased across the periods. Many factors determined the drop in the capitals. One of the possible causes was small land size, which was less likely to use large mechanisation (Mariyono, 2009b). Recently, the large types of machinery in rice agriculture have been replaced with small and mobile ones. The highest rate of positive growth of all inputs was agrochemicals, with an increase of more than 100 per cent across the periods. The use of agrochemicals dropped in 1994 but rose in 2004 and 2014. Labour and material inputs exhibited the same pattern, with initial high rates of growth but declined over the next two periods, eventually became negative in 2014. During the period 1994–2014, Indonesian agriculture underwent many changes/shocks related to central policy and economic crisis. At a farm level, the use of material input and agrochemicals was very sensitive to the shocks, particularly for agrochemicals. The sharp decrease in the capital was supposed to be affected by the economic crisis because the main components/parts of the machinery (e.g. tractor) were imported. During the economic crisis, the exchange rate of Indonesian currency (Indonesian Rupiah, IDR) dropped at almost one-fourth towards US$. Thus, farmers did not use tractors anymore and employed traditional tools such as hand hoe and cow-powered plough. The rate of land growth was continually positive and increasing over time. Despite limited available land, the areas cultivated to rice increased. Since 2004, Indonesian rice production

<table>
<thead>
<tr>
<th>Year</th>
<th>Land</th>
<th>Labour</th>
<th>Capital</th>
<th>Materials</th>
<th>Chemicals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.040</td>
<td>1.359</td>
<td>1.802</td>
<td>0.917</td>
<td>−0.260</td>
</tr>
<tr>
<td>2004</td>
<td>0.170</td>
<td>0.073</td>
<td>−3.062</td>
<td>0.348</td>
<td>1.090</td>
</tr>
<tr>
<td>2014</td>
<td>0.220</td>
<td>−1.213</td>
<td>−7.926</td>
<td>−0.223</td>
<td>2.440</td>
</tr>
</tbody>
</table>

**Table IX.** Growth rate of inputs
the government has revitalised agriculture by allocating more land to the strategic commodities including rice (Mariyono, 2009b). The land cultivation of these sampled farmers grew at an almost similar rate to that of national level, as shown in Figure 1.

The rate of input growth will contribute to the scale and allocative efficiency effects. The scale effect is determined by three components: input growth (as it has been previously discussed), returns to scale (the sum of output elasticity with respect to all inputs), and normalised elasticity (the ratio of output elasticity with respect to each input to the sum of output elasticity with respect to all inputs). As shown in Table VII, the translog production technology of rice agriculture exhibited decreasing returns to scale in 1994 and 2004 and increasing returns to scale in 2014. Overall, however, the production technology exhibited decreasing returns to scale[5]. The normalised elasticity resulting from output elasticity with respect to each input is given in Table X.

The normalised output elasticity of each input exhibited a similar pattern to the output elasticities[6]. The scale effect and its components are given in Tables XI and XII, respectively. The scale effect in the first two points in time was negative. The decreasing returns to scale in those periods has led to negative scale effects. In contrast, the scale effect was positive at the last point in time because of increasing returns to scale.

The last component of TFP growth is the allocative efficiency effect, which constitutes the gap between the normalised output elasticity and shares in input cost. The share of costs is given in Table XIII.

Generally, labour and agrochemicals showed a higher share in the cost of production. In small-scale rice agriculture, this condition is reasonable. Small-scale rice agriculture is usually labour and chemical intensive. Chemicals are used to increase the productivity of

<table>
<thead>
<tr>
<th>Year</th>
<th>Land</th>
<th>Labour</th>
<th>$\frac{\theta_{input}/\theta}{Capital}$</th>
<th>Materials</th>
<th>Chemicals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.838</td>
<td>Insignificant</td>
<td>0.040</td>
<td>0.121</td>
<td>0.002</td>
</tr>
<tr>
<td>2004</td>
<td>0.744</td>
<td>Insignificant</td>
<td>0.052</td>
<td>0.110</td>
<td>0.094</td>
</tr>
<tr>
<td>2014</td>
<td>0.682</td>
<td>Insignificant</td>
<td>0.029</td>
<td>0.113</td>
<td>0.176</td>
</tr>
</tbody>
</table>

Source: Author’s analysis

<table>
<thead>
<tr>
<th>Year</th>
<th>Land</th>
<th>Labour</th>
<th>$\frac{\theta_{input}}{(\theta_{input},\theta)}$</th>
<th>Capital</th>
<th>Materials</th>
<th>Chemicals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.034</td>
<td>Insignificant</td>
<td>0.072</td>
<td>0.111</td>
<td>$-2.9E-4$</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>0.126</td>
<td>Insignificant</td>
<td>$-0.159$</td>
<td>0.038</td>
<td>0.103</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>0.205</td>
<td>Insignificant</td>
<td>$-0.229$</td>
<td>$-0.025$</td>
<td>0.430</td>
<td></td>
</tr>
</tbody>
</table>

Source: Author’s analysis

<table>
<thead>
<tr>
<th>Year</th>
<th>Returns to scale ($\theta$)</th>
<th>($\theta-1$)</th>
<th>$\sum (\theta/\theta_{input})$</th>
<th>($\theta-1)\sum (\theta/\theta_{input})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.861</td>
<td>$-0.140$</td>
<td>0.216</td>
<td>$-0.030$</td>
</tr>
<tr>
<td>2004</td>
<td>0.937</td>
<td>$-0.063$</td>
<td>0.108</td>
<td>$-0.007$</td>
</tr>
<tr>
<td>2014</td>
<td>1.090</td>
<td>0.090</td>
<td>0.380</td>
<td>0.034</td>
</tr>
</tbody>
</table>

Source: Author’s analysis
land, and labour is often more appropriate than machinery such as a tractor. Capital consequently showed a low share of cost in small-scale rice agriculture. The land exhibited the smallest share of the cost because farmers studied here operated rice agriculture on their own land. The cost related to land was the land tax, which was relatively low in rural areas [7]. The shares of land, labour and capital costs tended to increase over time, whereas the shares of agrochemicals and materials tended to decrease. The dynamics of shares of the cost depended on the price of inputs and the level of use of these inputs.

With a positive rate of growth in inputs, the allocative efficiency effect will be positive, negative or zero respectively if the gap between normalised output elasticity with respect to each input minus the cost share of that input is positive, negative or zero. The gap between normalised output elasticity and input cost share is shown in Table XIV. Perfect allocative efficiency was not expected to be found in Indonesian rice agriculture, and consequently, the allocative efficiency affected TFP growth. From Table XIV, the land showed positive gaps over time which means that the use of land was low compared with other inputs. This gap decreased over time due to the increase in land tax. Capital, labour and materials all showed negative gaps, which means that the use of these inputs was economically excessive relative to the land use. The negative gap for capital increased, whereas the negative gap for materials decreased over time and the gap for labour fluctuated. Chemicals showed a negative gap in 1994 and 2004, but became positive in 2014.

We can see that overall allocative efficiency improved. The gaps will have an overall impact on TFP growth only if there is variation in input growth. As shown in Table IX, variation in input growth occurred. The total allocative efficiency effects are given in Table XV. Land and capital had positive allocative efficiency effects. This is because the gap of land was positive and land use grew positively. In 1994, capital had a negative allocative efficiency effect, after which the effect increased considerably. The considerable increase in the allocative efficiency effect was mostly due to a dramatic decrease in capital growth. Since the use of capital was no longer allocative efficient, the negative growth caused an increase in allocative efficiency. For the case of labour and materials, the allocative efficiency effects were negative in the 1994 and 2004, but these effects were increasing. In 2004, the rates of labour and material growth were negative, and at the same time, there was an increase in the cost of labour and materials, resulting in decreases in allocative

\[
S_i = \left( \frac{W_i \times \text{Input}_i}{\sum_i (W_i \times \text{Input}_i)} \right)
\]

<table>
<thead>
<tr>
<th>Year</th>
<th>Land</th>
<th>Labour</th>
<th>Capital</th>
<th>Materials</th>
<th>Chemicals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.044</td>
<td>0.260</td>
<td>0.060</td>
<td>0.280</td>
<td>0.356</td>
</tr>
<tr>
<td>2004</td>
<td>0.012</td>
<td>0.467</td>
<td>0.105</td>
<td>0.217</td>
<td>0.199</td>
</tr>
<tr>
<td>2014</td>
<td>0.123</td>
<td>0.405</td>
<td>0.203</td>
<td>0.124</td>
<td>0.146</td>
</tr>
</tbody>
</table>

**Table XIII.** Share in cost of input use

\[
\frac{\theta_i}{\theta} - S_{\text{Input}_i}
\]

<table>
<thead>
<tr>
<th>Year</th>
<th>Land</th>
<th>Labour</th>
<th>Capital</th>
<th>Materials</th>
<th>Chemicals</th>
<th>$\sum_i \left( \frac{\theta_i}{\theta} - S_{\text{Input}_i} \right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.794</td>
<td>−0.260</td>
<td>−0.021</td>
<td>−0.159</td>
<td>−0.354</td>
<td>1.588</td>
</tr>
<tr>
<td>2004</td>
<td>0.732</td>
<td>−0.467</td>
<td>−0.053</td>
<td>−0.106</td>
<td>−0.105</td>
<td>1.465</td>
</tr>
<tr>
<td>2014</td>
<td>0.550</td>
<td>−0.405</td>
<td>−0.174</td>
<td>−0.111</td>
<td>0.031</td>
<td>1.380</td>
</tr>
</tbody>
</table>

**Table XIV.** Estimated gap between normalised elasticity and share of input cost
efficiency. For the case of labour, the cost increased relatively large due to the significant decrease in labour growth. Regarding agrochemicals, the effect of allocative efficiency was positive and increasing. In 1994, agrochemicals decreased and the gap was negative. In the next two points in time, both growth and the gap were positive. The total effect was positive.

The allocative efficiency effects were considerable. The effects increased over time after beginning with a negative value. This indicates that allocative efficiency effect improved. The allocation of inputs was much more efficient after the recent, perhaps because farmers became more conscious of the incorrect allocation of some inputs.

Table XVI shows TFP growth, which combines growth in technological change, scale effects, allocative efficiency and technical efficiency. In the absolute value, TFP growth was high, particularly for 2014. The most significant contributor to TFP growth was the technological change, followed by the allocative efficiency effect, which came from both allocative efficiency and growth of inputs. With respect to the considerable magnitude of TFP growth, this could be due to the relatively long (ten-year) time periods. On an annual basis, the growth rate was 0.058, 0.172 and 0.437 in 1994, 2004 and 2014, respectively. This finding is in line with Fuglie and Wang (2012), in which TFP has increased considerably, meaning that it has replaced resource expansion and input intensification as the primary sources of growth in world agriculture.

Based on this finding, technological change and allocative efficiency effects were the most significant components of TFP growth for Indonesian rice agriculture. Previous studies on productivity growth, which used stochastic production technology, did not account for allocative efficiency effects. Depending on the level of allocative efficiency and input growth, the estimates of the productivity growth could be either under- or overestimated. In such studies, allocative efficiency effects are unexplained.

This study finds evidence of the impressive growth in TFP. Slow growth in 1994 was perhaps due to government neglect of the agricultural sector at the time (Mellor et al., 2003). The sector has become more central, partly because it was the only sector able to grow in the economic crisis in 1997. After that, the sector has received much more attention from the government (Mariyono, 2009b), resulting in a high TFP growth in 2004 and 2014.

The rate of change in technical efficiency was very small. Therefore, it is realistic to make efforts to increase this component. Enhancing the technological change will be effective if the appropriate new technology is available, and the existing technology has

<table>
<thead>
<tr>
<th>Year</th>
<th>Land</th>
<th>Labour</th>
<th>Capital</th>
<th>Materials</th>
<th>Chemicals</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.032</td>
<td>−0.354</td>
<td>−0.037</td>
<td>−0.146</td>
<td>0.092</td>
<td>−0.413</td>
</tr>
<tr>
<td>2004</td>
<td>0.124</td>
<td>−0.034</td>
<td>0.162</td>
<td>−0.037</td>
<td>−0.114</td>
<td>0.101</td>
</tr>
<tr>
<td>2014</td>
<td>0.168</td>
<td>0.491</td>
<td>1.380</td>
<td>0.002</td>
<td>0.075</td>
<td>2.116</td>
</tr>
</tbody>
</table>

Source: Author's analysis

<table>
<thead>
<tr>
<th>Year</th>
<th>Technological change</th>
<th>Allocative efficiency</th>
<th>Scale effect</th>
<th>Technical efficiency</th>
<th>TFP growth (f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.989</td>
<td>−0.413</td>
<td>−0.030</td>
<td>0.033</td>
<td>0.579</td>
</tr>
<tr>
<td>2004</td>
<td>1.601</td>
<td>0.101</td>
<td>−0.007</td>
<td>0.022</td>
<td>1.717</td>
</tr>
<tr>
<td>2014</td>
<td>2.211</td>
<td>2.115</td>
<td>0.034</td>
<td>0.011</td>
<td>4.371</td>
</tr>
</tbody>
</table>

Source: Author’s analysis
been fully adopted by all farm operators, in other words, if rice agriculture has been technically efficient. As this study has shown that the rice agriculture has not been technically efficient, Shapiro (1983) and Belbase and Grabowski (1985) suggested that efforts to improve technical efficiency may be more cost-effective than introducing new technologies as a means of increasing agricultural productivity. Concerning technology, Hassan et al. (2014) suggested that farmers adopt new technologies or improved ones, as this will help shift farmers production frontier, thus contributing to raise productivity growth to a higher level. It is also recommended that by expanding the scope of research and development, the net value of production and labour could help to raise productivity growth in the country. Prices of production inputs like fertiliser should be part of policy priorities because of its opposite impact.

Efforts to enhance technical efficiency have a direct impact, that is, increases in technological efficiency will directly improve TFP. Technical efficiency can be enhanced by providing appropriate training and extension programmes. As per suggestion of Wondemu (2016), in short to medium terms, a substantial increase in productivity is expected to occur through improving technical efficiency and exploiting economies of scope. In this regard, policy measures on encouraging more farmers to participate in extension services and improve their schooling will have a significant impact on the technical efficiency of the farmers. Luther et al. (2018) and Mariyono (2018) suggested that sending farmers to field schools can enhance agricultural productivity.

Addressing the scale effect, which also varies, requires careful policy formulation. Given the parameters of rice production technology, the scale effects may be improved by adjusting (either reducing or increasing) the use of inputs. Based on the increasing returns to scale of production technology in 2014, it is reasonable to increase the use of land, labour and chemical inputs which have positive normalised elasticity, and to rationally reduce the use of capital and material inputs, which have negative normalised elasticity.

The increase in the use of inputs, however, also influences allocative efficiency. For 2014, the increase in land use leads to increased allocative efficiency, but the increases in other inputs lead to decreased allocative efficiency. It is, therefore, the increase in land use will improve the scale and allocative efficiency effects. The increases in all effects can also be achieved by reducing capital and material inputs. Increases in labour and chemical inputs will lead to opposite impacts on the scale and allocative efficiency effects. The preferable policy is the one that is able to provide the greatest net positive impact.

In the era of a sustainable economy, producing more with less of inputs will be the major challenge in the next two decades. This is particularly true when the production provides an adverse impact on the human health and the environment. Considering that agricultural production leads to an environmental problem, measures to accelerate and sustain TFP growth can be conducted by jacking up investment in research and development, infrastructural facilities and improving input-use efficiency, particularly for environmentally detrimental inputs such as inorganic fertilisers and synthetic pesticides. Note that the synthetic pesticides and inorganic fertilisers should be applied wisely (Mariyono, Kuntariningsih and Kompas, 2018) because the use of agrochemicals has triggered external costs (Mariyono, Kuntariningsih, Suswati and Kompas, 2018). Biotechnology research to address biotic and abiotic stresses should be paid more attention. As suggested by Kumar et al. (2008), given the diminishing trend in public investment in agriculture which requires being reversed, the good option to accelerate TFP growth is carried out by developing appropriate technology, leading to an increase in yield potential, both for irrigated and rain-fed areas. Research problems in the rain-fed unfavourable ecosystems and breaking of the current irrigated yield ceilings are more complex and challenging. To make progress in them will require the mobilisation of the
best of science and the best of scientists in the National Agricultural Research System in partnership mode. This needs higher investment in agricultural research which has been convincingly justified in several studies.

5. Conclusion and policy implication

Indonesian rice agriculture needs to grow substantially in order to keep pace with the rising food needs of the national population that grows at a significant rate. As the main staple food in Indonesia, the increase in rice productivity is very important because it guarantees food security in particular, and it has many considerable effects on economic development and rural poverty alleviation in general. It is clear that productivity measures provide a key indicator of the performance of a country’s agricultural sector, a fact which has long been recognised.

Using an approach which decomposes the TFP growth of rice production into technological change, technical efficiency, scale effect and an allocative efficiency effect, the main sources of the productivity of rice were determined. The rate of growth in TFP was low in 1994, but quite high in 2004 and 2014. The growth rate of the TFP was driven predominantly by an impressive rate of growth in technological change, followed by improvements in the allocative efficiency effect. The high increase in productivity growth from 2004 to 2014 was enabled by a recovery from the economic crisis at the end of the 1990s. Farmers adopted better techniques, and the uses of all inputs were much more allocative efficient.

The government, in collaboration with and private sectors, should improve extension services, allocate more areas of cultivated to rice, enhance the scale and efficiency effects, and adjust the use of capital and material inputs rationally to keep the rice productivity growing. The actions will increase technical efficiency and improve scale and efficiency effects. Big machinery such as large four-wheel tractor can be replaced with hand tractor; big static rice huller can be replaced with mobile one. Increasing land cultivation can be conducted by improving the existing irrigation infrastructures such as irrigation channels and dams, and establishing ones to create more land for rice cultivation both in Java and other islands of Indonesia. With improved irrigation infrastructure, the cropping index of rice will increase. Selected material inputs can be adjusted by adopting good agricultural practices. Research and development programme should continue to produce technology and innovation because technological change is the highest source of TFP of rice production in Indonesia.

Notes

1. Agricultural practice including rice agriculture in Indonesia is common for farmers not to use fertilisers, pesticides, and tractors at all. In the absence of such inputs, there exist positive production. But by definition, if the functional form is a translog production technology, the production with no such input will be zero, and econometric estimation will be impossible because ln0 is mathematically undefined. The problem can be handled by summing the quantities of individual fertilisers and replacing the zero level of input use with a small positive value (Trewin et al., 1995). A study by Villano and Fleming (2006) uses this method to compare with a quadratic functional form rather with a translog model. The results show that both methods provide very similar measures of output elasticity with respect to inputs and estimates of technical efficiency. The translog model, however, provides more precise estimates than the quadratic model as the log-likelihood for the translog model is much greater than that for the quadratic model (and the variance of the technical inefficiency effects in the stochastic frontiers for the translog model is also greater than that for the quadratic model).

functional forms results in a more accurate technical efficiency estimate. More flexible functional forms reduce the error terms ($\epsilon = v-u$) and generate higher estimates of technical efficiency. Considering that a higher rate of efficiency represents a better estimate, the primal approach is also more accurate than the dual approach because 'studies using the primal approach leads to significantly higher technical efficiency estimates than those obtained from dual frontiers' (Thiam et al., 2001, p. 241).

3. Technological progress represents movements in both average and frontier production functions (Jansen and de Londono, 1994). In this case, farmers operate their farms closer to the frontier production, which is increasing over time.

4. A high rate of technological progress with a similar pattern of technological change occurred in rice agriculture in the Philippines (Villano and Fleming, 2006).

5. A formal test for returns to scale was also undertaken. The test shows that the translog production function of rice agriculture did not exhibit constant returns to scale.

6. The important difference between normalised elasticity and output elasticity is that the sum of normalised elasticity is exactly equal to unity.

7. Rice farming is commonly operated in a semi-subsistence manner. It is usual that farmers operated rice farming on their own land, and there was no rental cost for land. This is also the case for family labour where there was no need to pay family labour. If the rental cost of land and wage of family labour were imputed, rice farming would provide a negative profit. Thus, it was more realistic to not impute the rental cost of land and wage of family labour in this analysis.

References


BPS (2016), Statistik Indonesia (Indonesian Statistics), Badan Pusat Statistik, Jakarta.


About the author
Joko Mariyono is a member of academic staff in the Department of Magister Management, Postgraduate Programme, University of Pancasakti, Tegal – Indonesia. He holds PhD in economics from The Australian National University, Canberra. He specializes in rural and development economics, and his research expertise includes econometric and micro-economic modelling, micro-economic impact evaluation of innovation and dissemination of technology, economic sustainability of rural and agriculture development and agribusiness. He also provides consultancy services to research and development institutions in Indonesia and overseas. Joko Mariyono can be contacted at: mrjoko28@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Critical success factors for Kaizen implementation in the Nigerian construction industry

Temitope Seun Omotayo
Scott Sutherland School of Architecture and Built Environment,
Robert Gordon University, Aberdeen, UK

Udayangani Kulatunga
School of the Built Environment, University of Salford,
Greater Manchester, UK, and

Bassam Bjeirmi
Scott Sutherland School of Architecture and Built Environment,
Robert Gordon University, Aberdeen, UK

Abstract

Purpose – Kaizen in construction is a new paradigm stemming out of lean production systems. Construction companies in developing economies, such as Nigeria, have a task to innovate to liquidate in certain cases. With the aid of kaizen, which encompasses the benefit of stakeholder relationship improvement and management, profitability enhancement and delivery of projects to satisfied clients, construction companies can realise expected growth. An exploration of the critical success factors (CSF) and associated drivers within the limits of the scope is essential. The paper aims to discuss these issues.

Design/methodology/approach – Exploratory factor analysis statistical tests categorised the CSF identified in the literature review section. A detailed sampling approach extracted 135 questionnaires under the five-point Likert scale format from a larger population in Nigeria. An exploration of important drivers and iteration of eigenvalues combined with asymptotic significance of the drivers provided the CSF and significant drivers.

Findings – Construction management function, operational efficiency, construction business ethics and construction cost management were the CSF established from the exploratory factor analysis tests. It was confirmed that kaizen can be adopted in Nigerian construction companies with reflections on the principal drivers for the CSF.

Research limitations/implications – The exploration of the CSF and drivers associated with kaizen implementation can be applied to other developing countries with considerations of implementation in terms of cost and time.

Originality/value – The identification of CSF provides ample opportunity for consideration of kaizen in construction companies. The findings of this study are a basis for investigations into cost and time implications of kaizen adoption in construction companies.

Keywords Continuous improvement, Construction, Implementation, Kaizen, Developing economies

Paper type Research paper
Considering the prevailing concomitant drawbacks of ensuring project success within the construction industry, kaizen has a major role to play in terms of managing cost, time and quality (Kaur and Kaur, 2013; Junker, 2010). These drawbacks seem to transcend construction economies of developed and developing countries. Nonetheless, the circumstantial differences between the categories of construction economies in relation to studies on the aforementioned drawbacks have peculiarities which should be investigated when change and innovation are being discussed.

In developing countries such as Nigeria, the need for innovative and modern mechanisms of managing construction companies is currently imperative. The economy of Nigeria is out of recession at the moment and the construction industry requires efficient managerial approaches for an uplift. Productivity of construction companies in Nigeria is at its lowest ebb and waste reduction mechanisms are urgently required (Funso et al., 2016; Guerrini et al., 2011). This is why the exploration of the potential that kaizen has in the current economy is essential. In recent studies of project performance, the aforementioned challenges are evident in the UK and other developed construction economies. Hence, the application of kaizen has a vast transferrable import in undermining the vagaries of project success.

Kaizen is a part of the lean production system which essentially addresses waste reduction incrementally (Jin and Doolen, 2014; Junker, 2010; Khanna, 2009). The benefits of kaizen are improved productivity, waste reduction, profit, client satisfaction, quality enhancement and increased competitive advantage (Kumiega and Vliet, 2008; Macpherson et al., 2015). Although kaizen was originally developed as an integral part of the Toyota production system, understanding and use of kaizen in the construction industry is still new (Cheng, 2018; Singh and Singh, 2018). Unlike lean construction, which has been studied extensively in the built environment research milieu, kaizen has only been researched by very few authors in the construction industry. Some of the few research works carried out in the construction industry were done by Omotayo and Kulatunga (2017), Vivan et al. (2015), Kaur and Kaur (2013), Martin (1993) and Smadi. Over the years, the focus of kaizen researchers has been its application in the areas of manufacturing, food production and the health sector. The potential benefits of kaizen can improve the present condition of small and medium scale enterprises.

The application of kaizen in the construction industry begins with its implementation in the workplace. Therefore, the management of a construction company requires adequate training on kaizen implementation (Dakhli et al., 2017; Singh and Singh, 2018). However, there are some requirements which have to be in place for kaizen to be implemented. In developing countries such as India, China, Brazil and Malaysia, kaizen has been adopted in the manufacturing sector for continuous improvement and productivity enhancement (Shang and Pheng, 2013; Arya and Jain, 2014; Puvanasvaran et al., 2011; Vivan et al., 2015). Studies about continuous improvement in the construction industry will be elucidated in the next section. The case studies of continuous improvement in the construction industry focus on the research methodology adopted by the researcher. The process of implementing kaizen as a tool for incremental reduction of construction cost, quality and profitability enhancement will also be taken into consideration.

2. The potential application of kaizen in the construction industry

Kaizen is concerned with reducing waste in a particular process by incrementally addressing the existing process before they occur or when they occur. The processes here may be in the project planning or construction phase. Although kaizen has not been used in most construction industries, a kaizen model was developed by Vivan et al. (2015) based on an action research strategy using 76 building projects in Brazil. The results showed that the overall cost of the building reduced drastically over the phase of the project. Kaizen was
applied during construction to reduce cost. Kaizen costing is the cost reduction mechanism of kaizen during production. The model developed by Vivian et al. (2015) also eliminated the myth of standardisation for production in construction during the action research process.

Other authors such as Kaur and Kaur (2013), Martin (1993) and Smadi identified the use of kaizen and kaizen costing for offsite manufacturing of building components and construction. The case studies were conducted in a metal industry, a large construction and concrete company by Savolainen (1999), discussed the understanding and adoption of kaizen processes empirically. The findings reveal that the kaizen adoption process is iterative and the speed of implementation differs in these two companies. Omotayo and Kulatunga (2017) developed an IDEF0 model for the application of kaizen in the construction process. Omotayo and Kulatunga (2017) proposed a plan-do-check-act approach in the key construction process phases such as interim valuation, cashflow calculations, monitoring of building material cost, plant and labour. These implementations of the aforementioned construction process were not considered from a case study point of view. Imai (2012) authored a comprehensive approach of how kaizen can be applied in the workplace in the form of gemba kaizen. The continuous improvement approach is a managerial strategy in deciphering key problems within the workplace before they occur. Further studies of gemba kaizen in construction were conducted by Omotayo and Kulatunga (2017) within the confines of business process modelling and notation for improving small and medium scale construction businesses.

The use of new techniques in enhancing productivity of construction companies is crucial in the current economic climate. Consequently, criteria for adopting kaizen must be explored. The critical success factors (CSF) for analysing the adoptable mechanisms for kaizen in the construction economies of developed and developing countries may be similar. However, previous research and application indicated a vast acceptance of kaizen in developing countries.

3. Scope of the study
Researchers in developing countries such as India, Brazil, Malaysia and China have been investigating the application of kaizen in the construction industry. Hence, this research will focus on Nigeria as part of the wider scope of construction countries which are developing. The CSF in this study will be based on managerial and building construction site applications in Nigeria. Construction cost management (CCM) has a major role to play in implementing kaizen. This is due to the enormous benefit of kaizen, which is centred on cost reduction, waste reduction and profitability. In addition to this, cost management in construction has a major influence on planning and executing construction activities in developing countries (Aibinu and Jagboro, 2002; Odeh and Battaineh, 2002). The CSF for kaizen implementation intricately advances the course of struggling construction organisations after application.

Exploring the potential implementable CSF for kaizen in the Nigerian construction industry as a transferrable panacea in construction industries of developing economies is the research aim in this study.

4. Drivers for implementing kaizen in the construction industry
The strategies other researchers have suggested for implementing kaizen in countries around the world can be applied to the Nigerian construction industry. These strategies can be categorised as they relate to the various sections of the kaizen process within a construction industry for the company and construction process. Additionally, the CSF in this section are from third world countries (such as China, Brazil and India) similar to the Nigerian construction industry. Most third world construction industries have similar challenges, such as cost overruns, project financing and small and medium scale
construction companies, which are facing stiff competition from multinational construction companies. The challenges faced in third world construction companies are, therefore, very similar (Ballesteros-Pérez et al., 2015; Guo et al., 2016; Amoatey et al., 2015) and this allows for consideration of strategic transferability.

Shang and Pheng (2013) investigated the challenges facing the implementation of kaizen in Chinese construction companies, by interviewing project managers. The findings show that Chinese construction organisations do not have the culture of exposing the problems within the business because it can lead to losses. The nature of organisations in China is such that minor problems can be perceived as major problems and this leads to covering up minor problems. Kaizen demands the identification of difficulties which can lead to waste and eliminate them (Suárez-Barraza et al., 2011). Shang and Pheng (2013) further stated that most project managers in China decide to compress the duration of the project even though the projects cannot be completed in such an unrealistic time frame. The author also stated that kaizen could not work in this instance because there would be intense pressure to finish a project without considering the quality. The final challenge involves the inadequacy of construction professionals with the pertinent understanding of how kaizen works. It was noted that Chinese construction companies are facing the challenge of few construction professionals with an understanding of kaizen. Although most construction companies envy the Japanese production success stories, they do not have the required kaizen skills.

Borrowing from the experiences of Shang and Pheng (2013) in implementing kaizen in construction companies in China, the CSF are related to identified strategies for the implementation of kaizen as a strategy and kaizen costing in a company. Although there are no direct identified implementation strategies for the Nigerian context, there are basic requirements for kaizen implementation in terms of CSF in developing countries such as Brazil, China and India.

Shang and Pheng (2013), Arya and Jain (2014), Puvanasvaran et al. (2011), Berger (1997), Chukwubuikem et al. (2013) and Magnier-Watanabe (2011) have identified several CSF required for the implementation of kaizen costing in a company. Some of these strategies could be referred to as enablers or CSF required for kaizen costing. These CSF have been categorised as thus:

1. Organisational structure: Magnier-Watanabe (2011) noted that “kaizen required a horizontal organisation structure and opportunistic knowledge acquisition”, kaizen requires a lot of communication and teamwork and a less bureaucratic management structure. This allows a lot of information to flow within the organisation. This fosters improved relationship between the employer and the employees. There is a need for an ad-hoc, collective and innovative system within the organisation for kaizen costing to be effective.

2. Construction process standardisation: there is a need to improve the process of construction to a more standardised method if kaizen costing will be used (Shang and Pheng, 2013). Most construction companies’ modus operandi depends on the nature of the project, organisation policy and regulatory bodies.

3. Government and regulatory bodies: government policies, politics and construction regulatory bodies influence various construction industries around the world. If these bodies do not approve kaizen costing as a means of post-contract cost control, they cannot be used. However, some construction regulatory bodies are very flexible and do not get involved in innovation within construction companies.

4. Contract documentation and procurement: the type of procurement process adopted can affect the cost of the project. A design and build system will be different from the
traditional procurement system. Therefore, contractors can have more resources at their disposal to implement kaizen costing during a project and create their own team. Also, the level of involvement of stakeholders in the project goes a long way during the construction process. Clarity of exclusions and accuracy of estimates can also influence the time available for kaizen activities.

(5) Financial risk management: the quantity surveyors and project managers are mainly involved in managing the financial risk which can arise as a result of price fluctuations, inflation, changes in the building design, variation, claims, theft, fraudulent practices, kickbacks, payment delays, suppliers’ and sub-contractors’ cost. These factors along with the preliminary items of work can affect the financial position of a construction project, thereby affecting the success of kaizen costing.

(6) Communication and teamwork: kaizen costing activities on site cannot involve the cost or project managers alone, but a kaizen team. This team will have to work with every stakeholder, including the suppliers of building materials, clients and sub-contractors. Communication is essential during the implementation of kaizen costing. Therefore, regular site meetings and post-project review meetings are necessary. In most cases where BIM is implemented during the pre-tender process, communication can be easier with other stakeholders.

(7) Decision making: the decision made by the cost or project manager during kaizen implementation and follow-up can affect the overall performance of kaizen costing. The contractor or management function in cost management has to be involved in making some final decisions about claims or litigation that can affect the implementation process of kaizen costing. Shang and Pheng (2013) noted that the construction company needs to see “problems as opportunities” in order to make the best decisions during kaizen activities.

(8) Relationship management: in most situations where claims are raised by the subcontractor or contractor, there is a need for relationship management between the stakeholders. The availability of resources to execute kaizen costing during construction depends on the relationship between the contractor quantity surveyor and the client.

Based on the eight major sections associated with implementing kaizen, the drivers in Figure 1, have been identified.

The categorisation of the CSFs was identified through an extensive literature review search, which identified some cogent drivers when implementing kaizen.

5. Methodology
A survey research strategy was used with a large number of the population involved within the scope of this study. A survey research strategy provides an opportunity to explore and extract large data sets from the existing population (De Vaus, 2013). The CSF for the adoption of kaizen in Nigerian construction companies began with an extensive literature review. This culminated into the development of a Likert scale questionnaire for quantitative questionnaires, targeting building contractors, quantity surveyors and project managers. The research gaps identified in the literature review were modified to provide the focus on the main challenges of developing a good understanding of continuous improvement in Nigerian companies. The opinions and perceptions of key building contractors in the Nigerian construction industry were essential to gain a better understanding of construction industries in Nigeria.
Figure 1. Categories of drivers for the implementation of kaizen in the Nigerian construction industry.
A random sampling technique was adopted for the questionnaire survey. According to Bray and Rees (1995), “random sampling is defined as one for which each measurement or count in the population has the same chance (probability) of being selected”.

In this study, 84 construction companies were contacted for data collection purposes, this is the company sample size. In total, 250 respondents in these companies were chosen as the broader sample size in order to cover one-third of the population. Overall, 250 questionnaires were distributed to the 84 companies, and 135 were returned. The response rate is 54 per cent and this was achieved by following-up the respondents. The questionnaires which were not properly completed were returned to the respondents for amendments.

5.1 Questionnaire survey
A quantitative questionnaire provided relevant details for an exploratory factor analysis addressing the research objective of shaping the CSF based on a five-point Likert scale. The categorised drivers identified in Figure 1 were positioned in the 5-point Likert scale format of “not important” to “extremely important”. According to De Vaus (2013), survey questionnaires target a larger population for the purpose of deductive exploration. Hence, the primordial process of data collection was employed to extract the perceptions of construction stakeholders in Nigeria towards kaizen. The word “kaizen”, was not used as the main word during the data collection, “continuous improvement” was used instead, since the respondents may not have been used to the term “kaizen”, and so a brief participant information sheet was applied to educate the respondents on what continuous improvement in construction means and the associated benefits.

5.2 Exploratory factor analysis
According to Cornish (2007), factor analysis is data reduction through the multivariate method. Pallant (2016) also supported this by noting that factor analysis is a combination of various techniques with steps for reduction of the principal components. The author further advised that principal component analysis and factor analysis are distinct. Both approaches use the correlation pattern to produce a smaller number of linear combinations. Yong and Pearce (2013) stated that the main purpose of factor analysis is to provide a structured pattern, which makes it easier for the researcher to understand the logic behind the relationship. The author also stated that factor analysis could be used for exploratory factor analysis and confirmatory factor analysis. For this study, an exploratory factor analysis was used. The CSF were identified using literature synthesis and the drivers were listed and categorised. The categorisation does not provide the CSF for the implementation of kaizen, but only classified the drivers for the respondents to answer the questionnaires. The process of factor analysis starts with the validity of the cases. There are 135 cases in this study. According to Pallant (2016), 150 cases is the benchmark. However, smaller samples can be considered if the solutions have high loading marker variables above 0.80. Pallant (2016) stated that smaller samples with the good reliability of factor structures and the Kaiser-Mayer-Olkin (KMO) measure of sampling adequacy would have an index of 0.6 for a good analysis. If the KMO is less than 0.6, the factor analysis is not reliable. Therefore, some variables were reduced. The reduction process starts with the reliability of data with KMO. The principal component factor extractor was used for this study; this considers the best factors which reflect the comparison of the variables (Pallant, 2016). Other types of factor analysis extraction are principal factors, image factoring, maximum likelihood factoring, α factoring, unweighted least squares and generalised least squares.

The extraction process is followed by the decision making which can be based on Kaizer’s criterion, parallel analysis and scree plot (Field, 2009). The scree plot was used for this investigation. The scree plot displays the eigenvalues of the factors. The factors
above the elbow were retained. The factor rotation and interpretation was determined after the number of factors has been decided. This presents the pattern of loadings for easier interpretation.

6. Findings
The categorised drivers were tested using the KMO Measure of Sampling Adequacy to assess the adequacy of the data for factor analysis. This is based on the principal components. If the KMO is above 0.6, then the data is unreliable for the analysis. The first stage of the KMO test did not provide the required sampling adequacy. Four drivers were eliminated to give the KMO and Bartlett’s test in Table I. The four factors eliminated from this table are employee empowerment, availability of software packages, procurement method adopted and updating cost information during construction.

Therefore, some factors have to be eliminated. Some unimportant factors with less than 0.4 coefficient in the correlation matrix were eliminated. An eigenvalue of (1) was used for the principal components. The findings of the scree plot for the 48 factors which were later reduced to 44 factors for KMO adequacy provided four major categories from the elbow of the curve. This is illustrated in Figure 2.

In total, 11 eigenvalues were produced using the communalities. These values are above 1.0, and the scree plot in Table II indicates that the elbow breaks at the fourth point on the graph from the left. The component transformation matrix will be produced using the four identified factors.

| Kaiser-Meyer-Olkin measure of sampling adequacy | 0.639 |
| Bartlett’s test of sphericity | Approx. $\chi^2$ | 5,008.688 |
| | df | 946 |
| | Sig. | 0.000 |

Table I. KMO and Bartlett’s test after the second iteration

![Scree Plot](image)

Figure 2. Scree plot showing the eigenvalues
The component transformation matrix from the principal component extraction also made use of the Varimax with the Kaiser Normalisation rotation method. This method extracts the four major components. The four components identified from the scree plot extraction method are correlated with each other for further analysis to prove their correlation. The rotated matrix provides the coefficients of the factors, which are categorised into the four categories.

The extraction method was based on the principal component analysis. The rotation method employed Varimax with Kaiser Normalisation. The rotation converged in 20 iterations. The four factors extracted with an eigenvalue greater than 1 were used to re-categorise the drivers into four major components as listed in Table III.

The values of the factors which appeared in different components are selected based on the highest value. The drivers in Table IV have been ranked based on the outcome of the asymptotic significance in Sections 6.1 to 6.4.

The 44 drivers have been allocated to the CSF based on the pattern matrix. The names given to the categories are based on the similarities between the drivers. The most important drivers in each category are now tested for their level of importance using Kendall’s test for concordance. The four factors are management function, operational efficiency (OE), construction business management ethics and CCM as discussed in the sections below.

6.1 Assessing the drivers for construction management function (CMF) CSF
In assessing the drivers for management function, Kendall’s coefficient of concordance was used to address the asymptotic significance and for drivers within the CSF. Pallant (2016) noted that if the asymptotic significance is less than 0.050, then there is a high relationship between the drivers in relation to CMF. In Table IV, CMF1 representing variations and rework during construction has $p = 0.000 < 0.050$. This indicated that CMF1 is crucial in CMF and project success. Furthermore, CMF2 indicates excellent remuneration and motivation, and contractor/QS relationship are very important to implementing kaizen in the Nigerian construction industry. In Figure 3, CMF4 to CMF9 have $p < 0.050$. Therefore, having excellent relationships between the stakeholders involved in construction projects presents an opportunity for kaizen adoption. Kaizen depends on relationship management within the workplace for successful continuous improvement.

6.2 Assessing the drivers for OE CSF
The asymptotic significance for the drivers under OE has four drivers which are significant. OE1 has $p = 0.001 < 0.050$, hence, excellent working conditions in construction companies within Nigeria is extremely important for implementing kaizen. OE2, OE3 and OE4 in Table IV, have $p$-values, 0.006, 0.013 and 0.041 < 0.050 as illustrated in Figure 4. The deductions here imply that contractors’ decision, price and design risk and post-project reviews of cost information can propel the attainment of kaizen during construction processes.

In Figure 4, there are seven drivers which have $p$ values of > 0.050. Drivers OE5 to OE11 in Table IV, are not important in driving construction companies towards continuous improvement. Continuous improvement processes require excellent working conditions and

<table>
<thead>
<tr>
<th>Component-transformation matrix</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.612</td>
<td>0.523</td>
<td>0.443</td>
<td>0.393</td>
</tr>
<tr>
<td>2</td>
<td>−0.145</td>
<td>−0.659</td>
<td>0.480</td>
<td>0.561</td>
</tr>
<tr>
<td>3</td>
<td>−0.344</td>
<td>0.306</td>
<td>−0.522</td>
<td>0.718</td>
</tr>
<tr>
<td>4</td>
<td>−0.097</td>
<td>0.446</td>
<td>0.548</td>
<td>−0.126</td>
</tr>
</tbody>
</table>
it is based on post-project reviews of cost information and the overall processes involved in construction. Decision making by the contractor in construction companies has a major influence on the day-to-day running of the organisation.

6.3 Assessing the drivers for construction business ethics (CBE) CSF
Suppliers’ costs of building materials (CBE1) has a p value of 0.000 < 0.050. This is crucial for CBE because of associated kickbacks and corrupt practices between the subcontractor

<table>
<thead>
<tr>
<th>Drivers</th>
<th>Component 1</th>
<th>Component 2</th>
<th>Component 3</th>
<th>Component 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexible organisation policy</td>
<td>0.778</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Existing continuous improvement policy</td>
<td>0.681</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government regulations</td>
<td>0.631</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisation structure and communication</td>
<td>0.614</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architect/project managers’ decision</td>
<td>0.598</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project complexity</td>
<td>0.590</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contractor/QS relationship</td>
<td>0.585</td>
<td>0.503</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee experience</td>
<td>0.572</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variations and rework during construction</td>
<td>0.541</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent remuneration and motivation</td>
<td>0.540</td>
<td>0.409</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experienced QS and other staff</td>
<td>0.466</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent employee/employer relationship</td>
<td>0.459</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disputes and litigations</td>
<td>0.431</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarity of exclusions in the contract</td>
<td>0.426</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architect/QS relationship</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee empowerment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved teamwork</td>
<td>0.752</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial status of the company</td>
<td>0.730</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price and design risk</td>
<td>0.700</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent working conditions</td>
<td>0.667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Influence of construction professional bodies</td>
<td>0.605</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training of inexperienced employee</td>
<td>0.404</td>
<td>0.578</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-project reviews of cost information</td>
<td>0.578</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contractor decision making</td>
<td>0.562</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardised production process</td>
<td>0.483</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QS site experience</td>
<td>0.409</td>
<td>0.482</td>
<td>0.414</td>
<td></td>
</tr>
<tr>
<td>Regular site meetings</td>
<td>0.472</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stability of market conditions</td>
<td></td>
<td></td>
<td>0.727</td>
<td></td>
</tr>
<tr>
<td>Quality of cost information</td>
<td></td>
<td></td>
<td>0.704</td>
<td></td>
</tr>
<tr>
<td>Fraudulent practices and kickbacks</td>
<td></td>
<td></td>
<td>0.688</td>
<td>0.403</td>
</tr>
<tr>
<td>political stability</td>
<td></td>
<td></td>
<td>0.684</td>
<td></td>
</tr>
<tr>
<td>Subcontractors’ cost</td>
<td></td>
<td></td>
<td></td>
<td>0.672</td>
</tr>
<tr>
<td>Payment delays</td>
<td></td>
<td></td>
<td>0.643</td>
<td>0.446</td>
</tr>
<tr>
<td>Price fluctuations</td>
<td></td>
<td></td>
<td>0.639</td>
<td></td>
</tr>
<tr>
<td>Suppliers’ cost of materials</td>
<td></td>
<td></td>
<td>0.571</td>
<td></td>
</tr>
<tr>
<td>Improved contractor-client communication</td>
<td>0.473</td>
<td></td>
<td></td>
<td>0.538</td>
</tr>
<tr>
<td>Contractor/subcontractor relationship</td>
<td></td>
<td></td>
<td></td>
<td>0.727</td>
</tr>
<tr>
<td>Contractor/suppliers relationship</td>
<td></td>
<td></td>
<td></td>
<td>0.720</td>
</tr>
<tr>
<td>QS decisions</td>
<td></td>
<td></td>
<td></td>
<td>0.663</td>
</tr>
<tr>
<td>Communication among project professionals</td>
<td></td>
<td></td>
<td>0.432</td>
<td>0.662</td>
</tr>
<tr>
<td>Management of overheads on cost</td>
<td></td>
<td></td>
<td></td>
<td>0.636</td>
</tr>
<tr>
<td>Claims</td>
<td></td>
<td></td>
<td></td>
<td>0.618</td>
</tr>
<tr>
<td>In-depth knowledge of production process</td>
<td></td>
<td></td>
<td>0.509</td>
<td>0.564</td>
</tr>
<tr>
<td>Construction method</td>
<td></td>
<td></td>
<td>0.499</td>
<td>0.499</td>
</tr>
</tbody>
</table>

Table III. Rotated matrix table for the drivers
and suppliers. CBE2 and CBE3 in Figure 5 have p values of 0.002 and 0.029, respectively. The p values for CBE2 and CBE3 are therefore less than 0.050. In Table IV, political stability and the quality of cost information required for CBE in Nigeria are connected to corrupt practices in the Nigerian political structure and construction industry. This connection does not imply a positive outlook for implementation, but non-causal drivers associated with existing realities of construction business in Nigeria. Therefore, there has to be a favourable political system, access to excellent cost data for kaizen to be implemented and the cost information provided by the QS will aid the reduction of corrupt practices in the construction industry.

The extent to which the three most significant drivers can influence CBE in the Nigerian construction industry presents an opportunity for the adoption of kaizen.

6.4 Assessing the drivers for CCM

In Figure 6, the stability of market conditions (CCM1) has the highest p-value of 0.006. From Table IV, CCM2 to CCM4 have p-values of 0.009 to 0.043 which are below 0.050 grid in Figure 6; all indicate significance to the CCM CSF.

<table>
<thead>
<tr>
<th>Construction management function (CMF)</th>
<th>Operational efficiency (OE)</th>
<th>Construction business ethics (CBE)</th>
<th>Construction cost management (CCM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variations and rework during construction (CMF1)</td>
<td>Excellent working conditions (OE1)</td>
<td>Suppliers’ cost of materials (CBE1)</td>
<td>Stability of market conditions (CCM1)</td>
</tr>
<tr>
<td>Excellent remuneration and motivation (CMF2)</td>
<td>Contractor’s decision making (OE2)</td>
<td>Political stability (CBE2)</td>
<td>Updating cost information during construction (CCM2)</td>
</tr>
<tr>
<td>Contractor/QS relationship (CMF3)</td>
<td>Price and design risk (OE3)</td>
<td>Quality of cost information (CBE3)</td>
<td>QS decisions (CCM3)</td>
</tr>
<tr>
<td>Existing continuous improvement policy (CMF4)</td>
<td>Post-project reviews of cost information (OE4)</td>
<td>Improved contractor-client communication (CBE4)</td>
<td>Claims (CCM4)</td>
</tr>
<tr>
<td>Architect/project managers’ decision (CMF5)</td>
<td>Standardised production process (OE5)</td>
<td>Price fluctuations (CBE5)</td>
<td>Architect/QS relationship (CCM5)</td>
</tr>
<tr>
<td>Clarity of exclusions in the contract (CMF6)</td>
<td>Financial status of the company (OE6)</td>
<td>Fraudulent practices and kickbacks (CBE6)</td>
<td>In-depth knowledge of production process (CCM6)</td>
</tr>
<tr>
<td>Government regulations (CMF7)</td>
<td>Influence of construction professional bodies (OE7)</td>
<td>Payment delays (CBE7)</td>
<td>Accuracy of estimates (CCM7)</td>
</tr>
<tr>
<td>Excellent employee/employer relationship (CMF8)</td>
<td>Regular site meetings (OE8)</td>
<td>Subcontractors’ cost (CBE8)</td>
<td>Construction method (CCM8)</td>
</tr>
<tr>
<td>Flexible organisational policy (CMF9)</td>
<td>Training of inexperienced employee (OE9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disputes and litigations (CMF10)</td>
<td>QS site experience (OE10)</td>
<td>Contractor/suppliers relationship (CCM9)</td>
<td></td>
</tr>
<tr>
<td>Employee experience (CMF11)</td>
<td>Improved teamwork (OE11)</td>
<td>Management of overheads (CCM10)</td>
<td>Communication among project professionals (CCM11)</td>
</tr>
<tr>
<td>Project complexity (CMF12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experienced QS and other staff (CMF13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisational structure and communication (CMF14)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table IV. Categorisation of drivers from Table III into CSFs

Note: Drivers in italics are considered important to the CSF based on rotated matrix data.
Asymptotic significance for construction management function drivers

Figure 3.
Asymptotic significance for operational efficiency drivers

Figure 4.
Nigerian construction industry
Therefore, updating cost information and the QS decision before and during construction as a continuous improvement mechanism will make it easier for kaizen to fit into the project life cycle. Reduction of claims during construction projects is significant in the CCM driver for kaizen. Claims will be reduced when there is meticulous cost information.

7. Discussion of findings: Implementation of kaizen in the Nigerian construction industry

This section deduces the exploratory factor analysis and asymptotic significance findings for the adoption of kaizen in the Nigerian construction industry. The discussion below is also expanded to other developing countries. A general perspective of the findings in this discussion provides more insight into how the CSFs in Table IV are justified.

7.1 Construction management function

The management of construction activities starts with the decision made by the contracting management, in this instance, the contractor or group of contractors. The contractor will make certain decisions based on the type of contract, duration and conditions of contracts. Yngve (1995) noted that an organisation could not exist alone, it has to interact with other systems, other organisations. Certain drivers within and outside influence construction companies. Some of these are flexible organisational policy; existing continuous improvement policy; government regulations; organisational structure and communication; architect/project managers’ decision; project complexity; contractor/QS relationship; employee experience; variations and rework during construction; excellent remuneration and motivation; excellent employee/employer relationship; experienced QS and other staff; disputes and litigations and clarity of exclusions in the contract (please refer to Table IV).
Variation management in construction is an integral part of construction management. Variation management has a $p$-value of 0.000 $< 0.050$, indicating very high significance. Change management has been a major discussion point in construction management circles (Alaryan et al., 2014; Othman, 1997). In achieving kaizen in the construction process, variations and rework as part of the change management process should be minimised. This process can also be achieved through early adoption of kaizen in construction cost planning. Involvement of the design team in construction cost planning has an effect on the rapid attainment of cost reduction during construction. Hence, variations and rework is a significant driver for the CMF within the sphere of kaizen realisation in the construction milieu.

Excellent remuneration and motivation, contractor/QS relationship and existing continuous improvement policy are all related to what is happening within the organisation. The aforementioned drivers had $p$-values of 0.00, 0.003, 0.004 and 0.019, respectively. The interaction of the organisation with the external environment has to do with government regulations such as taxes, and financial incentives for construction companies in Nigeria. The major drivers for a successful business are the human resources, motivation and relationship management within the organisation. These factors are crucial for the implementation of kaizen within the organisation. The organisational strategy for kaizen implementation has to do with making problems visible, setting high targets and staying in touch with reality (Emiliani, 2005; Singh and Singh, 2015; Smadi, 2009; Suárez-Barraza et al., 2011). The owners of construction companies need the input of employees for better identification of problems. Therefore, adequate communication is required for kaizen in the workplace. Imai (2012) stated that kaizen in the workplace is known as gemba kaizen, this is important for good housekeeping and management for continuous improvement. Continuous improvement in the place of work can only exist when there is a form of motivation. The motivation of workers to keep track of what is wrong or requires improvement is crucial. Implementing kaizen, in this instance gemba kaizen, depends on the experience of the employees. It also depends on the relationship between the employer and the employee.
The findings in this study reveal that a proper consideration of government regulations before implementing kaizen is imperative. In Nigeria, there are no government regulations on CCM. Moreover, most construction companies have the obligation to advance the course of their project with any modern construction methods and standards available in the construction industry.

Findings from the empirical data suggest that there is a major impact of the abovementioned drivers which are significant to the success of the management function as key attributes for the application of kaizen in Nigerian construction companies. CMF as a finding under the CSF connotes a wider scope of how construction organisations can redress their performance in line with continuous improvement principles.

7.2 Operational efficiency
The modus operandi of most construction companies in the small and medium scale range depends on some drivers. These drivers are excellent working conditions; decisions made by the contractor; the experience of the Quantity Surveyor on site; training of inexperienced employees and other staff; regular site meetings; influence of construction professional bodies; improved teamwork; financial status of the company; price and design risk; standardized production process; and post project reviews of cost information. These drivers were also assessed in terms of their relevance to the implementation of kaizen during construction.

The findings indicate that excellent working conditions have a significant $p$-value of 0.001. The excellent working conditions entails health and safety standards, security, modern facilities and spaces in the workplace and construction sites. Crema et al. (2015) provided ways of improving the safety with lean and safety methods. Shang and Pheng (2013) also suggested the standardisation of the construction process for easier implementation of kaizen in the construction industry. The process of standardisation in construction does not have to do with the closed working environment as in a factory but a well-structured method statement, a work breakdown structure and information management system.

Contractors’ decision-making skills having a $p$-value of 0.006 is also a vital driver for OE. The decision made on construction sites by the contractor can be a positive or negative influence on the outcome of the construction project. The contractor makes very spirited decisions during variation management, claims and the cashflow. These decisions go a long way in influencing relationships between stakeholders and the final outcome of the construction project. Although the QS and architect also make these decisions, the final approval comes from the contractors.

The QS and Architect with a significant $p$-value of 0.013, on site also influences the decisions the contractor makes; this is based on some level of experience. The site quantity surveyor’s experience may not be of relevance to the OE of the site because the architect makes most of the decisions along with the contractor. The price and design risk is very significant to this study because the effect of fluctuations and variation are always influential. This affects the activities, which are related to incremental cost reduction. Design risks can lead to variations, which makes contractors claim for more work on the site.

The findings indicated a post-project review driver under OE with a $p$-value of 0.041. Post-project reviews of construction cost information is an indication of the extent to which construction companies have been able to add some element of continuous improvement. Continuous improvement in construction can be established further with the aid of kaizen principles, such as plan-do-check-act. This process is circular and adds an incremental understanding to the processes which have to be improved. A transference of errors in the construction process should be avoided at all cost when using historical cost information for
elemental cost planning. The elemental cost planning phase of construction projects may systematically include latent errors from previous projects when they have been used for multi-rate or single rate estimating. Furthermore, post-project reviews can be applied to other planning phases of construction projects, such as design, stakeholder relationship management, schedule management and method statement.

Overall, the findings show that decisions made by the contractor, architect and QS; reviewing previous project details can be strengthened in the construction process to other to establish kaizen within Nigerian construction companies.

### 7.3 Construction business ethics
This aspect of construction business has to do with communication, information management, relationship management with the client and ethical business policies during construction. This factor is related to the construction process phase. The drivers for this factor are improved contractor-client communication; subcontractors cost; political stability; payment delays; suppliers’ cost of materials; fraudulent practices, kickbacks and price fluctuations.

The findings indicated a significance of 0.000 for suppliers’ cost of building materials. Political stability with a $p$-value of 0.002 and quality of cost information as 0.029. The empirical findings are aligned with the present unethical condition of construction practises in Nigeria.

The practices involved in the contractor–client relationship has to do with how the clients, who can be government, quasi-government private clients, and organisations or companies, make arrangements for procurement with the contractors to deliver a project. In Nigeria, certain classes of contractors are forbidden to take part in the submission of tender. Small and medium scale construction companies in Nigeria have a lot of challenges with competitiveness in the industry (Ayanda and Laraba, 2011; Chukwudi and Tobechukwu, 2014). This fact was corroborated with the result of the asymptotic significance. Suppliers’ costs of materials is the most significant driver within the CBE. In most developing countries around the world, building material cost has a major significant effect on the final cost of construction projects (Omotayo and Kulatunga, 2017). Bribery and corruption in construction business is not peculiar to developing economies. In the UK, several cases of bribery have resulted in persecution of contractors (Transparency International UK, 2011). Inflation of the cost of building materials has a direct relationship with final construction cost. Therefore, suppliers’ building material cost has to be checked for unethical practices as part of kaizen implementation.

Political stability; this driver is more significant than subcontractors’ cost. Political stability has a way of influencing construction business in Nigeria. Oyedele (2015) noted that political stability and security in the northern part of Nigeria would make construction business difficult. In the Southern parts of Nigeria, thugs are bribed in order to avoid unnecessary violence. These types of conditions affect the implementation of kaizen. The political situation in most developing countries are volatile, hence, a more proactive approach to managing construction business is imperative. Kaizen possesses this problem of detection approach. Furthermore, the political system impacts on construction cost, due to government corrupt practices, deliberate padding of construction budget and institutionalised corruption.

The quality of cost information for the post-contract cost control process starts from the cost-planning phase and the working budget. Cost estimates are always very inaccurate (Park and Papadopoulou, 2012). Oyedele (2015) has also discussed the problems with cost estimates in Nigeria. The factors Oyedele (2015) identified are crucial and mostly unavoidable. Therefore, the quality of cost information seems more important compared to improving the contractor-client relationship. In recent time, most Nigerian auditing companies have adopted
The Knowledge Academy (2018) as a quality management system for construction process improvement (The Knowledge Academy, 2018). The Knowledge Academy (2018), can be used in enhancing the quality of construction processes and products in Nigeria by associating the plan-do-check-act principle with quality control of cost data.

The findings indicate that in adopting kaizen, the construction business and environment have to be standardised. Kaizen can be used to maintain the quality of cost information and delivery time (Brunet and New, 2003; Colenso, 2000; Doolen et al., 2008). The challenges of business ethics with kaizen has not been considered fully, but unethical practices on construction sites have a way of influencing the delivery of the project and overall cost.

7.4 Construction cost management
The drivers for the factors highlighted were claims; updating cost information during construction; stability of market conditions; the accuracy of estimates; in-depth knowledge of production process; architect-quantity surveyors’ relationships; quantity surveyors’ decision; construction method; communication among project professionals; contractor/suppliers relationship; and management of overheads on cost. The asymptotic significant findings reveal that stability of market conditions, updating cost information, QS decisions and claims have p values less than 0.050 at 0.006, 0.009, 0.019 and 0.043, respectively as indicated in Figure 6.

The stability of market conditions can determine the outcome of final construction cost. This was investigated by Adegoke (2016), Odediran et al. (2013) and Olatunji (2008), the stability of market conditions affects the prices of building materials, particularly cement. Prices fluctuations, foreign exchange rate, bank lending rates and the state of the economy all influences construction cost. Olatunji (2008) noted that the tender sums of most construction projects in Nigeria are higher during unstable market conditions. The mean score for unstable market conditions is 23.85, while the significant value is 0.006. This proves that this driver is very significant for the factor.

Excessive claims have a way of leading disputes in construction (Lord and Gray, 2011; Olanrewaju and Anavhe, 2014). The process of updating cost information during construction in the cashflow is vital for cost monitoring and control. The cost information can arise from materials purchase, overheads relating to site office, maintenance of site utilities and payment of subcontractors, suppliers and labourers. The process of updating cost information is a corrective process, and it is carried out during variation management.

Decisions made by the QS may be opposed by the architect or contractor. Decisions to adopt a new building element based on design changes have cost, time and quality implications. The contractor’s decision may contradict what the QS suggests. Most conflicts in construction arise as a result of financial differences between the budget and actual expenses. Variation management and rework may result in a number of disputes.

It can be deduced from the findings that additional details included during construction may jeopardise planned schedule and time. Thus, CCM as a CSF have to adopt kaizen for continuous cost improvement purposes. The aforementioned underlying drivers of CCM will have a major impact on the final account of a construction project.

8. Conclusion
Several organisations may have been making use of kaizen in the form of continuous improvement, however, it is still very new to the construction industry. With the wider scope and complexity of construction projects around the world, a more holistic philosophy into managing construction projects for enhanced productivity is sacrosanct. In developing economies, the gross domestic product growth rate is high, but infrastructure delivery has remained the same. Hence, an economy depends on the construction industry for growth.
Most construction companies in developing economies have productivity and other associated issues which limit their growth. Continuous improvement has been proven to provide major impact on the growth of construction companies in developing countries. Implementing continuous improvement demands a lot of financial resources, cost and time. Nonetheless, certain CSFs need to be in place for construction organizations to adopt kaizen. CMF, OE, CBE and CCM covers the intricate drivers for a successful kaizen implementation. Within the Nigerian scope as an extrapolation for developing economies, the aforementioned CSFs demand attention for successful project delivery and kaizen. CMF explored the drivers which include financial motivation, relationship between the upper and lower management, change management in terms of handling variations and rework and the relationship between the organisation and external systems, in this instance government regulations. An exploratory view of these drivers indicate that construction companies in developing economies need a lot of good housekeeping for effective realisation of kaizen.

OE in construction, business ethics and cost of construction has been emphasised in previous sections to indicate drivers that should be delimited and may establish a decline in the CSFs. Kaizen is concerned with waste minimisation and elimination for effective profitability. Therefore, drivers which propagate intense waste during construction should be of prime significance to the project life cycle and kaizen. Certain attitudinal challenges were highlighted in form of drivers. These are political stability, contract and QS decisions and suppliers’ material cost. These attitudes expressed by key stakeholders in the construction sector cannot be eliminated easily. Change management is essential in adopting kaizen.

Change management of organisational culture and perception towards innovation and new ideas is therefore needed. Kaizen is construction is definitely possible because it has been adopted in other developing countries. The construction economies which are yet to delve into this paradigm of continuous improvement require a change management strategy based on the identified CSFs. The possibility of adopting kaizen in Nigeria and other developing countries is high. However, careful consideration of implementation cost and duration is essential.

References


Further reading


About the authors
Temitope Seun Omotayo holds a PhD in Construction and Project Management from the University of Salford and a Master’s Degree in Quantity Surveying (Mechanical and Electrical) from the same institution. His undergraduate degree in quantity surveying was from Federal University of Technology Akure, Nigeria. With a postgraduate certificate in Teaching, Temi will be a fellow of the higher education academy. He also coordinates distance learning modules, supervises undergraduate and postgraduate dissertations. At the moment he is currently co-supervising a PhD student along with Dr David Moore. Temitope Seun Omotayo is the corresponding author and can be contacted at: t.s.omotayo@rgu.ac.uk

Dr Udayangani Kulatunga is Reader at the School of the Built Environment and also Programme Director for BSc Quantity Surveying offered in Sri Lanka. She is also Director of the flagship research group, the Centre for Disaster Resilience, which provides strategic advice and practical guidance based on rigorous research, informed by industry and community members on man-made and natural disasters. Dr Kulatunga has more than 13 years’ experience in the academia. Dr Kulatunga has published widely in refereed journals and international conferences. Her research interests include cost management, procurement and administration, disaster management and performance measurement.

Bassam Bjeirmi has an extensive experience in the development and delivery of training programs for the private and public sectors. His main subject area is construction and generic project management where he has delivered training programs in the UK and overseas for companies including Shell, Sonatrach and ADNOC in Nigeria, Algeria and the Arabian Gulf, respectively. Over the years he has provided consultancy to a number of public and private organisations on project implementation, monitoring and control. Other areas included the capture of knowledge, its dissemination and its effect on organisational learning.
Aligning product variety with supply chain and business strategy

Juneho Um
Business School, Liverpool Hope University, Liverpool, UK

Neungho Han
Institute of Foreign Trade Research Units, Sungkyunkwan University, Jongno-gu, Republic of Korea, and

Tonci Grubic and Asad Ghalib
Business School, Liverpool Hope University, Liverpool, UK

Abstract
Purpose – The purpose of this paper is to demonstrate the strategic alignment between product variety and supply chain (SC) focus through cost leadership or differentiation to improve business performance.
Design/methodology/approach – The research investigated product variety-related capabilities and strategies in SCs including level of variety, SC agility, SC cost efficiency, cost leadership, differentiation and business performance, and aimed to justify the theory by testing structural equation modelling using survey data from the UK and South Korea.
Findings – Differentiation links high product variety and agile SC while cost leadership is aligned with low product variety and SC efficiency. High product variety negatively impacts on cost leadership strategy. Also, product variety should be mediated by cost leadership or differentiation strategy to improve business performance. Companies in South Korea display higher SC agility, cost leadership and cost efficiency than companies in the UK, while the UK companies exhibit a higher level of product variety and differentiation than those in South Korea.
Research limitations/implications – The findings contribute to theoretical development of variety issues at the alignment of business strategy and SC management according to the level of product variety.
Originality/value – The findings can help international companies set up specific variety-related strategies to achieve global competitiveness.

Keywords Supply chain, Business performance, Structural equation modelling, Product variety

Paper type Research paper

1. Introduction
Globalisation promotes the need to customise products and services according to the target market or country, and provision of product scope is an issue in deciding if globalisation raises marginal variety’s profit (Qiu and Zhou, 2013). Deregulation of trade with supporting technologies has promoted global markets, so manufacturers are required to appropriate strategies in order to provide product variety to different markets. In addition, a noticeable movement for manufacturers is to extend the variety of their products to provide more options for customers in today’s challenging global markets. Consumers can match their needs and preferences with the product variety provided by manufacturers (Lancaster, 1990). For example, mega-stores such as Tesco and Wal-Mart provide wide product variety in their quest to provide value to consumers, enabling the stores to improve market share and revenue. Some researchers stress the advantages of product variety based on long-tail theory (Anderson, 2006; Zhou and Duan, 2012), that product variety can improve sales, while others criticise high levels of variety provision mainly due to cost increase. Also, an excess of product information and provision can result in selection confusion for customers and lead to forecasting difficulty for manufacturers (Wan et al., 2012). For example, the market share of Aldi keeps growing, focusing on cost leadership with low product variety,
while Tesco have lost their market share to an increasing number of low-cost providers. From the operations perspective, an increase in product variety incurs a cost burden due to diseconomies of scale in the supply chain (SC). The emergence of a global SC has detonated the trade-off between provision of product variety and SC performance. Increases in product variety in the global SC incur both production costs and market mediation costs (Randall and Ulrich, 2001). Also, global purchasing and logistic costs increase with product variety provision. High-variety ambitions need to be profitably achieved without severe deterioration of SC performance, especially considering the cost profile.

Many manufacturers have recognised that a trade-off exists between product variety and SC performance (Fisher and Ittner, 1999; Thonemann and Bradley, 2002; Um et al., 2017). Thus, several researchers (Yeh and Chu, 1991; Blecker and Abdelkafi, 2006; Scavarda et al., 2010; Patel and Jayaram 2013) have investigated how to mitigate the impact of product variety on the cost profile of SC processes. Modularity, cellular manufacturing and postponement, typical of functional-level management such as SC integration (Scavarda et al., 2010; Um et al., 2017), SC flexibility and agility are essential in responding quickly to diverse customer needs and improving cost efficiency (Fisher et al., 1999; Swafford et al., 2008). SC flexibility and agility-based management activities in particular have been emphasised to satisfy their high-variety provisions and ambition (Um et al., 2017). In addition, partnerships with suppliers (Tummala et al., 2006; Cousins et al., 2011) and close customer relationships (Child et al., 1991; Wang and Feng, 2012) can be requisite criteria when the manufacturer considers a high level of product variety or the introduction of new products. Thus, business performance such as sales growth and SC performance including cost efficiency and customer service are influenced by the level of product variety provision and focused strategies such as cost leadership or differentiation by individual manufacturers. Theoretically, a high level of product variety environment employs a differentiation strategy to enhance customer service, while a low level of product variety environment considers the cost leadership policy resulting in cost efficiency (see Agarwal et al., 2006; Stavrulaki and Davis, 2010). However, most of the empirical studies on the relationship between product variety and business performance have dealt with potential mitigation strategies or functional approaches without explicitly considering the strategic alignment of SC and business strategies involved with product variety issues.

The level of the existing product variety reflects the organisation’s strategic goals and movement. Therefore, all variety-related management capabilities including customisation (i.e. from make to stock to design to order), cost leadership and differentiation, and performance including SC cost efficiency, agility and business-related performance, can differ according to the different level of product variety provided by manufacturers. Also, variation is extended to the national level based on the theoretical approaches in managing variety concerns. The fundamental question relates the matching level of variety offered with organisational or national focus and its reflective strategies. For example, based on the product types (Fisher, 1997) or level of customisation and product variety (Agarwal et al., 2006; Stavrulaki and Davis, 2010), manufacturers are required to focus on either lean, agile or leagile SCs (Lee, 2002).

The main objective of this study is to explore the appropriate alignment of product variety with business and SC strategy. Then, the research tests findings by comparing the differences between UK and Korean manufacturers in the level of product variety, management strategies and performance. The study also explores the structural process to achieve better business performance. This is achieved by evaluating differences in the level of variety and variety-related strategies including cost leadership and differentiation, and performance including SC agility and cost efficiency. Inspired by literature reviews on product variety and SC performance-related research (see Randall and Ulrich, 2001; Stavrulaki and Davis, 2010), the study also examines how differences in product variety
between the UK and Korea influence SC strategies and performance from global perspectives. The UK and Korea have different economic backgrounds and outputs, such as economic growth rates (i.e. 1.8 vs 2.0 per cent), GDPs (i.e. $2,888 vs $1,147bn), trade volume (i.e. $674 vs $527bn in exports), inflation rates (i.e. 1.8 vs 1.1 per cent) and income distribution (i.e. 0.32 vs 0.30 in Gini coefficient) (International Monetary Fund, 2016; Central Intelligence Agency, 2016). Therefore, the data were collected from two countries that have different economic backgrounds for a cross-examination to generalise the research outcomes better in the global business environment. The key aims for this study are to identify: appropriate alignment of business strategies with product variety in the SC; the relationships between variety-related factors responsible for improving business performance; and the differences in product variety, and SC performance, that exist between the UK and Korea, suggesting the appropriate SC implications for companies in each country.

The research makes two significant contributions. First, it establishes the relative extended theory of product variety management at corporation and national level. A corollary of this contribution concerns the subsequent implications for company strategies and policy. Second, the research findings from manufacturers explain how different levels of product variety can impact on business performance positively within the SC. The findings also suggest important managerial and practical alignment for the adoption of different approaches to product variety under different strategic backgrounds, from the perspective of manufacturers.

2. Research background and development of hypotheses
2.1 Product variety and SC performance
The term “product variety” is employed by academics and industry with a number of different conceptual meanings (Stablein et al., 2011). Thus, the concept and scope of product variety must be defined clearly. For example, variety can be defined as a number or collection of different things of a particular class of the same general kind, and product line breadth refers to the stock-keeping units (SKUs) within a brand category (ElMaraghy et al., 2013). Therefore, increase in product variety refers to the introduction of new products and can be calculated as the number of SKUs in the SC (Wan et al., 2012). MacDuffie et al. (1996) defined it to refer to company choices about the breadth and depth of different product lines using three dimensions for types of variety: first, fundamental (i.e. different core models and designs at the fabrication and design stage); second, intermediate (i.e. different technical options dependent on core design at the assembly stage); and third, peripheral (i.e. particular options and accessories independent of core design at the distribution and sales stage). Such an explanation is a reference to internal variety that is generally regarded as the variance involved in creating the product within a firm or SC. On the other hand, external variety is the extent of different and distinguishable products offered by manufacturers in the marketplace that is available to the end consumers. This study employed internal variety based on the concept of MacDuffie et al. (1996) from the perspective of manufacturers, considering potential varieties the manufacturers are able to provide.

There is a robust relationship between product variety and the SC cost escalation (Scavarda et al., 2010). A more cost-efficient provision of product variety can generally be attained through three operations strategies: first, use of the modularity concept by changes in product architecture, which can result in minimising the complexity and associated cost in SC functions (see Forza and Salvador, 2002; Holweg and Pil, 2004; Caridi et al., 2012); second, flexibility in the operations process using group technology (see Child et al., 1991; Berry and Cooper, 1999); and third, postponement strategy which delays product configuration decisions in the SC process, such as form and time postponement (Holweg and Helo, 2014). However, since functional-level strategies suggested do not cover the entire SC,
the level of product variety should be matched with organisational SC strategies in advance. For example, stock-outs that result from a high product variety strategy may ultimately hurt sales performance (Wan et al., 2012).

Beamon (1999) developed an SC performance framework through a mixture of measures, and suggested three systemic types of performance as critical components: resource, output and flexibility. The resource measure refers to a high level of SC efficiency (e.g. the total costs of resources, inventory, manufacturing, distribution in the SC) while the output measure refers to a high level of customer service (e.g. customer satisfaction, customer response times, on-time deliveries, order fill rate, customer complaints, backorder/stock-out, manufacturing lead time and shipping errors). Many organisations are endeavouring to capture the benefits of flexibility in SCs (Gopal and Thakkar, 2012). SC flexibility is the internal capability for responding to the changing environment at the functional level. Its attributes are three fundamental processes: procurement, manufacturing and distribution flexibility (Swafford et al., 2006). However, agility represents an externally focused competence concentrating more on speed at the business level, such as market responsiveness, delivery reliability, lead time, product customisation and frequency of product introduction (Swafford et al., 2008). Therefore, SC flexibility such as change delivery time, production volume, production mix, delivery schedule and implementation of emerging change can be regarded as an important antecedent of SC agility. Agility itself is a dynamic capability concerned with speed in unpredictable markets with significant demand turbulence to achieve better customer service. Therefore, the role of SC agility should be stressed in the high product variety environment. This study adapts the concept of SC cost efficiency from Beamon (1999) and SC agility from Swafford et al. (2008), respectively.

2.2 Research model and hypotheses
Theorising seeks to produce findings grounded in evidence and geared towards understanding how relations lead to results, and it extends knowledge (Theodore et al., 2017). One obvious piece of evidence is that product variety and provision of customer options have developed for companies to increase sales volume and market share; however, companies need to accept cost increase in SC and employ appropriate SC strategies to mitigate the impact of product variety. In short, based on the level of product variety, SC strategy and competitive criteria can differ between manufacturers, which can impact on business performance differently. For example, different levels of product variety compared with competitors can result in different strategies, such as different degrees of focus on either cost leadership, differentiation or hybrid strategies. Product customisation is a factor which increases product variety (Silveria, 1998) that can have a negative impact on cost leadership (Um et al., 2017). A high product variety environment typically has unpredictable demand, short life cycle, close customer relationship and high margin (Fisher, 1997; Stavrulaki and Davis, 2010). Thus, focus should be on SC agility, differentiation and customer service (Um, 2017). Instead, a low product variety environment has more predictable demand, close supplier relationship and mass production (e.g. economy of scale) focusing on SC cost efficiency and cost leadership.

Cost leadership is related to cost reduction as a market winner, and differentiation is related to new product development or variety increase with high quality (Kim, 2006). Hallgren and Olhager (2009) insisted that the three strategies of Porter (2004) can fundamentally be reduced to two, since the company must choose between cost leadership and differentiation strategies even with a focused strategy. In this study, cost leadership has two elements: low price and low manufacturing unit cost. Differentiation has three components that are related to customer service (i.e. product), technology and marketing differentiation, respectively. For example, deliver a high-quality product with volume flexibility, develop a new product quickly with design flexibility, and control the sales and distribution network with a distinctive brand image (see Kim, 2006; Hallgren and
Capabilities such as agility and cost efficiency contribute positively to business performance, either acting alone or in concert with other capabilities (Rosenzweig et al., 2003). The variety-related characteristics in SC are summarised in Table I.

H1a–H1b were designed to investigate strategic differences and performance according to the level of product variety. In terms of alignment of product variety with the strategies, the research considered SC agility, SC cost efficiency, differentiation, cost leadership and business performance. The performance of an SC can be attributed to a match or mismatch between the type of product (i.e. innovative or functional) and the design of the SC (Fisher, 1997).

Stavrulaki and Davis (2010) also highlighted alignment between the key aspects of a product including product variety, and its SC processes and strategy. From the perspective of variety, a high level of product variety corresponding with a high level of customisation typically focuses on differentiation, variety management strategies and customer relationships that enhance SC agility, while a low level of product variety corresponding with a low level of customisation generally focuses on cost leadership and cost efficiency. Also, based on the long-tail effect, the change in the consumption pattern when more niche products are being provided enables demand to shift from the hits to the niches over time, which can improve customer service and satisfaction (Anderson, 2006; Zhou and Duan, 2012). Thus, the study proposes the strategic alignment matching strategies with product variety and the SC, shown in Figure 1 (the strategic alignment model) and the following hypotheses:

**H1a.** High product variety is associated with increased focus on differentiation.

**H1b.** SC agility is associated with increased focus on differentiation.

<table>
<thead>
<tr>
<th>Product variety</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand uncertainty</td>
<td>Predictable</td>
<td>Unpredictable</td>
</tr>
<tr>
<td>Production focus</td>
<td>Economy of scale</td>
<td>Diseconomy of scale</td>
</tr>
<tr>
<td>Level of customisation</td>
<td>Low customisation</td>
<td>High customisation</td>
</tr>
<tr>
<td>Production structure</td>
<td>Make to stock</td>
<td>Make/assembly to order</td>
</tr>
<tr>
<td>Product life cycle</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>Product type</td>
<td>Functional</td>
<td>Innovative</td>
</tr>
<tr>
<td>Relationship focus</td>
<td>Suppliers</td>
<td>Customers</td>
</tr>
<tr>
<td>Supply chain strategy</td>
<td>Lean SC</td>
<td>Agile SC</td>
</tr>
<tr>
<td>Market strategy</td>
<td>Cost leadership</td>
<td>Differentiation</td>
</tr>
<tr>
<td>Supply chain focus</td>
<td>Cost</td>
<td>Customer service</td>
</tr>
<tr>
<td>Profit margin</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

**Sources:** Adapted from Fisher (1997) and Stavrulaki and Davis (2010)
Cost efficiency is dominant for mass production systems with a low level of customisation and variety, while more enhanced customer service is required for companies where product variety matters (see Agarwal et al., 2006). Therefore, business strategies should be considered in line with the level of product variety and SC strategies. According to Silveria (1998), the most significant factors motivating an increase in product variety are the ability to customise the product and the demands made by customers. In short, product proliferation and variety increase together (Hu et al., 2011). Under this circumstance, an agile system is appropriate for differentiation that focuses on product variety and customisation and is negatively associated with a cost-leadership strategy that focuses on cost efficiency (Hallgren and Olhager, 2009). Thus, the study proposes the following two hypotheses:

\[H2a.\] Low product variety is associated with increased focus on cost leadership.

\[H2b.\] Cost efficiency is associated with increased focus on cost leadership.

Business performance is measured through accounting data that represents the company's performance and market valuation (Vickery et al., 2003) such as return on assets (ROA), return on investment (ROI) and return on sales (ROS). Tan et al. (1999) recommended nine diverse measures of business performance that involve market share, ROA and overall competitive position (market share growth, sales growth, ROA growth, production cost, customer service levels, product quality and competitive position). Vickery et al. (1999) supported the use of ROI, ROS, market share and the performance growth of each (e.g. market share growth) for business performance. Panayides (2007) argued that multiple measures of performance would reflect a firm's improvements more accurately. Thus, this study employed ROI, ROS, market share growth and sales growth to indicate business performance.

Product variety incurs cost increase in SC mainly due to the increased complexity in manufacturing, purchasing and logistics functions in SC. Also, trade-off exists between product variety and SC performance while a relationship between product variety and business performance is still arguable. Kekre and Srinivasan (1990) investigated the market benefits and cost disadvantages of product variety and concluded that product variety increases market share and a firm’s profitability. To improve profitability, firms should make competitive moves, and new product introductions (NPI) are one of these moves that might positively influence market share and returns (Otero-Neira et al., 2010). However, in the case of mature firms, increased variety may not increase total demand that improves profitability. Instead, firms can increase variety to retain market share by a differentiation approach. Besides cost increases, an extensive array of options can at first seem highly appealing to consumers, yet can reduce their motivation to purchase the product. To address this variety issue, appropriate strategic approaches are crucial to achieve better business performance. Thus, the study proposes the following hypotheses based on the structural equation model (Figure 2):

\[H3.\] Differentiation and cost efficiency are associated with business performance.

\[H4.\] Differentiation and cost leadership mediate the relationship between product variety and better business performance.

Different location and country characteristics affect the overall performance of a firm differently, and the level of internationalisation of each country may impact on economic performance (Antonio et al., 2015). As well as to confirm the findings, the study applies a comparative analysis to the cases of the UK and Korea, considering the different strategic focus of companies in each country and business performance measured according to the level of product variety.
3. Research methodology

A questionnaire considering the product variety issues was sent to companies identified as manufacturers, based on their standard industrial classification (SIC) code. The data were collected from two countries with the intention of conducting a cross-examination and comparison. The survey was conducted by post with a package including a covering letter and return stamped envelope for the UK companies. In Korea, in order to obtain an acceptable level of response, e-mail and direct interview surveys were chosen. In total, 212 companies responded to the survey from the UK and 152 from Korea (total = 364), yielding an acceptable 18 per cent overall response rate (see Frohlich, 2002; Anseel et al., 2010). As a whole, 84.1 per cent of the participants had positions above assistant manager and sales representative. In total, 59.1 per cent of the firms were small or medium sized and the remaining 40.9 per cent were large firms. Table II presents the industry types of the survey respondents. The demographic analysis of the responses indicated that the participating firms spanned a diverse group of manufacturing industries, which allowed for generalisation of the findings (Gatignon and Xuereb, 1997). Product variety was measured as fundamental, intermediate and peripheral using a five-point scale (1 = 1–5, 2 = 6–10, 3 = 11–15, 4 = 16–20, 5 = above 20) based on the core product family (MacDuffie et al., 1996). Also, strategies including differentiation, cost leadership, cost efficiency, SC agility

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>UK Frequency</th>
<th>Percentage</th>
<th>Korea Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing industry type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food, beverage, tobacco</td>
<td>17</td>
<td>8.0</td>
<td>9</td>
<td>5.9</td>
</tr>
<tr>
<td>Wood and furniture</td>
<td>21</td>
<td>9.9</td>
<td>11</td>
<td>7.2</td>
</tr>
<tr>
<td>Chemical materials and products</td>
<td>15</td>
<td>7.1</td>
<td>13</td>
<td>8.6</td>
</tr>
<tr>
<td>Non-metal mineral products</td>
<td>10</td>
<td>4.7</td>
<td>5</td>
<td>3.3</td>
</tr>
<tr>
<td>Fabricated metal products</td>
<td>29</td>
<td>13.7</td>
<td>4</td>
<td>2.6</td>
</tr>
<tr>
<td>Computer and communication products</td>
<td>9</td>
<td>4.2</td>
<td>17</td>
<td>11.2</td>
</tr>
<tr>
<td>Electronic parts and components</td>
<td>19</td>
<td>9.0</td>
<td>22</td>
<td>14.5</td>
</tr>
<tr>
<td>Electrical machinery and equipment</td>
<td>18</td>
<td>8.5</td>
<td>21</td>
<td>13.8</td>
</tr>
<tr>
<td>Transport equipment</td>
<td>23</td>
<td>10.8</td>
<td>15</td>
<td>9.9</td>
</tr>
<tr>
<td>Textiles and leather</td>
<td>5</td>
<td>2.4</td>
<td>3</td>
<td>2.0</td>
</tr>
<tr>
<td>Paper products</td>
<td>3</td>
<td>1.4</td>
<td>8</td>
<td>5.3</td>
</tr>
<tr>
<td>Machinery and equipment</td>
<td>23</td>
<td>10.8</td>
<td>10</td>
<td>6.6</td>
</tr>
<tr>
<td>Basic metal products</td>
<td>5</td>
<td>2.4</td>
<td>3</td>
<td>2.0</td>
</tr>
<tr>
<td>Clothing and footwear</td>
<td>5</td>
<td>2.4</td>
<td>6</td>
<td>3.9</td>
</tr>
<tr>
<td>Other</td>
<td>10</td>
<td>4.7</td>
<td>5</td>
<td>3.3</td>
</tr>
<tr>
<td>Total</td>
<td>212</td>
<td>100</td>
<td>152</td>
<td>100</td>
</tr>
</tbody>
</table>

Table II. Industry analysis

Supply chain and business strategy
and business performance were measured using a five-point Likert scale (1 = poor, 5 = excellent). Questions of cost leadership, differentiation and business performance compared to the company’s competitors were asked.

The study sample was examined to determine whether non-respondent manufacturers differed significantly from those responding regarding key characteristics (i.e. sales and number of employees), as suggested by Armstrong and Overton (1977), Gerbing and Anderson (1988) and Wiengarten et al. (2012) to check the existence of non-response bias. The result revealed that the sample did not suggest the presence of non-response bias. In addition, a measurement invariance test was applied since samples were taken randomly from the UK and South Korea, to identify the critical assumption that the basic structure of the model is stable across cultures, and that individuals in different countries use its scale in a similar manner (Turker, 2009; Malham and Saucier, 2014). Thus, through multi-group confirmatory factor analyses (CFA) the results of fit indexes suggest that measurement invariance is supported across the two countries ($\chi^2$/df = 2.09, RMSEA = 0.054, SRMR = 0.606, CIF = 0.903). For common method bias, Harman’s one-factor test was conducted using all variables. Since no single factor was apparent in the un-rotated factor structure, common method variance does not exist.

4. Results

4.1 Reliability and validity

Both exploratory factor analysis (EFA) and CFA were conducted using AMOS 22 to examine reliability and validity since items were collected from different studies. First, all constructs were assessed with CFA and the measurement model offered a satisfactory fit ($\chi^2_{213} = 347.507$, GFI = 0.925, SRMR = 0.037, RMSEA = 0.042 and CFI = 0.965) with acceptable factor loadings (> 0.59). Composite reliability (CR) also presented internal consistency (CR > 0.754). Convergent validity was confirmed as all factor loadings exceeded 0.5, as well as acceptable average variance extracted (AVE > 0.506). Table III indicates the result of CFA indicators. In addition, there is no case where the square of the correlation between a pair of constructs is greater than the AVE of the constructs. Thus, discriminant validity using the procedures suggested by Fornell and Larcker (1981) was confirmed (see Table IV).

Second, Cronbach’s $\alpha$ was tested to measure the reliability of the scale items before the EFA. The results yielded acceptable $\alpha$ values (Nunnally, 1978) and six structures showed acceptable internal consistency ( > 0.795). EFA was then applied to check construct validity (McDonald, 1981; Hattie, 1985). Six factors were extracted with eigenvalues greater than 1, and six structures explained 68.0 per cent of the total variance and all of the loadings above the minimum cut-off, 0.5 (Hair et al., 2010). EFA results showed all items with a high within-factor loading (i.e. convergent validity) and a low cross-factor loading (i.e. discriminant validity). The results of separate EFAs for both the UK and Korea also showed acceptable reliability and validity with the same items loaded. Thus, the variables of the construct have a consistent and stable structure across the groups (Floyd and Widaman, 1995). EFA and CFA confirmed the stability of all constructs for structural equation modelling (SEM) and $t$-test. Table V presents the EFA pattern matrix.

4.2 Structural equation modelling

To test our proposed theoretical and empirical model, SEM analysis was conducted. The result from samples shows that increased product variety and SC agility improve differentiation, while a decrease in product variety and an increase in cost efficiency improve cost leadership; thus, $H1a$ and $H1b$ and $H2a$ and $H2b$ are accepted. Also, business performance is associated with cost leadership and differentiation that mediates the relationship between product variety and business performance. Therefore, $H3$ and $H4$
are accepted. Table VI indicates the results of the SEM. Both high product variety and SC agility have a significant impact on the differentiation at 0.001 and 0.01 levels, respectively. In addition, low product variety and cost efficiency are significantly related to cost leadership at 0.05 and 0.001 levels, respectively. Product variety did not impact directly on
business performance ($p = 0.557$); instead, cost leadership and differentiation mediate the relationships at 0.05 and 0.001 levels, respectively.

### 4.3 Comparison between UK and Korea

To test the research findings through a cross-examination a t-test was conducted. Regarding business strategies, cost leadership ($p < 0.01$) and differentiation ($p < 0.1$) varied significantly between the two countries. In the case of performance measures, SC agility ($p < 0.001$), SC cost efficiency ($p < 0.01$) and business performance ($p < 0.05$) showed significant differences. With regard to mean value, the UK (mean = 3.51) exhibited a
sharper focus on differentiation than Korea (mean = 3.36), while Korea (mean = 3.42) focused more on cost leadership than the UK (mean = 3.18). In addition, Korea scored better on cost efficiency than the UK. However, Korea (mean = 3.39) exhibited superior SC agility to the UK (mean = 3.11). Lastly, the UK (mean = 3.70) had better business performance than Korea (mean = 3.52). The results imply that the UK focuses on a higher level of product variety and differentiation strategy (see Stavrulaki and Davis, 2010), which leads to superior business performance compared with Korea. Korea focuses more on cost leadership with lower-level product variety, which leads to higher cost efficiency. Table VII explains the comparison of \( t \)-tests between the two countries.

### 5. Discussion

The study demonstrated the alignment of business strategy between product variety and SC through SEM analysis. The test justifies the appropriate business strategy according to companies’ level of product variety and matching SC focus. A high product-variety policy is matched with agile SC as a differentiation strategy, while a low product-variety policy matches lean SC as a cost leadership strategy. High product variety with efficient SC can follow a “mass customisation” strategy (e.g. automobile and computer industries such as BMW and Dell) while low product variety with an agile SC can be regarded as a “hybrid” strategy (e.g. perishable food and electronic components industries such as Aldi and Intel). Therefore, the survey results support the concept that companies employing differentiation strategies by providing different product variety focus more on SC agility, customer service and relationship, which results in a high margin. Companies employing a cost leadership strategy with low product variety ensure SC efficiency through economy of scale, which leads to cost reduction.

However, product variety does not always guarantee improved business performance \((p = 0.557)\). Given the cost burden from the provision of product variety, companies should link their strategic alignment and capabilities with either SC agility or cost efficiency. Thus, decisions on differentiation, cost leadership or hybrid strategies can be considered based on the existing level of product variety and/or intention of NPI as well as SC capability supported. Especially business strategies such as differentiation and cost leadership significantly mediate the relationship between product variety and business performance. Any mismatched alignment cannot achieve the better business performance. The perception that product variety has the potential to positively influence market share and returns (Otero-Neira et al., 2010) is supported only through appropriate and matching business strategies (i.e. mediating factor). The results also support that an agile system is matched with differentiation, while cost efficiency is matched with a cost-leadership strategy (Hallgren and Olhager, 2009). At the functional level, the cost burden can be minimised by use of advanced technologies and appropriate strategies such as modularity, cellular manufacturing, information technology, postponement and SC integration strategies.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>UK</th>
<th>Korea</th>
<th>Total</th>
<th>( T )</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product variety</td>
<td>3.74</td>
<td>3.04</td>
<td>3.44</td>
<td>5.195****</td>
<td>0.000</td>
</tr>
<tr>
<td>SC agility</td>
<td>3.11</td>
<td>3.39</td>
<td>3.23</td>
<td>-3.581***</td>
<td>0.000</td>
</tr>
<tr>
<td>SC Cost efficiency</td>
<td>3.31</td>
<td>3.52</td>
<td>3.40</td>
<td>-2.963**</td>
<td>0.003</td>
</tr>
<tr>
<td>Cost leadership</td>
<td>3.18</td>
<td>3.42</td>
<td>3.28</td>
<td>-3.156**</td>
<td>0.002</td>
</tr>
<tr>
<td>Differentiation</td>
<td>3.51</td>
<td>3.36</td>
<td>3.45</td>
<td>1.858*****</td>
<td>0.064</td>
</tr>
<tr>
<td>Business performance</td>
<td>3.70</td>
<td>3.52</td>
<td>3.63</td>
<td>2.591*</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Notes: *\( p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.1\)

Table VII. \( t \)-test of structures by country
The study extends the concept to the national level to justify the implications of the research findings. The UK shows higher levels of variety, differentiation and business performance while Korea exhibits higher performance in cost efficiency, cost leadership and SC agility, as supported by Stavrulaki and Davis (2010) and Agarwal et al. (2006). Higher levels of the production-dominant variety (i.e. increased production cost outweighs increased market mediation cost) are positively related with high-volume production (Randall and Ulrich, 2001). Thus, Korea appears to focus more on scale-efficient production (e.g. production-dominant variety) with a lower level of product variety than the UK, which has unstable demand due to greater product variety. However, it is notable that Korea has higher SC agility performance than the UK, which runs counter to the expectation that the higher the level of product variety the higher the level of agility. One explanation is Korea’s high dependence on exports and manufacturing-based industry, as supported by Antonio et al.’s (2015) research. SC activities for a quick response (i.e. agility) as a form of distinctive competence enable firms to achieve competitive export advantages (Piercy et al., 1998). Also, SC agility is a critical factor affecting overall global competitiveness (Swafford et al., 2006), especially for an export-based country. Christopher et al. (2006) concluded that agility and responsiveness are increasingly fundamental to competitive success in global business activity, such as global sourcing, offshore manufacturing and export. Thus, a global SC should be sufficiently agile to allow firms to improve their business performance and manage demand and supply uncertainty by being more responsive to unexpected change. Thus, a country concentrating on export competitiveness must achieve a global SC network structure with a high level of agility. For example, the total annual volume of exports in Korea ($527bn) is lower than for the UK ($674bn); however, export dependability accounted for a higher percentage of the GDP in Korea (36 per cent) as compared to the UK (23 per cent) in 2015 (International Monetary Fund, 2016; Central Intelligence Agency, 2016).

At the organisational level, the proximity of production facilities to the target market also enhances SC agility (Lee, 2004). Thus, companies that have a high level of product variety require careful consideration of their focus on local production and distribution to improve SC agility as well as variety management strategies such as modularity and postponement for cost reduction. In addition, better logistics performance can boost SC agility (see Arvis et al., 2012). Nowadays, the dominant trend in manufacturing industry is not towards pure customisation or pure standardisation, but towards the middle position of customer involvement (Lampel and Mintzberg, 1996), which is mass customisation. Thus, manufacturers that provide a low level of product variety may consider a structural shift to mass customisation in order to benefit from their SC agility capability. However, the movement should match the SC strategy and be supported by variety management activities.

6. Conclusions
The study investigated the relationships between product variety, SC and business strategies to improve business performance. First, the study demonstrated the importance of strategic alignment between product variety and SC focus, and suggested a theoretical justification for the choice of business strategies through the proposed model. A high level of product variety matches agile SC, while low product variety is aligned with SC cost efficiency. However, the level of product variety improves business performance only through matching business strategies such as cost leadership and differentiation. Therefore, an appropriate business strategy, considering the level of product variety and matching SC strategy, achieves the better business performance. Second, the overall comparison between the two countries found manufacturers in the UK demonstrating higher levels of product variety and differentiation (see Stavrulaki and Davis, 2010) than manufacturers in Korea. Instead, Korea displayed higher cost leadership attributes and cost efficiency than the UK. However, Korea shows better agile SC performance
than the UK, which indicates the potential to increase product variety without incurring a significant cost burden.

This research makes several contributions. For theoretical implications, confirmative and exploratory factor analyses formed the concepts of each structure. Then, SEM confirmed the strategic relations between product variety and SC focus to improve business performance. Also, the comparison between manufacturers in the UK and Korea supports the fact that a high product-variety context focuses more on differentiation while a low product-variety context focuses on cost efficiency and cost leadership (Stavrulaki and Davis, 2010). Regarding practical contributions, the findings suggest manufacturers appropriate SC strategies based on the existing level of product variety in order to achieve competitive advantage. Such insight is particularly valuable for manufacturing concerns that are considering changing the heterogeneity of their product base through increased product variety for better business performance.

There are several limitations associated with this research. First, it investigated exclusively manufacturing industries in the UK and Korea. This particularity may limit the generalisability of the findings to other populations with different competitive, environmental, economic and cultural characteristics (Hughes and Morgan, 2008; Antonio et al., 2015). In addition, the growth and development of SCs is not driven only by internal motives, but also by a number of external factors, such as increasing globalisation, reduced barriers to international trade, advances in information technology, environmental concerns and government regulations (Gunasekaran et al., 2004). An appropriate topic for future research concerns the practical approach for multinational corporations to cope with variety and global SC issues without sacrificing costs.

References


Further reading


Corresponding author

Juneho Um can be contacted at: juno9782@hotmail.com

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com
Measuring resource efficiency and resource effectiveness in manufacturing

Mangesh Gharfalkar
School of Science and Engineering, Teesside University, Middlesbrough, UK

Zulfiqur Ali
Teesside University, Middlesbrough, UK, and

Graham Hillier
Teesside University, Middlesbrough, UK and

Centre for Process Innovation Ltd, Redcar, UK

Abstract
Purpose – The purpose of this paper is twofold: first, to identify and analyse existing resource efficiency (RE) and resource effectiveness measures and indicators (REMI); and second, to identify gaps and develop a new indicator of “operational resource effectiveness” (OREf) suitable for manufacturing units.

Design/methodology/approach – Research methodology consists of three stages; gap identification, development and testing. Through review of academic literature, 40 REMIs are identified and analysed. A survey of manufacturers is carried out to validate the hypothesis and seek inputs on the development of the new indicator. The proposed indicator is tested by comparing OREf index of two manufacturing units with each other, with resource intensity per unit (RIPU), waste intensity per unit (WIPU) and with four other REMIs.

Findings – Analysis of 40 REMIs clearly points towards the absence of a hypothesised REMI. In total, 78 per cent of manufacturers surveyed in north England substantiate the hypothesis. Inverse correlation established between the proposed OREf indicator, RIPU, WIPU and other comparisons is likely to validate the output generated by the proposed indicator.

Research limitations/implications – Testing of this indicator is limited to two dissimilar manufacturing units that shared data.

Practical implications – The proposed indicator is useful for comparing the operational resource effectiveness of individual factories over a period as well as with other factories. RIPU and WIPU captured in this indicator also represent operational RE that can be used to initiate improvement action.

Originality/value – Inclusion of both, the resource consumption and the waste generation along with discount/multiplying factors that capture the circularity aspects is likely to be the distinguishing feature of this indicator.

Keywords Sustainability, Manufacturing, Performance indicators, Performance measures, Resource efficiency, Resource effectiveness

Paper type Research paper

1. Introduction

“Humankind has consumed more aluminium, copper, iron and steel, phosphate rock, diamonds, sulphur, coal, oil, natural gas, and even sand and gravel over the past century than over all earlier centuries put together, and the pace continues to accelerate” (Tilton, 2003). With rapidly increasing consumption of energy and material resources in the developed as well as the developing world, the issue of resource scarcity is becoming vital. The resource efficiency (RE) programme by United Nations Environment Programme (UNEP) emphasise that to meet the needs of the growing population, it is necessary to “decouple resource use and environmental degradation from the economic growth”. This will necessitate consumers in making social and environmental concerns, part of their

Mangesh Gharfalkar acknowledges the support offered by Teesside University through a PhD scholarship to enable this research.
buying decisions. It will require producers to change their design, production and marketing processes (UNEP, 2014). Duflou et al. (2012) argue that while the manufacturing sector plays a vital role in the world economy, it consumes significant amounts of energy and other natural resources and releases solid, liquid and gaseous wastes that lead to increased stress on the already fragile environment. Parker (2007) observes that unless new approaches to manufacturing are found and implemented, global population growth alone is expected to cause emissions and waste production to increase by at least 40 per cent by 2050.

Measuring, monitoring and improving RE and/or resource effectiveness can be one of the approaches to addressing the issue of resource scarcity highlighted above. This research aims to identify and analyse some of the existing RE and resource effectiveness measures and indicators (REMs); identify gaps and develop a “new indicator” of “operational resource effectiveness” (OREft) suitable for manufacturing units.

2. Literature review
2.1 RE and resource effectiveness
Cambridge dictionary defines “efficiency” as “good use of time and energy that does not waste any” and being “effective” is defined as “successful or achieving the results you want”. Efficiency and effectiveness can be differentiated between how well something is done (efficient) and how useful something is (effective) (Dffen, 2015). In his book titled “The Effective Executive”, Peter Drucker aptly differentiates the two by stating that “efficiency is doing the thing right and effectiveness is doing the right thing”. Kao et al. (1995) argue that a conversion process normally involves many intricate activities, many inputs and many outputs that limit the level to which efficiency gains can be achieved. Fearne and Fowler (2006) observe that there is evidence to suggest that focus on “efficiency” considerations undermines the need for delivering projects “effectively” against the set objectives.

UNEP (2010) defines RE from the perspective of value chain and product life cycle as “reducing the total environmental impact of the production and consumption of goods and services, from raw material extraction to final use and disposal”. In a policy document, Jansen (2013) highlights the fact that the current focus of RE of European Union Member States is restricted to improving the efficiency of use of input “natural resources” such as fossil fuels, rare earth metals, and water. It further elaborates on the European Commission’s (EC) flagship initiative of “Resource Efficient Europe” that defines resources to include all-natural resources that act as inputs to a nation’s economy. The EC captures the essence of RE by defining it as “A way to deliver more with less (natural resources)”. Similarly, the Australian Environment Protection Agency (EPA) defines RE as “doing more with less – creating more value with less impact” (EPA – Tasmania, 2013). The Australian EPA further describes RE in business terms as “process optimisation to limit consumption of energy, water and materials and output of waste products”. Although “RE” policies cannot by themselves reduce exposure to sudden shortages or rise in prices, they can surely reduce their impacts. Shortages and sudden price rises on world market are quite often created by speculation, man-made and natural disasters, geopolitical crises or rising demand in a specific application. Economic resilience and “environmental sustainability” can only be achieved with contributions from all members of the value chain across the globe working towards achieving RE. Otherwise, pressure on reducing resource consumption in only one economic block could see shifting of economic activities to less efficient parts of the world. This in turn is likely to increase pressure on Earth’s bio capacity as a whole (Euromines, 2011).

In the context of “environmental sustainability”, there is no formal definition of “resource effectiveness”. It could be defined as “To manage and optimise consumption of non-renewable and hazardous natural resources with an objective of achieving environmental sustainability”. Management and optimisation could include complete
elimination or reduction in the consumption of non-renewable natural resource(s) and/or replacement of non-renewable natural resource(s) with renewable natural resource(s). It could also include complete elimination or reduction in consumption of hazardous natural resources and/or replacement of hazardous natural resources with environmentally benign natural resources.

The strategic objective of “environmental sustainability” cannot be achieved even with 100 per cent RE at each stage of the supply chain. This is because non-renewable natural resources are finite. Therefore, to achieve the strategic objective of “environmental sustainability”, manufacturers may have to be “resource efficient” as well as “resource effective”. The “circular economy” business model seems to be the desirable approach to doing things right (efficiently) as well as doing the right things (effectively). The “circular business model” ensures not only recovery, reprocessing and reuse of waste streams but also replacement of non-renewable natural resources with renewable natural resources and replacement of hazardous resources with environmentally benign resources. Gharfalkar et al. (2015) capture the circularity aspect in the “5Rs of Resource Effectiveness” (Figure 2). In the context of manufacturing, it could be termed as “Resource Effective Manufacturing” (REftM). REftM could be defined as “Manufacturing environmentally benign products using nil or reduced quantity of non-renewable and hazardous natural resources that eliminates or reduces the generation of environmentally damaging waste streams”.

2.2 Need for measuring RE or resource effectiveness
Huysman et al. (2015) observe that the transition towards more resource efficient economies that is necessitated by challenges related to natural resources will need “quantitative indicators” that are able to track consumption of “natural resources” and the impacts associated with production and consumption systems. The EC highlights the importance of changing consumption patterns and improving products where consumers would buy products that last longer and/or products that could be easily reused or recycled. To achieve the objective of “sustainable development”, the EC’s initiative on “Resource Efficient Europe”, emphasises the need for mandatory as well as voluntary “measures of resource efficiency”. It highlights the need for developing robust and easily understandable “indicators” that will provide signals and measure the progress of RE. The EC wants Member States to put in place incentives to motivate companies to “measure, benchmark and improve their RE systematically” (EC, 2011). Therefore, to improve RE and/or resource effectiveness, it is necessary to assess it using appropriate measures and/or indicators of RE or resource effectiveness (REMI). Gaussin et al. (2013) observe that as indices become more comprehensive, they get more complicated and often include large number of “difficult-to-quantify” parameters such as societal impact.

2.3 Measures and indicators of RE and effectiveness
Oxford Dictionary defines a “measure” as “to ascertain the size, amount or degree of (something) by using an instrument or device marked in standard unit” and defines an “indicator” as “a thing that indicates the state or level of something”. Cambridge dictionary defines a “measure” as “to discover the exact size or amount of something” and defines an “indicator” as “something that shows what a situation is like”. For example, while the amount of solid waste generated can be considered as a “measure”, solid waste generated per unit of production could be considered as an “indicator” that affects environmental sustainability.

This section deals with the identification of existing RE and resource effectiveness measures and/or indicators (REMI). The literature search is conducted by identifying peer reviewed articles published in English language using the “Discovery” database search engine. All fields (titles, subject terms (keywords) and abstracts) of literature in these databases are Boolean searched using the search phrases “Resource Efficiency Indicator” or
“Resource Efficiency Index” “Resource Efficiency Measure” or “Resource Effectiveness Indicator” or “Resource Effectiveness Measure” or “Resource Effectiveness Index” for the period beginning 1987–2017. The publication of the Brundtland Commission report in 1987 made “sustainable development” prominent for the first time. Therefore, the cut off year for literature search is set as 1987. Overall criteria for selection of relevant literature and the number of useful articles identified through this process are summarised in Table I. Forty REMIs that are identified because of this search are summarised in Table II.

3. Research methodology
As depicted in Figure 1, the research methodology consists of three stages: gap identification, development and testing. The research is based on the foundation of two streams of investigation: literature survey and industry survey. Apart from identification of some of the existing REMIs, the literature survey aimed to understand the “resources” that are relevant for achieving “environmental sustainability” in manufacturing. It also aimed to understand the contextual background of measuring RE and/or resource effectiveness in achieving “environmentally sustainability”. Both these lines of investigation are used to identify gaps in some of the existing REMIs that are used for the development of a “new indicator”.

This research attempts to overcome some of the problems of complexity and assumptions by focussing on a few but important elements of “resource consumption” and “waste generation” for which operational data is easily available within a manufacturing unit (Figure 3). Scope of this research is limited to developing an aggregate indicator for measuring “operational resource effectiveness” (OREft) of existing manufacturing units.

The proposed indicator is based on the following hypothesis:

An indication of resource effectiveness of a manufacturing unit can be obtained by combining “input measures” that capture “consumption of key natural resources” with “output measures” that capture “generation of waste”, based on operational data that is easily available within the manufacturing unit.

3.1 Gap identification
This stage that has two strands of investigation involves identification of some of the existing REMIs and areas for improvement in the identified REMIs:

(1) literature survey; and
(2) industry survey.

Published literature is used to identify some of the existing REMIs. Identified REMIs are analysed using a set of qualitative and quantitative criteria. The quantitative criteria are summarised in Table I. The second strand of investigation includes a web-based survey of manufacturers in north England to understand if they use any REMIs. Both strands of investigation attempt to capture the REMIs that are in use and whether any of those REMIs capture both, “resource use” and “waste generation” in its measurement. It also

<table>
<thead>
<tr>
<th>Criterion/description</th>
<th>Total documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articles/documents published in English language between 1 January 1987 and 14 July 2017</td>
<td>149</td>
</tr>
<tr>
<td>Scholarly and peer reviewed journal articles, conference proceedings, dissertation/thesis, book chapter and reports</td>
<td>90</td>
</tr>
<tr>
<td>Articles/documents restricted to engineering, environmental sciences, business and ecology disciplines</td>
<td>26</td>
</tr>
</tbody>
</table>

Table I. Summary of database search
attempts to understand whether the current measurements are based on operational data available within the manufacturing unit. The industry survey also assesses the level of data availability for various elements of the proposed “Operational Resource Effectiveness” (OREf) indicator identified in Figure 3. Findings are used in the development of the new indicator.

3.1.1 Criteria for analysis of identified REMIs. Mostly qualitative analysis of REMIs has been undertaken. For example, Moffatt et al. (2001) assess a number of RE measures based on three sets of qualitative criteria such as robustness, practicality and usefulness to policy makers.
Similarly, Hirschnitz-Garbers and Srebotnjak (2012) use a set of six qualitative criteria such as LCA compatibility, coverage of industries, sustainability impact coverage, policy relevance, required data efforts and data availability. Each of these measures is qualitatively ranked as low, medium or high under each of the six key criteria.

In this research, qualitative as well as quantitative analysis of 40 identified REMIs is carried out. Points are allocated to different criterion under each of the three categories, whose scores are summarised in Table III. Since all categories do not score equally, they are mean normalised for parity. Each of the three categories is further divided into individual criterion that is scored individually depending on its relevance and importance to "environmental sustainability".

As the focus of this research is on developing an aggregate “operational resource effectiveness” (ORE\textsubscript{a}) indicator suitable for manufacturing units, lower criterion scores for REMIs occur at global or national level (score = 0) than those that can measure RE for a product across its life cycle (score = 5). Examples of criterion scores for different boundary line suitability of REMIs are as follows:

- REMI suitable for measuring RE at global and/or national level only = 0.
- REMI suitable for measuring RE of individual factory (Gate2Gate) = 1.
- REMI suitable for measuring RE of individual process (Gate2Gate) = 2.
- REMI suitable for measuring RE of each product (Gate2Gate) = 3.
- REMI suitable for measuring RE of product across supply chain (Cradle2Gate) = 4.
- REMI suitable for measuring RE of product across life cycle (Cradle2Grave) = 5.

The “resource consumption” related category (Table III) scores more than the other two categories. This is because it deals with various aspects of “resource consumption”, which is the key element of environmental sustainability. Also, this category has the maximum number of subcategories in it. Each subcategory is further divided into number of individual criterions. For example, the subcategories include capturing of different types of energy (non-renewable, recovered and renewable), materials (non-renewable, reused, reprocessed and renewable), water...
(fresh and recovered), etc. The assessment criterion assigns higher scores for REMIs that separately capture and discount “recovered” resources and maximum score for capturing and discounting “renewable” resources. Examples of the individual criterion scores for different types of energy captured by each REMI in its calculation are as listed below:

- REMI does not capture consumption of renewable energy = 0.
- REMI aggregates consumption of renewable and non-renewable energy = 1.
- REMI separately captures consumption of renewable and non-renewable energy = 2.
- REMI discounts consumption of renewable energy = 3.

Also, since the aim is to develop an aggregate $ORE_d$ indicator that “simultaneously” capture number of “key elements of resource efficiency or resource effectiveness” in its calculation, higher scores are allocated to REMI that capture more “key elements of resource efficiency or resource effectiveness” in its measurement. As hypothesised, “consumption of key natural resources” and resultant “waste generation” are considered as the “key elements of resource efficiency or resource effectiveness” (Figure 3). Therefore, while most other criteria are scored on a band of 0–5 in increments of 1, a score of 0 or 5 is allocated to each of the “key elements of resource efficiency or resource effectiveness”. These include key natural resources such as “energy”, “materials”, “water” and “land” use on the “consumption side” and “greenhouse gases”, “effluent” and “solid waste” on the “output side”. Individual criterion scores for these “key elements of resource efficiency or resource effectiveness” are listed below. A REMI can score 5 in more than one “key elements of resource efficiency or resource effectiveness” only if those “key elements of resource efficiency or resource effectiveness” appear simultaneously in its calculation:

- REMI captures energy consumption in its measurement = 5.
- REMI captures consumption of materials in its measurement = 5.
- REMI captures consumption of water in its measurement = 5.
- REMI captures land use in its measurement = 5.
- REMI captures generation of greenhouse gases (GHGs) gases in its measurement = 5.
- REMI captures generation of effluent waste in its measurement = 5.
- REMI captures generation of solid waste in its measurement = 5.

3.1.2 Method used for industry survey. A web-based survey is carried out with manufacturers in north England. The target audience include businesses from the manufacturing, engineering and processing industry, classified as “manufacturers” by the office of national statistics (ONS). Financial Analysis Made Easy) database is used to e-mail manufacturers. In total, 86 responses are received. The survey consists of total 44 questions but not all questions are applicable for all respondents. It is divided into four sections: 1 (consent form), 2 (about the respondent and his/her business), 3A (reasons for not measuring RE), 3B (how RE is measured in the organisation) and 4 (inputs for the development of the new indicator). Sections 1, 2 and 4 are applicable for all respondents.

3.2 Development stage
Based on the foundation of the hypothesis statement, this stage includes following aspects in the development of a conceptual framework and the algorithm for the new indicator of operational resource effectiveness ($ORE_d$):

(1) Seek inputs from the gaps identified from analysis of 40 REMIs and from the results of the survey of manufacturers in north England.
(2) Identify elements or variables of the proposed ORE$_t$ indicator. This includes decision on the resources and waste categories to be included in the proposed indicator. The 5Rs of resource effectiveness (Figure 2) and alternative hierarchy of resource use proposed by Gharfalkar et al. (2015) are also used in this decision making.

(3) Introduction of circularity factors to differentiate various categories of resource use and waste generation. In absence of any academic research; policy guidelines or industry practices on circularity factors, ratios of Fibonacci numbers are used for this purpose. Even if the ratios of integer numbers were used, the relative outcome would have been the same.

(4) Decision on the units of measurement of each of the identified elements (variables) of the proposed indicator. To make the indicator unit free, all elements of the proposed ORE$_t$ indicator including production units are converted into the same unit of mass. For example, on the “resource consumption” side, energy is converted into tons of oil equivalent, water and materials into tons. On the “waste generation” side, GHG is converted into tons of carbon equivalent, effluent and solid wastes into tons.

### 3.2.1 Theory behind the proposed ORE$_t$ indicator

As resource effectiveness can be considered as one of the performance measures for achieving environmental sustainability, it is necessary to understand the philosophy of performance measurement. Neely et al. (1995) define performance as the efficiency and effectiveness of an action and performance measurement as the process of quantifying action. Stefan (2004) defines performance measure as a metric used to quantify the efficiency and/or effectiveness of an action that supports strategic objective. Bernolak (1997) observes that the data requirements should be limited to the necessary detail and frequency.
The concept of “overall equipment effectiveness” (OEE) provided by Seiichi Nakajima is identified as suitable for developing the proposed OREft indicator. While OEE is calculated by multiplying three different types of efficiencies: namely, availability, performance and quality, OREft of a factory can be calculated by multiplying the efficiency or effectiveness of different elements of “resource use” with the efficiency or effectiveness of different elements of “waste generation” identified in Figure 3. The proposed indicator takes into consideration following underlying principles that are used for the development of the hypothesis statement:

1. Natural resources are scarce. Therefore, for achieving the strategic objective of “environmental sustainability”, the RE and/or resource effectiveness indicator should take into consideration consumption of key natural resources and ignore other resources such as time, money or manpower.

Figure 3.
Elements of proposed OREft indicator
(2) An indication need not be accurate and therefore it may not be necessary to capture all variables of environmental sustainability in its measurement. Therefore, the proposed indicator should capture only the most important variables of environmental sustainability (not all) such as energy, raw materials, water and waste.

(3) Consumption of every natural resource has an impact, and a different impact, on the environment. Therefore, the indicator should not only capture the consumption of key natural resources but also the generation of waste.

(4) Many of the existing REMIs are complex and dependent on data outside the organisation and also on assumptions. Complex indicators are often not measured and monitored especially if they are dependent on data from multiple sources and/or if they are based on a set of assumptions. For adoption by the industry, measures or indicators must be based on readily available operational data rather than on assumptions.

3.2.2 Scope and system boundaries of proposed $O_{R_E}$ indicator. For the purpose of this research, resources are grouped into two categories depending on their importance to “environmental sustainability”. The first group is defined as “primary resources” and includes the “natural resources” that are primarily responsible for “environmental sustainability”. The second group is defined as “secondary resources” and comprise of “natural” and “human made resources” that play a secondary role in “environmental sustainability”:

(1) Primary resources: raw materials, consumables (water), energy (oil, gas, coal, etc.), waste streams.

(2) Secondary resources: time, human capital and money capital.

Since the strategic objective is to support “environmental sustainability”, scope of the proposed indicator is limited to primary resources such as raw materials, water, energy and waste. It excludes secondary resources such as time, money (capital) or human capital.

On system boundaries, Huysman et al. (2015) observe that RE indicators have been developed for systems at micro-scale of specific processes and products to meso-scale and macro-scale of sectors and countries. At micro-scale, some indicators capture products and processes from factory entry gate to factory exit gate (Gate2Gate) while others consider full life cycle. Some indicators evaluate RE at regional or national level while others consider a more global perspective by including resources that are embodied in imported products.

The proposed indicator developed around the system boundary of a “business unit” or a “factory” is defined as the Gate2Gate $O_{R_E}$ indicator. It can measure “operational resource effectiveness” for each “business unit” or a “factory” from its entry gate to exit gate (Gate2Gate). As in the case of the OEE, and as hypothesised, the scope of the proposed indicator is restricted to operational data. This aspect is substantiated by the industry survey (Figure 5). Also, an indicator that aims to be perfect by attempting to capture all aspects of environmental sustainability end up being too complex, lacks data availability and unless mandatory is not accepted by the industry.

3.2.3 Elements of the proposed $O_{R_E}$ indicator. The elements (variables) of the proposed indicator identified in Figure 3 are based on the circularity principles of the “5Rs of Resource Effectiveness” (Figure 2). To capture key elements of “resource use” and “waste generation”, the framework considers the third “R” that consists of “recovery” options such as “reuse” and “reprocessing”. These “recovery” options lead to the conversion of a “waste” into a “non-waste” (resource). The European waste directive 2008/98/EC defines “waste” as “any substance or object which the holder discards or intends to discard or is required to discard” (Directive, 2008). Elements of the proposed indicator
takes into consideration the resource flows that could be measured in physical units of materials, energy and water flows as summarised in Table IV. To support the primary objective of “environmental sustainability” only “primary resources” categorised below are considered in the proposed indicator:

(1) Virgin resources:
- renewable virgin resources; and
- non-renewable virgin resources.

(2) Recovered resources:
- reused (via repair, recondition, refurbish, remanufacture); and
- reprocessed (upcycled, recycled, down-cycled).

3.2.4 Equations of the proposed Gate2Gate ORE_{dt} indicator. The concept behind “Material Intensity per Unit Service” (MIPS or M1 in Table II) is used for capturing each element of the proposed indicator identified in Figure 3. MIPS is calculated as mass of material input (MI) per total units of service (S) (Hinterberger and Schmidt-Bleek, 1999). Like MIPS, the proposed indicator captures consumption of different resources and generation of different wastes per unit of production in a factory. The proposed indicator is based on the resource flows that can be measured in physical units of materials, energy and water flow on the input side and flow of waste streams such as GHG, effluent, solid and hazardous waste on the output side (Figure 3). If product and/or process wise operational data for each element of the proposed ORE_{dt} indicator identified in Figure 3 is available, then product and/or process wise ORE_{dt} can be also assessed. But if it is not available, then all products manufactured in a factory need to be assigned the ORE_{dt} of that factory:

\[
\text{Gate2Gate ORE}_{dt} = \text{Resource intensity per unit} \times \text{waste intensity per unit},
\]

\[
\text{Gate2Gate ORE}_{dt} = \text{RIPU} \times \text{WIPU}. \tag{1}
\]

On the resource consumption side equations, following abbreviations are used:
- RIPU, resource intensity per unit.
- WIPU, waste intensity per unit.
- EIPU, energy intensity per unit.
- MIPU, material intensity per unit.
- WtrIPU, water intensity per unit.

<table>
<thead>
<tr>
<th>Input</th>
<th>System boundaries</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Materials</td>
<td>Nations</td>
<td>Products</td>
</tr>
<tr>
<td>Energy</td>
<td>Regions</td>
<td>Waste</td>
</tr>
<tr>
<td>Water</td>
<td>Corporations</td>
<td>Emissions</td>
</tr>
<tr>
<td></td>
<td>Processes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Products</td>
<td></td>
</tr>
</tbody>
</table>

Table IV. System boundaries for mass balance

Source: Jasch (2002)
RIPU = EIPU + MIPU + W_{tr}IPU, \hspace{1cm} (2) \hspace{1cm} \text{Measuring RE}

where:

\begin{align*}
EIPU &= \text{(Energy consumption)/(Production units)}, \hspace{1cm} (2a) \\
MIPU &= \text{(Material consumption)/(Production units)}, \hspace{1cm} (2b) \\
W_{tr}IPU &= \text{(Water consumption)/(Production units)}. \hspace{1cm} (2c)
\end{align*}

Next level of elements of resource use as identified in Figure 3 is captured as follows:

\begin{align*}
\text{Energy} &= \text{New energy} + \text{Recovered energy}, \hspace{1cm} (2a.1) \\
\text{New energy} &= \text{Renewable energy} + \text{non} - \text{renewable energy}. \hspace{1cm} (2a.2)
\end{align*}

As explained in the previous sections, consumption of only primary raw material(s) is considered in the consumption of materials:

\begin{align*}
\text{Material} &= \text{Virgin material} + \text{Recovered material}, \\
\text{Virgin material} &= \text{Renewable material} + \text{Non} - \text{renewable material}, \\
\text{Recovered material} &= \text{Reused material} + \text{Reprocessed material}, \\
\text{Water consumption} &= \text{Fresh water} + \text{Recovered water}.
\end{align*}

On the waste generation side equations, following abbreviations are used:

- WIPU, waste intensity per unit.
- GHGIPU, greenhouse gases emissions intensity per unit.
- E_{eff}IPU, effluent intensity per unit.
- SWIPU, solid waste intensity per unit.
- E_{eff}, effluent.
- SW, solid waste.
- Haz, hazardous.
- Nhaz, non-hazardous.

\begin{align*}
\text{WIPU} &= \text{GHGIPU} + \text{SWIPU} + \text{E}_{\text{eff}}\text{IPU}, \hspace{1cm} (3)
\end{align*}

where:

\begin{align*}
\text{GHGIPU} &= \text{(Quantity of greenhouse gas generated)/(Production units)}, \hspace{1cm} (3a) \\
\text{SWIPU} &= \text{(Quantity of solid waste generated)/(Production units)}, \hspace{1cm} (3b) \\
\text{E}_{\text{eff}}\text{IPU} &= \text{(Quantity of effluent generated)/(Production units)}. \hspace{1cm} (3c)
\end{align*}
The next level of elements of waste generation includes hazardous and non-hazardous waste. They are further classified into waste that is sent for recovery and waste that is sent for disposal. GHGs are hazardous and are invariably released to the atmosphere. Therefore, GHGs are captured under hazardous waste and do not include the next level of recovery and/or disposal. Once the practice of carbon capture is well established, these levels may be added to the downstream equations of GHG:

\[
GHG = Haz \, GHG, \quad (3a.1)
\]

\[
E_{fl} = Haz \, E_{fl} + Nhaz \, E_{fl}, \quad (3b.1)
\]

\[
Haz \, E_{fl} = Haz \, E_{fl} \text{ for recovery} + Haz \, E_{fl} \text{ for disposal}, \quad (3b.1.1)
\]

\[
Nhaz \, E_{fl} = Nhaz \, E_{fl} \text{ for recovery} + Nhaz \, E_{fl} \text{ for disposal}, \quad (3b.1.2)
\]

\[
SW = Haz \, SW + Nhaz \, SW, \quad (3c.1)
\]

\[
Haz \, SW = Haz \, SW \text{ for recovery} + Haz \, SW \text{ for disposal}, \quad (3c.1.1)
\]

\[
Nhaz \, SW = Nhaz \, SW \text{ for recovery} + Nhaz \, SW \text{ for disposal}. \quad (3c.1.2)
\]

To encourage “circularity”, each element of “resource consumption” is allocated a different “incentive” or a “discount” or a “multiplying” factor called “circularity” factor. For example, in the case of energy use, manufacturers need greater incentive to the use of renewable energy over recovered energy than over non-renewable energy. Similarly, in the case of materials, there must be more incentive for use of renewable materials over recovered materials over non-renewable materials. Within the recovered materials category, “reused” materials are considered more resource efficient than “reprocessed” (recycled, upcycled, down-cycled) materials. Same logic is applied for the use of fresh and recovered water. Since there is no precedence or research in the use of such “circularity factor”, the use of ratios of Fibonacci numbers starting with 1 for deriving the “circularity factor” has been proposed. These factors are used in the detailed equations of the proposed Gate2Gate OREft indicator to encourage circularity/environmental sustainability ($\alpha = 1/1$, $\beta = 1/2$, $\gamma = 1/3$, $\lambda = 1/5$).

“Circularity factors” are based on the hierarchy between different recovery options as proposed in the “Hierarchy of Resource Use” by Gharfalkar et al. (2015). Reuse could take place via repair and reuse, recondition and reuse, refurbish and reuse, remanufacture and reuse or any other operation and reuse. Reprocessing could include either recycling, upcycling or down cycling. Further, the hierarchy between various reuse options is based on the “Hierarchy of Reuse Options” as proposed by Gharfalkar et al. (2016).

Circularity factors for energy use. The circularity factors for energy are based on the hierarchy of energy use where renewable energy is at the top, recovered energy at the middle and non-renewable energy at the bottom of the hierarchy:

- Circularity factor for non-renewable energy: $\alpha = 1/1 = 1 \rightarrow$ no discount as it does not support environmental sustainability.
- Circularity factor for recovered energy: $\beta = 1/2 = 0.5 \rightarrow$ medium discount for encouraging circularity.
- Circularity factor for renewable energy: $\gamma = 1/3 = 0.33 \rightarrow$ maximum discount for supporting environmental sustainability.
Circularity factor for material use. The circularity factors for material use are based on the hierarchy of material use where renewable materials are at the top, followed by reused materials (repaired/reconditioned/refurbished/remanufactured), reprocessed materials (recycled/upcycled/down-cycled) and non-renewable materials at the bottom of the hierarchy in the same order:

- Circularity factor for non-renewable materials: $\alpha = 1 \rightarrow$ no discount as it does not support environmental sustainability.
- Circularity factor for reprocessed materials: $\beta = 1/2 = 0.50 \rightarrow$ it is less resource efficient than reused.
- Circularity factor for reused materials: $\gamma = 1/3 = 0.33 \rightarrow$ more resource efficient than reprocessed.
- Circularity factor for renewable materials: $\lambda = 1/5 = 0.20 \rightarrow$ maximum discount for supporting environmental sustainability.

Circularity factors for water use. The circularity factors for water use are based on the hierarchy of water use where recovered water is at the top and fresh water the bottom of the hierarchy:

- Circularity factor for fresh water: $\alpha = 1 \rightarrow$ no discount.
- Circularity factor for recovered water: $\beta = 1/2 = 0.5 \rightarrow$ maximum discount.

With above inputs of circularity factors, the equations for energy, material and water consumptions are as mentioned as follows:

**Energy consumption**

$$\text{Energy consumption} = \alpha(\text{Non-renewable}) + \beta(\text{Recovered}) + \gamma(\text{Renewable})$$

$$= (\text{Non-renewable}) + 0.5(\text{Recovered}) + 0.33(\text{Renewable}),$$

**Material consumption**

$$\text{Material consumption} = \alpha(\text{Non-renewable}) + \beta(\text{Reprocessed}) + \gamma(\text{Reused})$$

$$+ \lambda(\text{Renewable}) = (\text{Non-renewable}) + 0.5(\text{Reprocessed})$$

$$+ 0.33(\text{Reused}) + 0.20(\text{Renewable}),$$

**Water consumption**

$$\text{Water consumption} = \alpha(\text{Fresh}) + \beta(\text{Recovered}) = (\text{Fresh}) + 0.5(\text{Recovered}).$$

Circularity factors for waste generation. The circularity factors for waste generation are based on the hierarchy of waste where non-hazardous waste for recovery is at the top, followed by non-hazardous waste for disposal, hazardous waste for disposal and hazardous waste for recovery at the bottom of the hierarchy in the same order. These circularity factors are used for differentiating between hazardous and non-hazardous waste at the primary level as well as waste going for recovery and waste going for disposal at the secondary level.

**Circularity factors for hazardous waste:**

- For disposal: $\alpha = 1 \rightarrow$ No discount/incentive.
- For recovery: $\beta = 1/2 = 0.50.$

**Circularity factors for non-hazardous waste:**

- For disposal: $\gamma = 1/3 = 0.33.$
- For recovery: $\lambda = 1/5 = 0.20 \rightarrow$ maximum discount/incentive.
The final equation for the Gate2Gate ORE\textsubscript{at} after consideration of circularity factor is as mentioned in the following equations:

\[
\text{Gate2Gate ORE}_{at} = \text{RIPU after circularity} \times \text{WIPU after circularity},
\]

\[
\text{RIPU after circularity} = \text{EIPU after circularity} + \text{MIPU after circularity} + \text{WtrIPU after circularity},
\]

\[
\text{WIPU after circularity} = \text{GHGIPU after circularity} + \text{EfflIPU after circularity} + \text{SWIPU after circularity},
\]

\[
\text{Gate2Gate ORE}_{at} = ((\text{Non - renewable energy}) + 0.5(\text{Recovered energy})
+ 0.33(\text{Renewable energy}) + (\text{Non - renewable material})
+ 0.5(\text{Reprocessed material}) + 0.33(\text{Reused material})
+ 0.20(\text{Renewable material}) + (\text{Fresh water})
+ 0.5(\text{Recovered water}) \times ((\text{GHG}) + (\text{Haz E}_{at} \text{ for disposal})
+ 0.5(\text{Haz E}_{at} \text{ for recovery}) + 0.33(\text{Nhaz E}_{at} \text{ for disposal})
+ 0.2(\text{Nhaz E}_{at} \text{ for recovery}) + (\text{Haz SW for disposal})
+ 0.5(\text{Haz SW for recovery}) + 0.33(\text{Nhaz SW for disposal})
+ 0.2(\text{Nhaz SW for recovery})).
\]

Finally, the Gate2Gate ORE\textsubscript{at} index of individual factory is derived as per Equation (6). This equation ensures that the Gate2Gate ORE\textsubscript{at} index can be measured on a scale of 0–1. It is assumed that higher the Gate2Gate ORE\textsubscript{at} index score, better the manufacturing unit in terms of its resource effectiveness:

\[
\text{Gate2Gate ORE}_{at} \text{Index} = 1/(\text{Gate2Gate ORE}_{at}).
\]

### 3.3 Testing stage

Main objective of this stage is to test the validity of the proposed Gate2Gate ORE\textsubscript{at} indicator. There are various definitions of validation. Oxford Dictionary defines “to validate” as “to check or prove the validity or accuracy of”. Cambridge dictionary defines it as “to make something officially acceptable or approved, especially after examining it” and/or “to prove that something is correct”. Kirchner \textit{et al.} (1996) defines ‘validity’ as the “adequacy for specific purpose”. Bockstaller and Girardin (2003) consider an indicator to be validated “if it is scientifically designed, if the information provided by it is relevant and if it is useful and used by the end users”.

General framework and methods for the validation of indicators at conceptual and output stage as suggested by Bockstaller and Girardin (2003) are summarised in Table V. They clarify that the design or conceptual validation is important when the possibility of no other validation exist. Therefore, it is not necessary for a new indicator to be subjected to all types of validation:

1. Conceptual validation: to assess whether the indicator is scientifically founded.
2. Output validation: to assess the soundness of the outputs of the indicator.
Following validation methods are adopted at the conceptual and output stage of the proposed Gate2Gate ORE\textsubscript{Ei} indicator:

1. Conceptual validation: review by experts’ method is used for validating the concept of the proposed indicator. Manufactures are considered as the experts in this case and a web-based “industry survey” is used to seek their inputs.

2. Output validation: although 6 of the 86 manufacturers surveyed agreed to share data for testing of the indicator, only two shared their data: rubber products manufacturing unit and cast-iron foundry unit. The indicator is validated by establishing its relationship with consumption of resources per unit of production (RIPU) and generation of waste per unit of production (WIPU). The indicator is also validated by comparing the index with four other resource REMIs. It could not be compared with more REMIs due to lack of data as required for calculating other REMIs. Gate2Gate ORE\textsubscript{Ei} index of the two manufacturing units is also compared with each other to understand if and why one manufacturing unit is more resource effective than the other.

4. Results/findings

4.1 Findings of the analysis of 40 REMIs

Outcome of the analysis of 40 REMIs using a set of quantitative and qualitative criteria is graphically depicted in Figures 4–6. The graph in Figure 4 captures mean normalised scores of each of the 40 REMIs. These are further grouped into different blocks in two matrices as in Figures 5 and 6. Proposed Gate2Gate ORE\textsubscript{Ei} indicator is also scored using the same set of criteria and plotted on the graph and the two matrices. In the “score vs complexity matrix” (Figure 5), the Y-axis is grouped into three levels of scores: low score of 0 to less than 1, medium score between 1 to less than 2 and high score between 2 and 3. In the “data availability vs complexity matrix” (Figure 6), the Y-axis is grouped into three levels of data availability: low, medium and high. “Low” indicates that a REMI is based on 100 per cent assumptions; “medium” indicates that it is based on a combination of operational data and assumptions, while “high” indicates that it is based on 100 per cent operational data. For both the matrices, the X-axis is grouped into three levels of complexity: low, medium and high.

The analysis confirms that 75 per cent of REMIs score below 1 against the mean normalised maximum possible score of 3 and the remaining 25 per cent score between 1 to less than 2. Although, only 2 out of the 40 REMIs (M5 and M6) cross the half way mean normalised score of 1.5 (Figure 4), both, M5 and M6 are complex and not based on 100 per cent operational data (Figure 5). Thirteen of the 40 REMIs (M2, M3, M12A, M12B, M14, M15, M16, M20, M21, M23, M24, M27 and M28) are low on complexity and high on data availability, but none of them simultaneously capture resource use and waste generation in their measurement (Figures 5 and 6 and Table VI).

<table>
<thead>
<tr>
<th>Type of validation</th>
<th>Question</th>
<th>Methods of validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual validation</td>
<td>Is it scientifically founded?</td>
<td>Peer review</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Review by experts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comparison of approaches</td>
</tr>
<tr>
<td>Output validation</td>
<td>Is it realistic or does it inform about the reality?</td>
<td>Validation through comparison with a set of measured data.</td>
</tr>
<tr>
<td>(Empirical validation)</td>
<td></td>
<td>Global expert validation</td>
</tr>
</tbody>
</table>

Source: Bockstaller and Girardin (2003)
The three REMIs (M5, M6 and M15) that simultaneously capture both, resource consumption and waste generation in its measurements (Table VI) are high on complexity and not 100 per cent based on easily available operational data. For example:

1. **M5 (ecological footprint – compound)** relates to a country’s use of resources to its land base. It involves estimation of net average per capita consumption of about 50 biotic resources, estimation of per capita land appropriated to produce each good or service and estimation of average annual per capita energy consumption for over 100 categories of traded goods. This is further converted to the amount of forested land necessary to sequester the emitted CO2. Finally, the total ecological footprint is estimated by adding
all the appropriated land areas. This is very complex, data intensive and low on data availability. Also, M5A is suitable for assessing RE only at national level.

(2) M6 (ecological footprint – component) is suitable for calculating footprint values for individual activities or components at local and personal level. But it is also complex, data intensive and low on data availability for all the 24 components that it takes into consideration in its measurement.

(3) M15 (EMC) involves combining data from economy-wide material flow accounts such as direct material consumption (DMC) with data from life cycle analysis (LCA) by multiplying the mass of selected base materials with the LCA impact coefficients. In total, 13 different impact categories of LCA are aggregated into one score by weighting. M13 is complex and not good on data availability.

(4) Although M13, M14, M16, M17, M18, M22, M23 and M30 are high on data availability and low on complexity, all of them are low on score.

(5) Finally, while M24, M35, M36 and M37 are high on data availability and medium on complexity, none of them capture both, the resource use as well as waste generation in its measurement.

| M1, M2, M3, M7, M9, M12, M13, M18, M19, M20, M22, M23, M24, M25, M26, M28, M29, M30, M31, M32, M33, M34, M35, M36, M37 | M4, M8, M11, M14, M16, M17, M27, M38, M39, M40 | M5, M6, M15 Proposed Gate2Gate OREfi indicator |

Table VI. REMI grouping based on the aspects it captures in its measurement
None of the 40 REMIs provide incentives to encourage circularity in recovery, reprocessing or reuse of waste resources. With this major gap identified in the analysed REMIs, incentive/multiplying factor defined as circularity factors are used in the development of the proposed indicator.

Analysis of 40 REMIs confirms that a REMI as per the hypothesis statement does not exist.

4.2 Results of the industry survey

Key findings summarised here relate to the 86 responses by manufacturing, engineering and processing businesses in north England. These respondents are hereafter being called as “manufacturers”. Statistically, 86 responses represent the overall population of manufacturers in England at 94.1 per cent expected incidence rate with ±5 per cent error and 95% confidence level. The use of 90% confidence levels with a margin of error of ±5 per cent is considered reasonable for most audits/surveys (Bristol, 2015). Calculations of whether 86 responses represent the overall population of manufacturers in England are based on the sample size calculation mentioned as follows (Bristol, 2015):

\[ n = \left[ c^2 \times N \times p \times (1-p) \right]/\left( A^2 \times N \right) + \left( c^2 \times p \times (1-p) \right) \]

where \( n \) is the sample size required; \( N \) the whole target population in question; \( p \) the average proportion of records expected to meet the various criteria; \( (1-p) \) the average proportion of records not expected to meet the criteria; \( A \) the margin of error deemed to be acceptable (e.g. for 5 per cent error either way, \( A = 0.05 \)); \( c \) a mathematical constant defined by the confidence interval chosen (how sure we need to be of the result):

- To be 95 per cent sure of the result the constant \( c = 1.96 \).
- To be 90 per cent sure of the result the constant \( c = 1.645 \).
- To be 80 per cent sure of the result the constant \( c = 1.28 \).

Four key findings of the industry survey are summarised as follows:

1. 78 per cent of manufacturers surveyed in north England agree that a good “resource effectiveness” indicator should include both, consumption of key natural resources and waste generation in its measurement (Figure 7).

2. 54 per cent of manufacturers surveyed in north England either strongly agree (16 per cent) or agree (38 per cent) that a good “resource effectiveness” indicator should be based on 100 per cent operational data (Figure 8).

3. Both the above findings substantiate the hypothesis statement.

4. Considering current availability of data, 51 per cent of manufacturers surveyed in north England recommend a system boundary of Factory Gate2Gate for the new indicator. Only 6 per cent recommend a system boundary of Cradle2Gate for each product and 8 per cent recommend a system boundary of Cradle2Grave for each product (Figure 9). This input is used to define the system boundary of the proposed resource effectiveness indicator as Factory Gate2Gate.

4.3 Results of case studies

4.3.1 Gate2Gate \( \text{ORE}_{it} \) index vs RIPU and WIPU. Correlation between the resource intensity per unit (RIPU), waste intensity per unit (WIPU) and Gate2Gate \( \text{ORE}_{it} \) index of the rubber products manufacturing and foundry unit is analysed. For both the units, it is
observed that the Gate2Gate ORE\textsubscript{t} index is inversely proportional to RIPU as well as inversely proportional to WIPU of that manufacturing unit (Figures 10–13).

4.3.2 Comparison of Gate2Gate ORE\textsubscript{t} index of two manufacturing units. Although it is not advisable to compare RE of two diverse manufacturing units such as a rubber product manufacturing and a foundry unit, a comparison of their Gate2Gate ORE\textsubscript{t} indices is carried out for academic purpose. It is assumed that the two units manufacture similar products. For a manufacturing unit to be resource efficient/productive than the other, it is necessary to have lower values of resource intensity per unit (RIPU) and/or waste intensity per unit (WIPU) \textit{vis-à-vis} the other unit. RIPU is consumption of resources per unit of production and WIPU is generation of waste per unit of production. With this logic,
Figure 9. Preferred system boundary for a good resource effectiveness indicator.

Figure 10. Rubber unit: resource intensity per unit (RIPU) vs Gate2Gate ORE\textsubscript{E} Index.

Figure 11. Rubber unit: waste intensity per unit (WIPU) vs Gate2Gate ORE\textsubscript{E} Index.
the comparison of Gate2Gate ORE_{gt} index for these two units during 2013, 2014 and 2015 indicate that the first unit (in this case the rubber products manufacturing unit) is more resource efficient/productive than the second unit (foundry) in each year (Figure 14). Reasons why the first unit has a better Gate2Gate ORE_{gt} index and therefore could be considered more resource efficient/productive than the second unit are mentioned below:

- Three years’ average consumption of resources per unit of production (RIPU) of the first unit is 6.00, which is 17 per cent lower than that of the second unit whose average RIPU is 7.19 (Table VII).

- Three years’ average generation of wastes per unit of production (WIPU) of the first unit is 1.77, which is 22 per cent lower than that of the WIPU of the second unit, which is 2.27 (Table VII).

- This means on an average, the first unit consume less resources per unit of production and generates lower waste per unit of production as compared to the second unit. Therefore, it may be inferred that the first unit is more resource efficient/productive than the second unit.

4.3.3 Comparison of Gate2Gate ORE_{gt} index with other REMIs. For validation/testing purpose, Gate2Gate ORE_{gt} indices of the two manufacturing units are also compared with four REMIs described as follows:

(1) Resource productivity (M12A): calculated as the monetary output per unit of all resources aggregated together. Since the two manufacturing units did not share monetary data,
monetary output is replaced with tons of production output. This is divided by the aggregate of energy, material and water resources converted into equivalent tons.

(2) Material productivity (M20): calculated as monetary output per unit of direct material consumed. Monetary output is replaced with tons of production output.

(3) Total material consumption (M21): it measures the total amount of materials directly used by a nation or a company or a business unit.

(4) Water productivity (M23): calculated as monetary output per unit of fresh water consumed. Monetary output was replaced with tons of production output.

Values of the Gate2Gate ORE\textsubscript{ft} index and the four REMIs for the foundry unit are summarised in Table VIII. Comparison of the Gate2Gate ORE\textsubscript{ft} index with resource productivity (Figure 15), material productivity (Figure 16) and water productivity (Figure 17) shows similar trend between the compared indicators for 2012 to 2015. This is in line with the expectation that lower the resource/material/water productivity, lower the RE and vice versa. Comparison of the Gate2Gate ORE\textsubscript{ft} index with total material consumption shows an opposite trend (Figure 18). This is also in line with the expectation that lower the material consumption, higher the RE. Similar trends are observed for the plastic unit.

<table>
<thead>
<tr>
<th>Year</th>
<th>RIPU Rubber</th>
<th>WIPU Rubber</th>
<th>Gate2Gate ORE\textsubscript{ft} Index</th>
<th>Code</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>6.55</td>
<td>7.21</td>
<td>2.34</td>
<td>ORE\textsubscript{ft}</td>
<td>0.057</td>
<td>0.060</td>
<td>0.072</td>
<td>0.065</td>
</tr>
<tr>
<td>2014</td>
<td>6.17</td>
<td>6.71</td>
<td>1.61</td>
<td>M12A</td>
<td>0.110</td>
<td>0.113</td>
<td>0.123</td>
<td>0.107</td>
</tr>
<tr>
<td>2015</td>
<td>5.28</td>
<td>7.64</td>
<td>1.36</td>
<td>M20</td>
<td>0.136</td>
<td>0.140</td>
<td>0.146</td>
<td>0.132</td>
</tr>
<tr>
<td>Average</td>
<td>6.00</td>
<td>7.19</td>
<td>1.77</td>
<td>M21</td>
<td>2015</td>
<td>2106</td>
<td>2516</td>
<td>2001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate2Gate ORE\textsubscript{ft} Index</td>
<td>0.057</td>
<td>0.060</td>
<td>0.072</td>
<td>0.065</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource productivity</td>
<td>0.110</td>
<td>0.113</td>
<td>0.123</td>
<td>0.107</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material productivity</td>
<td>0.136</td>
<td>0.140</td>
<td>0.146</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total material consumption</td>
<td>2015</td>
<td>2106</td>
<td>2516</td>
<td>2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water productivity</td>
<td>0.688</td>
<td>0.724</td>
<td>0.953</td>
<td>0.673</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Measuring RE

1877

Figure 15. Foundry unit: Gate2Gate OREₘₖ index vs resource productivity M12A

Figure 16. Foundry unit: Gate2Gate OREₘₖ index vs material productivity M20

Figure 17. Foundry unit: Gate2Gate OREₘₖ index vs water productivity M23
5. Conclusions
To summarise, none of the 40 analysed REMIs that were identified through the literature survey, capture both, the “resource use”, and “waste generation” using 100 per cent operational data in its measurement. Also, none of these REMIs provide incentives to encourage circularity in recovery, reprocessing or reuse of waste. 78 per cent of surveyed manufacturers agreed that a good “resource effectiveness” indicator should include both, consumption of key natural resources and waste generation in its measurement. Also, 54 per cent of the manufacturers agree that a good “resource effectiveness” indicator should be based on operational data. Both these responses clearly validate the hypothesis. Finally, the inverse correlation established between the Gate2Gate ORE\textsubscript{it} index and the RIPU and the WIPU of a foundry and plastic products manufacturing unit, comparison of the Gate2Gate ORE\textsubscript{it} indices of these two units with each other and with four existing REMIs, validates the output generated by the new ORE\textsubscript{it} indicator. The ORE\textsubscript{it} indicator is a “new indicator” of “operational resource effectiveness” suitable for manufacturing units. Unlike many REMIs, the new ORE\textsubscript{it} indicator is based on readily available operational data, not assumptions. In addition to the fact that the proposed indicator captures “resource consumption” and “waste generation” in its measurement, inclusion of “circularity factors” that capture the circularity of resource use and recovery and reuse of waste streams is the key distinguishing feature of this indicator. In terms of its practical implications, the proposed indicator can be used for comparing the operational resource effectiveness of individual factories over a period as well as with other manufacturing units. It also captures useful information such as resource intensity per unit and waste intensity per unit, which also reflect operational RE or resource productivity that can be used to initiate improvement action. Adoption of this indicator across manufacturing supply chain can lead to an overall improvement in the RE, resource productivity, as well as resource effectiveness across the supply chain.

As far as the limitations of this research and the Gate2Gate ORE\textsubscript{it} indicator are concerned, testing of this indicator was limited to two dissimilar manufacturing units that shared data. The validation could have been more effective if more units manufacturing similar products had shared their operational data. Also, lack of availability of data for any supply chain, restricted the testing of this indicator to Gate2Gate boundary of each manufacturing unit. For a unit manufacturing a variety of products within the same campus, unless product or process wise resource consumption and waste generation data is available, this indicator cannot evaluate the resource effectiveness of individual products or processes within that campus. All products or processes within a campus are allocated the
same resource effectiveness as that of the manufacturing unit. Also, the indicator considers consumption of only the primary raw materials. It does not differentiate between different raw materials as they are aggregated together by weight.

Suggestions for future investigations include, conducting an industry survey in other regions of the UK and/or Europe to create a database of Gate2Gate ORE\textsubscript{ft} indices of similar and dissimilar manufacturing units. Further investigations may also be carried out targeting specific industrial segments such as the foundry or the plastic injection moulding units. This may help in identifying units with high Gate2Gate ORE\textsubscript{ft} index, whose best practices could then be shared within the industry segment for overall improvement of that segment.

The linear system of “make-use-dispose” is not environmentally sustainable. To achieve real long-term environmental sustainability, evolution of “closed loop resource effective business models” is inevitable. These business models are likely to have renewable natural resources as inputs and outputs that are environmentally benign. This goal of environmental sustainability can be achieved if resource effectiveness is assessed at each stage of a product life cycle. Manufacturing is just one stage of this cycle. It may not be important how accurate or precise an indicator is but whether it gives some indication of resource effectiveness that could be used for initiating improvement actions. The proposed ORE\textsubscript{ft} indicator is a new indicator that could be used by manufacturers for achieving this objective.

**References**


Further reading


About the authors
Mangesh Gharfalkar is PhD Researcher at the School of Science and Engineering, Teesside University, Middlesbrough, UK. He has a Degree in Production Engineering, MBA in Marketing and Fellowship in Manufacturing Management. He began his career as a Lecturer of Production Engineering and worked in the manufacturing sector for more than two decades. His skills and expertise include: energy conservation, productivity improvement and prevention, recovery and reuse of waste. His current research focus is on resource efficiency/effectiveness in manufacturing. Mangesh Gharfalkar is the corresponding author and can be contacted at: mgharfalkar@yahoo.com

Zulfiqur Ali is Research Professor and Director of Academic Research and Innovation Partnerships at Teesside University, Middlesbrough, UK. His research interests include micro and nanofabrication approaches for creating high density functional devices. He is on the Editorial Board for Micro and Nano systems from Bentham Science Publishers. He has been an International Panel Member for the Higher Education Academy (HEA) of Ireland. He has coordinated a number of EU projects and is currently leading the Teesside University activity in the Horizon 2020 GateOne project on smart system integration for SMEs.

Graham Hillier is the Director of Strategy and Futures at the Centre for Process Innovation, Wilton Centre, UK. He has a Degree in Metallurgy, PhD Degree from the University of Cambridge and an MBA. He is a Chartered Engineer, a Fellow of the Institute of Materials, Minerals and Mining and a Fellow of the Royal Society of Arts, Manufactures and Commerce. He is also visiting Professor in the Department of the Built Environment, University of Salford. He has business background in recycling operations, new product development, business development and low carbon energy.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Cost of quality measurement in food manufacturing companies: the Greek case

Evangelos Psomas and Christina Dimitrantzou
Department of Business Administration of Food and Agricultural Enterprises, University of Patras, Patras, Greece

Fotis Vouzas
Department of Business Administration, University of Macedonia, Thessaloniki, Greece, and

Nancy Bouranta
Department of Business Administration of Food and Agricultural Enterprises, University of Patras, Patras, Greece

Abstract

Purpose – The purpose of this paper is to focus on the cost of quality (CoQ) of food manufacturing companies. The study aims at empirically validating the conceptual structure of the core dimensions of CoQ (prevention, appraisal, internal and external failure cost) and determining their level and relationships. Determining the reasons for not measuring the CoQ as well as the barriers-difficulties and benefits of the CoQ measurement is also an aim of the present study.

Design/methodology/approach – Greek food manufacturing companies were approached through a structured questionnaire and 91 participated in the study. Exploratory and confirmatory factor analysis, descriptive statistics and correlation analysis are applied for data analysis.

Findings – The structure of the core dimensions of CoQ is empirically validated, while all of them range within medium levels and are inter-related. The reasons for companies not measuring the CoQ as well as the barriers-difficulties of companies measuring the CoQ are not significant. On the contrary, significant benefits are derived from the CoQ measurement.

Research limitations/implications – The small sample of the food manufacturing companies operating in one country and the subjective business evidence collected are the main limitations of the present study.

Practical implications – Food companies can be motivated to establish a robust CoQ measurement system, which would reflect the level of the CoQ dimensions. According to the results of the CoQ measurement, a food company can make decisions to increase the prevention and appraisal costs and decrease the internal- and external-failure costs in order to be more competitive in the market.

Originality/value – The paper provides deeper insights into the level and inter-relations of empirically validated CoQ dimensions. Very few studies on the CoQ in food manufacturing companies have been carried out in Greece and the present study makes the picture of the CoQ status in this sector more clear.

Keywords Greece, Cost of quality, Food manufacturing companies, PAF model

Paper type Research paper

1. Introduction

In the modern global market, business organizations are confronted with highly competitive and complex circumstances (Wudhikarn et al., 2015), due to the internationalization of transactions, the rapid growth of technology and the changing socio-political and economic environment (Mahmood et al., 2014). Such a dynamic trading environment provides many potential opportunities for companies but also hidden risks (Arvaiova et al., 2009). As far as Greek companies are concerned, a complicated tax system, coupled with limited cash liquidity and a decrease in consumer spending power constitute a hard business environment in which companies have to survive and develop (Chatzipetrou and Moschidis, 2016).

In such a business environment, consumers demand lower priced and yet superior quality products while companies concentrate on the identification of their cost drivers and
increasing their market share and profit (Stapleton et al., 2004; Wudhikarn et al., 2015). In other words, companies have to improve the quality of their products and services and simultaneously focus on the cost factor because not only do they have to satisfy the customers’ needs and requirements with regard to quality but they also have to provide the products/services with the minimum cost (Snieska et al., 2013; Prashar, 2014; Malik et al., 2016). The measurement and evaluation of all the quality costs seems to be a critical issue for a company’s improvement in all business areas, since it plays an important role in increasing efficiency, reducing unfruitful expenditure and enhancing profitability (Sahu and Sridhar, 2013; Satanova et al., 2015). Thus, in order for companies to obtain a competitive advantage, they should evaluate the Cost of Quality (CoQ) and more specifically its most commonly used dimensions including the prevention, appraisal, internal and external failure cost (PAF model) (Snieska et al., 2013).

Although the concept of CoQ is not recent, very few studies have been conducted on the measurement of CoQ so far (Mahmood and Kureshi, 2015). Bhatti et al. (2014) state that though there is a reasonable amount of detailed advice available on the CoQ, there are only a few published practical examples that give insights into the costs that should be included or excluded in quality costing, and how these costs are collected (Bhatti et al., 2014). Similarly, Snieska et al. (2013), note that the scientific literature still has questions on how to calculate all the constituents of CoQ dimensions in detail. Moreover, Pires et al. (2017) mention that there is limited research that describes the way CoQ information is used in the management process. Chatzipetrou and Moschidis (2017) support the view that not much attention has been paid by researchers to the depiction and analysis of CoQ practices in Greek manufacturing and especially in food and beverages (F&B) enterprises.

Sailaja et al. (2014) proposed to extend the research in the field of CoQ by incorporating in the CoQ measurement the hidden elements of CoQ, given that their study focuses only on direct and visible costs which are easy to track and measure. Further studies are also needed to find out how companies make use of the CoQ information to formulate a survival strategy at a time of economic downturn and a growth strategy for their future expansion (Cheah et al., 2011). According to Pires et al. (2017), how exactly CoQ information is used by managers in their day-to-day tasks, as well as the reasons for the use of the CoQ information by managers are questions that should be thoroughly investigated. Lim et al. (2015) suggest the investigation of quality cost models for more complex situations where failures are not necessarily independent. Chatzipetrou and Moschidis (2017) proposed further research among F&B companies which could shed light on the deeper causes that hinder the implementation of quality costing mainly in relation to the economic restrictions on the Greek economy since summer 2015.

In order to fill the above-mentioned research gap and contribute to the literature, this study focuses on the CoQ in food manufacturing companies. Based on a sample of Greek food manufacturing companies operating in a business environment where an economic downturn and financial crisis dominates, the study aims to empirically validate the structure of the core dimensions of CoQ (prevention, appraisal, internal and external failure cost) and determine their level and relationships. The study also aims at determining the reasons for companies not measuring the CoQ, and the barriers-difficulties and benefits of companies measuring the CoQ. To the best of the authors’ knowledge, the only recent research studies which focus on the CoQ in the Greek food sector are the studies of Chatzipetrou and Moschidis (2016), Chatzipetrou and Moschidis (2017) and Moschidis et al. (2018). More specifically, Chatzipetrou and Moschidis (2016) focus on food retail companies meaning Greek supermarkets and examine, through multiple correspondence analysis, the use of the PAF model and its content and how these are diversified according to a number of demographic variables (e.g. size, Hazard Analysis of Critical Control Points...
These authors extend their previous work on supermarkets to the food manufacturing sector (Chatzipetrou and Moschidis, 2017). Additionally, Moschidis et al. (2018) examine through multiple correspondence analysis, the relationship between the quality costing system as defined by the PAF approach and the quality management maturity level of food manufacturing companies. The present study takes a different direction. It examines the reasons for not measuring CoQ, the barriers-difficulties of CoQ measurement as well as its possible benefits. In addition, the current empirical research enriches the CoQ instrument, using a larger variety of CoQ elements, compared to the previous studies in the Greek food sector and validates its structure using exploratory and confirmatory factor analysis. Thus, both academics and practitioners can benefit from the present study, given that it expands the knowledge already gained in terms of CoQ in the food manufacturing sector.

The rest of the paper is structured as follows: in the first part, the existing CoQ literature is reviewed focusing on the definition of CoQ, the PAF model, the reasons for not measuring the CoQ and the barriers-difficulties and benefits of the CoQ measurement. The research questions of the study are then formulated. In the next part of the paper, the methodology of a research study carried out in Greek food companies is described. This is followed by the presentation of the study findings and their discussion. In the following part, the final conclusions and practical implications are presented. Finally, the paper ends with the limitations of the study and future research recommendations.

2. Literature review

2.1 Cost of Quality (CoQ)

The inception of the concept of CoQ can be attributed to Juran and Feigenbaum in the 1950s (Sailaja et al., 2014). However, there is no general agreement on a single broad definition of CoQ, and hence, different definitions exist in different industries and among researchers (Plunkett and Dale, 1987; Juran and Godfrey, 2000; Schiffauerova and Thomson, 2006b; Malik et al., 2016). A large number of different terms have been used in order to analyze the term “cost of quality,” such as “quality cost,” “economics of quality,” “poor quality cost,” “price of non-conformance” (PoNC) or “cost of poor quality” (Kiani et al., 2009). The generally accepted interpretation of CoQ includes the cost of all efforts made by a company in order to provide a product that can meet the necessary requirements and the customers’ needs (Chatzipetrou and Moschidis, 2016). According to Pires et al. (2017), CoQ is commonly understood as the monetary expression of the efforts undertaken by the organization to ensure the intended levels of quality.

One of the most important concerns of organizations is how to achieve a balance between the desired level of quality and the expenses associated with it (Heravi and Jafari, 2014). There was a mistaken notion that the achievement of better quality requires higher costs because better quality would somehow cost more and make production difficult (Chopra and Garg, 2011). This was the myth that prevented many companies from investing more in CoQ related programs (Sahu and Sridhar, 2013). An optimal quality improvement policy provides a trade-off between the minimization of the CoQ and the maximization of the quality of conformance, in order to achieve high quality in the most economic way (Sailaja et al., 2014; Duarte et al., 2016).

2.2 Prevention-appraisal-failure (PAF) model

One of the most widespread models of quality costing is the PAF model suggested by Juran in 1951 and Feigenbaum in 1956, classifying CoQ into prevention, appraisal and failure cost (Plunkett and Dale, 1987; Juran and Godfrey, 2000; Schiffauerova and Thomson, 2006b, Chatzipetrou and Moschidis, 2016). Failure cost is further classified into two sub-dimensions, namely, internal failure and external failure cost. In general, the four
categories of quality costs, which are emphasized by many companies, are described as follows (Bendell et al., 1995; Goulden and Rawlins, 1997; Juran and Godfrey, 2000; Roden and Dale, 2001; Superville and Gupta, 2001; Dale and Wan, 2002; Chopra and Garg, 2011; Chiariini, 2015; Guinot et al., 2016; Farooq et al., 2017; Moschidis et al., 2018):

- Prevention costs: these costs are associated with the design, implementation and maintenance of the quality management system. Prevention costs are planned and are incurred before actual operation. Some examples of prevention costs are quality planning and certification, supplier capability surveys, process capability evaluations, quality education and training, etc.

- Appraisal costs: these costs are associated with the evaluation of purchased materials, processes, intermediate and final products and services to assure conformance with the specified requirements. Some examples of appraisal costs are the costs of incoming, in-process and final inspections/tests, product and process audits, calibration of measuring and test equipment, etc.

- Internal failure costs: these costs occur when the results of work fail to reach quality standards, while these failures are detected before transfer to the customer takes place. Some examples of internal failure costs are scrap, rework, re-inspection, retesting, machine stoppages, etc.

- External failure costs: these costs occur when products or services fail to reach quality standards, while these failures are not detected until after the products or services are transferred to the customer. Some examples of external failure costs are customer returns, warranty claims, product recalls, replacement, compensation payments to the customer, etc.

An alternative to the PAF approach was introduced by Philip Crosby in 1979 and is known as Crosby’s model (Crosby, 1979; Chatzipetrou and Moschidis, 2016). According to Crosby (1979), quality means “conformance to requirements” and non-conformance refers to failure to do things right the first time. The price of conformance is the cost involved in making certain that things are done right the first time, which includes actual prevention and appraisal costs, while the PoNC is the money wasted when work fails to conform to customer requirements, usually calculated by quantifying the cost of correcting, reworking or scrapping, which corresponds to actual failure costs (Crosby, 1979; Schiffauerova and Thomson, 2006a; Vaxevanidis et al., 2009).

2.3 Reasons for not measuring the CoQ and barriers-difficulties of the CoQ measurement
A wide variety of reasons why companies do not measure CoQ and the barriers-difficulties encountered by those measuring CoQ have been suggested by researchers. Some of them are the following (Bamford and Land, 2006; Eldridge et al., 2006; Sower et al., 2007; Arvaiova et al., 2009; Cheah et al., 2011; Guinot et al., 2016; Pires et al., 2017):

- lack of understanding of the CoQ concept;
- lack of knowledge of how to track the CoQ;
- absence of management interest in tracking CoQ;
- difficulty in collection of quality data;
- financial difficulties of covering the required expenses for CoQ measurement;
- lack of adequate accounting and computer systems necessary to track CoQ;
- lack of cooperation with other departments;
- management philosophy and company culture which do not promote CoQ evaluation;
• management belief that there is no value in any efforts to fully measure CoQ;
• lack of experienced manpower to accomplish the task;
• management deems CoQ to be an extra workload;
• organizations do not see the benefits of CoQ; and
• organizations focus on areas which perceive to be more important.

2.4 Benefits of the CoQ measurement
Numerous research studies have been dedicated to the determination of the CoQ benefits. Reviewing the literature, it is apparent that there is a general consensus among authors regarding the benefits gained after CoQ measuring and reporting. So, according to the literature, the main benefits of implementing the CoQ system are the following (Stapleton et al., 2004; Miguel and Pontel, 2004; Schiffauerova and Thomson, 2006a; Bamford and Land, 2006; Uyar, 2008; Kim and Nakhai, 2008; Kiani et al., 2009; Walsh and Antony, 2009; Arvaiova et al., 2009; Chopra and Garg, 2011; Khataie and Bulgak, 2013; Sahu and Sridhar, 2013; Malik et al., 2016; Pires et al., 2017; Chatzipetrou and Moschidis, 2017):
• increase in profit;
• increase in sales volume;
• product/service quality improvement;
• increase in company competitiveness;
• achievement of significant cost reductions;
• improvement of productivity;
• increase in customer and employee satisfaction;
• decrease in customer complaints;
• elimination of all forms of waste;
• provision of vital information for management to be aware of the size of the problem;
• identification of high-cost problem areas;
• prioritization of improvement actions with the highest potential payoff;
• evaluation of the worth of individual quality activities;
• increase in the overall level of quality awareness; and
• explanation of the conflicting cost behaviors that historically have been destructive to practitioners and researchers.

2.5 Inter-relations between CoQ dimensions
Many researchers have sought to understand the correlation between different CoQ dimensions. A large body of current literature shares general assumptions according to which (Omachonu et al., 2004; Sower et al., 2007; Ramdeen et al., 2007; Kiani et al., 2009; Chopra and Garg, 2011; Sailaja et al., 2014; Omar and Murgan, 2014):
• there is a relationship between the prevention cost and the appraisal cost as well as between the internal failure cost and the external failure cost;
• there is an inverse relationship between the prevention cost plus the appraisal cost and the failure cost: in other words, a strong negative relationship exists between the
conformance cost (prevention and appraisal cost) and the non-conformance cost (internal and external failure cost);

- increasing the prevention and appraisal costs will decrease the total CoQ;
- the prevention cost has the most effect on total CoQ rather than the appraisal cost;
- there is a strong negative relationship between the prevention cost and the internal failure cost, and between the prevention cost and the external failure cost; and
- there is a strong negative relationship between the appraisal cost and the internal failure cost, and between the appraisal cost and the external failure cost.

2.6 Research questions
Bearing in mind the purpose of the present study, the literature gap identified, the research proposals suggested by many authors and the above review of the literature, the following research questions are formulated in order for them to be examined in Greek food manufacturing companies:

RQ1. Do food companies measure CoQ?

RQ2. Which are the main reasons for the food companies not measuring CoQ?

RQ3. Which are the main barriers-difficulties encountered during the CoQ measurement by food companies?

RQ4. Which are the main benefits derived from the CoQ measurement by food companies?

RQ5. Do the four core CoQ dimensions, namely, prevention cost, appraisal cost, internal failure cost and external failure cost, reflect the measured CoQ elements identified in the literature?

RQ6. Which are the levels of the dimensions of CoQ in food manufacturing companies?

RQ7. Which are the relationships between the dimensions of CoQ in food companies?

3. Research methodology
3.1 Questionnaire development
In order to answer the above research questions, a research project was carried out focusing exclusively on the Greek food sector. A postal survey was employed in order to investigate the CoQ in this sector. A structured questionnaire was used as the data-collection method, the design of which was based on a comprehensive literature review in the field of CoQ. A draft of the questionnaire was first reviewed by academics and professionals and then subjected to a pilot test by practitioners, proving its appropriateness and achieving the content validity of the questionnaire items.

The final version of the questionnaire consists of five parts. The first part includes questions on the demographic profile of a company. The second, third and fourth parts contain statements with regard to the reasons why companies do not measure CoQ, the barriers-difficulties encountered when companies measure CoQ and the benefits of the CoQ measurement, respectively. Finally, the fifth part of the questionnaire contains statements describing CoQ elements. The questionnaire items describing the reasons for companies not measuring CoQ and the barriers-difficulties companies face with the CoQ measurement were drawn from the studies of Bamford and Land (2006), Eldridge et al. (2006), Sower et al. (2007), Arvaiova et al. (2009), Cheah et al. (2011), Guinot et al. (2016) and Pires et al. (2017). The questionnaire items describing the benefits companies derive due to CoQ measurement were drawn from the studies of Stapleton et al. (2004), Miguel and Pontel (2004),
Schiffauerova and Thomson (2006a), Bamford and Land (2006), Uyar (2008), Kim and Nakhai (2008), Kiani et al. (2009), Walsh and Antony (2009), Arvaiova et al. (2009), Chopra and Garg (2011), Khataie and Bulgak (2013), Sahu and Sridhar (2013), Malik et al. (2016), Pires et al. (2017) and Chatzipetrou and Moschidis (2017). The questionnaire items describing the CoQ elements most commonly referred to in the literature were drawn from the studies of Roden and Dale (2001), Dale and Wan (2002), Schiffauerova and Thomson (2006a), Vaxevanidis et al. (2009), Chopra and Garg (2011), Chiari et al. (2015), Guinot et al. (2016) and Chatzipetrou and Moschidis (2016). Respondents were asked to indicate the degree of agreement or disagreement with these statements, using a seven-point Likert scale (from 1: strongly disagree to 7 strongly agree).

3.2 Sample

The Greek food manufacturing companies constitute the population of the organizations of interest of the present research study. In the period during which the present study was carried out, the database of ICAP (the largest business information and consulting firm in Greece) included 277 food manufacturing companies operating all over Greece, with full access to their contact information. The questionnaire was addressed to the CEOs of these companies, allowing them to decide who would be the best member of staff to answer the questionnaire. A total of 91 companies responded positively and returned the questionnaire completed in full, which represents a response rate of 32.8 percent.

In order to exclude the risk of non-response bias, the companies participating in the research were compared with those which refused to do so, in terms of their company size (based on the number of employees) (Mann–Whitney test). No differences were found among these two groups. Moreover, it is also worth noting that several non-responding companies stated, when contacted, that the major reason for them not participating in the research project was lack of time (Psomas et al., 2013). Thus, the possibility of facing a problem with non-response bias was rejected. Moreover, since the questionnaire was completed by a single respondent from each company, the common method variance was checked by applying the single-factor test (Martinez-Costa and Martinez-Lorente, 2008). This method produced poor results as indicated by the low percentage of the variance extracted (below 25 percent), confirming that the common method variance is not a substantive problem.

The majority of the food manufacturing companies participating in the present study are small- and medium-sized enterprises (SMEs) (92.3 percent), based on the number of employees. More specifically, according to the Commission Recommendation 2003/361/EC (2003), concerning the definition of SMEs, the responding companies can be categorized as follows: 29.7 percent as micro enterprises (< 10 employees), 34.1 percent as small enterprises (11–50 employees), 28.6 percent as medium enterprises (51–250 employees) and 7.7 percent as non-SMEs ( > 250 employees). All the sample food companies implement the HACCP, as they are obliged by law to do so (Chatzipetrou and Moschidis, 2017), however, a rate of 78 percent implement a Food Safety Management System (FSMS) certified according to the ISO 22000 standard. Moreover, it is worth noting that 44 percent of the sample companies are ISO 9001 certified, while 37.3 percent are both ISO 22000 and ISO 9001 certified.

3.3 Data analysis

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) was applied in order to validate the structure of the core dimensions of CoQ (Psomas et al., 2013; Psomas and Antony, 2015), while their level is evaluated through descriptive statistics (Tye et al., 2011; Arvaiova et al., 2009) and their relationships through correlation analysis (Omachonu et al., 2004). Descriptive statistics were also calculated to describe the companies’ profile, the reasons why companies do not track CoQ, the barriers-difficulties encountered when companies measure CoQ and the benefits derived from CoQ measurement (Tye et al., 2011). The statistical package SPSS is used for data processing.
4. Results

4.1 Descriptive statistics

A percentage of 29.7 percent of the sample companies stated that they do not measure and evaluate CoQ elements. Table I presents the major reasons why these companies do not track CoQ. From Table I, it is apparent that the first and most significant reason for the companies not measuring CoQ is the ignorance of the benefits of CoQ measurement. This is followed by the lack of knowledge and experience of how to track CoQ. The remaining and less significant reasons for not measuring CoQ, are the financial difficulties of covering the required expenses and the lack of interest of top management.

The most significant barriers-difficulties encountered by those companies which measure the CoQ (70.3 percent of the sample companies) are, in descending order according to their mean value, presented in Table II: the extra paperwork involved, the difficulty of staff in identifying sources or causes of CoQ, the difficulty in identifying quality cost elements and the difficulty in assigning CoQ elements to its dimensions. The remaining barriers-difficulties are less significant (the mean values range from 3.40 to 2.82). It is worth noting that the least significant difficulty is the refusal of the auditors to fill out records determining how the time is distributed during the audit.

Table III shows the most significant benefits from the CoQ measurement. The identification of areas that can be further improved is rated as the most important benefit to be gained. The rest of the benefits are only slightly less significant than the first one (Table III).

<table>
<thead>
<tr>
<th>Reason for not measuring CoQ</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ignorance of the benefits of CoQ measurement</td>
<td>4.30</td>
<td>1.81</td>
</tr>
<tr>
<td>Lack of knowledge and experience of how to track CoQ</td>
<td>4.11</td>
<td>1.75</td>
</tr>
<tr>
<td>Financial difficulties of covering the required expenses for CoQ measurement</td>
<td>2.90</td>
<td>1.78</td>
</tr>
<tr>
<td>Lack of interest of top management</td>
<td>2.41</td>
<td>1.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Barriers-difficulties in measuring CoQ</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extra paperwork</td>
<td>4.12</td>
<td>1.58</td>
</tr>
<tr>
<td>Difficulty of staff in identifying sources or causes of CoQ</td>
<td>4.03</td>
<td>1.42</td>
</tr>
<tr>
<td>Difficulty in identifying quality cost elements</td>
<td>3.86</td>
<td>1.38</td>
</tr>
<tr>
<td>Difficulty in assigning CoQ elements to the four dimensions (prevention, appraisal, internal failure and external failure)</td>
<td>3.65</td>
<td>1.31</td>
</tr>
<tr>
<td>Lack of information and responsibility from staff</td>
<td>3.40</td>
<td>1.65</td>
</tr>
<tr>
<td>Difficulty in costing the sources of CoQ elements</td>
<td>3.40</td>
<td>1.57</td>
</tr>
<tr>
<td>Difficulty in calculating the time spent by staff in respective processes</td>
<td>3.38</td>
<td>1.39</td>
</tr>
<tr>
<td>Refusal of the auditors to fill out records determining how the time is distributed during the audit</td>
<td>2.82</td>
<td>1.62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Benefits of measuring CoQ</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying areas for further improvement</td>
<td>5.68</td>
<td>0.96</td>
</tr>
<tr>
<td>Investments in prevention actions for achieving continuous improvement</td>
<td>5.55</td>
<td>1.07</td>
</tr>
<tr>
<td>Economic determination of the company’s profit due to improved quality</td>
<td>5.45</td>
<td>1.03</td>
</tr>
<tr>
<td>Determining quality problems and how to work more efficiently</td>
<td>5.39</td>
<td>1.09</td>
</tr>
<tr>
<td>Economic determination of the company losses due to poor quality</td>
<td>5.29</td>
<td>1.08</td>
</tr>
<tr>
<td>CoQ determination constitutes a good marketing tool</td>
<td>5.17</td>
<td>1.57</td>
</tr>
</tbody>
</table>
4.2 Exploratory factor analysis (EFA)

The number of the responding companies in the present study is deemed, according to Hair et al. (2005), large enough for multivariate data analysis. So, the CoQ elements identified in the literature were used as measured variables of an EFA. The result is the establishment of four latent factors (Kaiser–Meyer–Olkin = 0.840, Bartlett’s test of Sphericity = 828.76, \( p = 0.00 \), eigenvalue > 1, measures of sampling adequacy > 0.75, factor loadings > 0.608 except for one item with a loading equal to 0.44, cumulative variance = 76.28 percent). According to Hair et al. (2005), factor loadings from ±0.30 to ±0.40 are minimally accepted, while a minimum of three items should represent each factor. So, based on these criteria, few low loading items (< 0.4) were excluded from the data analysis. The extracted latent factors are explained using the measured variable loadings and can be labeled in accordance with the core dimensions of CoQ namely, prevention cost, appraisal cost, internal failure cost and external failure cost. The mean value of each CoQ dimension is calculated based on the mean values of the respective items (Table IV). The prevention cost is the most highly evaluated CoQ dimension, followed by the appraisal cost. The internal failure cost is rated as the third highest dimension, while the least important dimension is the external failure cost.

4.3 Confirmatory factor analysis (CFA)

No case was excluded from the CFA, since all the ratios of Mahalanobis \( D^2 \)/independent variables were lower than 3, which indicate that no outliers exist (Hair et al., 2005). The univariate statistics of the variables used, including the skewness (from −0.585 to −0.01), kurtosis (from −1.215 to 0.144) and the standardized residuals (±2.5) show that there are no significant departures from normality, based on the thresholds suggested by Hair et al. (2005). Moreover, given that none of the inter-item Pearson correlation coefficients in this study is greater than 0.9, the possibility of multicollinearity in the data is very low (Singh et al., 2011). From the above, it is obvious that the basic assumptions of multivariate analysis are not violated.

CFA is performed to further validate the measures for all the factors considered in this study (Figure 1). In doing so, the goodness of fit of the model to the measured data is established. More specifically, the basics of goodness of fit, the absolute fit indices, the incremental fit indices and the parsimony fit indices indicate an acceptable fit of the proposed model (Table V). From the above, it is apparent that the results consistently support the structure of the latent factors – CoQ dimensions revealed as discussed earlier in the EFA stage (Bayraktar et al., 2009).

From Table VI, it is apparent that the vast majority of the standardized regression weights are above 0.73 and that just three of them are between 0.579 and 0.651. Moreover, the respective squared multiple correlations are satisfactorily high, with the exception of three which range between 0.336 and 0.424. Thus, the factor loadings are adequately high and a high amount of measured variable’s variance is explained by a latent factor. In other words, the measured variables – CoQ elements are well loaded onto the factor – CoQ dimension they are intended to measure. This means that the unidimensionality of the CoQ dimensions is ensured (Sadikoglu and Zehir, 2010).

<table>
<thead>
<tr>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevention cost</td>
<td>4.55</td>
</tr>
<tr>
<td>Appraisal cost</td>
<td>4.48</td>
</tr>
<tr>
<td>Internal failure cost</td>
<td>4.37</td>
</tr>
<tr>
<td>External failure cost</td>
<td>4.21</td>
</tr>
</tbody>
</table>

Table IV. CoQ dimensions
Prevention costs
Appraisal costs
Internal failure costs
External failure costs

CFA model

<table>
<thead>
<tr>
<th>The basics of goodness of fit</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$</td>
<td>112.319</td>
</tr>
<tr>
<td>Probability level</td>
<td>0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Absolute fit indices</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$/ Degrees of freedom ($\chi^2$/df)</td>
<td>1.582</td>
</tr>
<tr>
<td>Root mean square of approximation (RMSEA)</td>
<td>0.080</td>
</tr>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.863</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incremental fit indices</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Incremental fit index (IFI)</td>
<td>0.949</td>
</tr>
<tr>
<td>Tucker-Lewis coefficient (TLI)</td>
<td>0.933</td>
</tr>
<tr>
<td>Comparative fit index (CFI)</td>
<td>0.948</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parsimony fit indices</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Parsimony comparative fit index (PCFI)</td>
<td>0.739$^a$</td>
</tr>
<tr>
<td>Parsimony normed fit index (PNFI)</td>
<td>0.681$^a$</td>
</tr>
</tbody>
</table>

**Note:** $^a$Relatively high values represent better fit of the model – given that the respective values for the saturated model is 0.0

---

Food manufacturing companies

**Figure 1.**
CFA of CoQ elements

**Table V.**
Goodness of fit measures

Food manufacturing companies
Reliability analysis was also applied by the internal consistency method calculating the Cronbach’s α coefficient (Sadikoglu and Zehir, 2010) and the composite/construct reliability index (Avella and Vazquez-Bustelo, 2010). Table VII shows that all the Cronbach’s α coefficients and the construct reliability indexes are above 0.798, indicating that the selected CoQ elements reliably estimate the CoQ dimensions (Hair et al., 2005; Sadikoglu and Zehir, 2010). Table VII also shows the construct validity of the CoQ dimensions. According to Hair et al. (2005), construct validity is confirmed by evaluating convergent validity (AVE > 0.502) (Kim, 2009); discriminant validity (AVE > Corr²) (Kim et al., 2012); and nomological validity (significant correlations among the extracted latent factors) (Singh et al., 2011).

4.4 Correlation analysis

Table VIII displays the results of the correlation analysis of the CoQ dimensions (prevention cost, appraisal cost, internal failure cost and external failure cost). All the relationships are

<table>
<thead>
<tr>
<th>CoQ dimensions</th>
<th>Prevention cost</th>
<th>Appraisal cost</th>
<th>Internal failure cost</th>
<th>External failure cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevention cost</td>
<td>1</td>
<td>0.676*</td>
<td>−0.607*</td>
<td>−0.524*</td>
</tr>
<tr>
<td>Appraisal cost</td>
<td>0.676*</td>
<td>1</td>
<td>−0.512*</td>
<td>−0.551*</td>
</tr>
<tr>
<td>Internal failure cost</td>
<td>−0.607*</td>
<td>−0.512*</td>
<td>1</td>
<td>0.677*</td>
</tr>
<tr>
<td>External failure cost</td>
<td>−0.524*</td>
<td>−0.551*</td>
<td>0.677*</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: “Relationship significant at the 0.01 level (two-tailed)
statistically significant \( (p < 0.01) \). The analysis indicates that the strongest relationships are those between the internal failure cost and the external failure cost (0.677), and between the prevention cost and the appraisal cost (0.676), which are almost the same.

In addition, the findings show that the negative relationship between the prevention cost and the internal failure cost \( (-0.607) \) is stronger than the negative relationship between the prevention cost and the external failure cost \( (-0.524) \). The relationship between the appraisal cost and the external failure cost is also negative \( (-0.551) \). Finally, the least significant relationship concerns the negative relationship between the appraisal cost and the internal failure cost \( (-0.512) \).

5. Discussion

The small-medium size of the Greek companies of the sample is justified taking into consideration their small market size, the low GDP per capita and the tendency toward self-employment (Hyz, 2011). The Greek food companies studied by Psomas and Fotopoulos (2010) are also SMEs.

The majority of the companies participating in the present study do measure and evaluate CoQ. This is justified given that the majority of the sample food companies have been ISO 22000 certified for many years now. This means that these companies have many years of experience in implementing an FSMS according to an international standard, which is process- and customer-oriented (Chountalas et al., 2009), and this may have made them adopt more inherently the total quality principle concerning the evaluation of CoQ. The tendency of these companies to evaluate their CoQ is also justified due to the current economic downturn and financial crisis which prevails in the Greek business environment. It seems that these circumstances may force companies be more aware of their COS. This view is not in line with Moschidis et al. (2018), according to which the unbalanced Greek economic environment, which has disorganized Greek enterprises, does not constitute a fertile ground for quality costing techniques. However, the findings of the present study with regard to the high rate of food companies measuring CoQ are in line with the respective findings of Miguel and Pontel (2004), according to which approximately 65 percent of the respondents in their research study assess such costs. The percentage of the sample companies of the present study assessing CoQ is considered to be much higher than the respective percentage of the studies conducted by Sower et al. (2007) and Glogovac and Filipovic (2018), where only 34 and 42 percent, respectively, of the respondent organizations systematically track CoQ.

According to the findings, there are not such significant reasons for not measuring CoQ, given that the sample food companies assessed half of the proposed reasons as being of medium importance and the remainder as being of low importance. This is very optimistic and reflects the possibility of future CoQ measurement by those food companies which do not measure the CoQ at present. The ignorance of the benefits of CoQ measurement is the first most significant reason for companies not measuring CoQ. The study findings of Prickett and Rapley (2001) and Tye et al. (2007) reveal that the most frequent reason for not measuring CoQ is the lack of knowledge and experience of how to track CoQ, which is in line with the findings of the present study, since this reason is considered by the sample food companies as the second most significant reason for not measuring CoQ. However, the findings from the study of Prickett and Rapley (2001) suggest that the financial difficulties of covering the required expenses is another common reason for companies not tracking CoQ, while the present study reveals that this reason is not of high importance. It is worth noting that the research of Sower et al. (2007) comes to the opposite finding compared to the present study, meaning that the primary reason for their responding companies not measuring CoQ is the lack of interest by top management. They also rank, contrary to the present study, the ignorance of the benefits of CoQ measurement and the lack of knowledge and experience of how to track CoQ as less important reasons for not tracking CoQ.
The majority of the food companies participating in the present study do not face significant difficulties in measuring the CoQ, given that all the difficulties proposed are considered by these companies as of medium or low importance. This may be justified taking into consideration the ISO 22000 implementation and certification by the majority of the sample companies. The records these companies keep in the context of the FSMS may help them collect data with regard to CoQ, and thus, not face significant difficulties in measuring CoQ. The low level of these difficulties is a significant message for those companies not measuring the CoQ and can act as a strong incentive for them to collect data for CoQ evaluation. Contrary to the findings of present study, the difficulty in identifying quality cost elements is rated as the least significant difficulty in the study of Rasamanie and Kanapathy (2011), while the extra paperwork is not rated as a significant difficulty in the study of Chatzipetrou and Moschidis (2017) in Greek food companies.

The benefits the sample food companies derive from the CoQ measurement are worth discussing. According to the present study findings, significant benefits are derived, especially with regard to the internal business environment of a company. This is an optimistic finding and another strong incentive for those companies not measuring CoQ. The findings of present study are similar to the findings of the studies of Prickett and Rapley (2001), Tye et al. (2007), Uyar (2008) and Glavan et al. (2009), according to which the identification of the areas that can be further improved, the economic determination of the losses that a company incurs due to the possible lack of quality and the CoQ determination as a good marketing tool are among the main benefits derived from CoQ measuring and reporting.

The findings of the present study validate the structure of the core CoQ dimensions. The level of all the four dimensions of CoQ is not considered high by the sample food companies. Thus, it is apparent that these companies do not invest much money in prevention and appraisal activities. This may be justified due to the medium level of the internal and external failure costs of these companies, which may motivate them to underestimate the importance of prevention and appraisal activities. The medium level of the internal and external failure costs may be due to the HACCP implementation by all the sample food companies and the ISO 22000 implementation and certification by the majority of the sample companies. The internal and external audits accomplished in the context of the FSMS make the sample food companies reduce significantly the incidences of failures in the internal and external business environment. The medium level of the prevention and appraisal cost of the sample companies demonstrates that these companies are not mature enough from the total quality perspective. A high level of total quality maturity may make companies be more proactive and invest a higher proportion of their budget in prevention and appraisal activities. Moschidis et al. (2018) also support the view that the more mature a company’s quality management system is, the more emphasis the company places on appraisal quality costs and effective use of quality costs information. However, it seems that this is not the case as far as the food companies participating in the present study are concerned.

The relationships among the CoQ dimensions are also worth discussing. Two main issues are reflected in these relationships. The first issue deals with the high positive relationship between the two dimensions of cost of conformance (the correlation coefficient between the prevention and appraisal cost is 0.676) and the high positive relationship between the two dimensions of cost of non-conformance (the correlation coefficient between the internal and external failure cost is 0.677). The positive sign of the relationship between the prevention and appraisal cost may be justified by the process monitoring and auditing implemented by the sample companies in the context of the FSMS as well as the proactive character of this system. The positive sign of the relationship between the internal and external failure cost means that improving internal failure costs (decreasing them) results in improving external failure costs (decreasing them). Moreover, increasing the internal failure costs also results in an increase in the external failure costs.
The second issue revealed from the relationships among the dimensions of CoQ is related to the negative relationships between the prevention cost and the two dimensions of cost of non-conformance (internal and external failure cost) and the negative relationships between the appraisal cost and the two dimensions of cost of non-conformance (internal and external failure cost). This means that increasing prevention and appraisal costs leads to decreasing both internal and external failure costs. However, according to the study findings, all these relationships are of a medium level, which shows the need for them to be strengthened by enhancing the costs of the food companies investing in effective prevention and appraisal activities. In other words, the inadequate prevention and appraisal activities the sample food companies have implemented until now do not ensure the absolute decrease in the internal and external failures.

The positive relationship between the prevention and appraisal cost revealed in the present study is in line with the findings of the studies of Chopra and Garg (2011) and Sailaja et al. (2014). Contrary to these studies, as well as the present study, Ramdeen et al. (2007) and Abdelsalam and Gad (2009) show that there is a negative relationship between the prevention and appraisal cost. It seems that in those cases, the preventive actions are sufficiently effective so as to reduce the necessity to implement appraisal activities. The findings of the present study regarding the positive relationship between the internal and external failure cost are in accordance with the findings of the studies of Chopra and Garg (2011) and Sailaja et al. (2014).

Several researchers such as Ramdeen et al. (2007), Sower et al. (2007), Chopra and Garg (2011) and Sailaja et al. (2014) have also asserted the negative relationship between the prevention cost and the internal and external failure cost. In contrast, Su et al. (2009) observe that the relationship between the prevention cost and the external failure cost cannot be supported and they also reach the conclusion that the prevention cost, and the internal failure cost do not have a significant real-time relationship, and that a perfect inverse relationship appears when there is a six-month time delay.

The present study findings are consistent with the findings of Chopra and Garg (2011) and Sailaja et al. (2014), according to which there is an inverse relationship between the appraisal cost and the internal failure cost. Furthermore, many studies such as those of Sower et al. (2007), Su et al. (2009) and Sailaja et al. (2014), similar to the present study, reveal that there is a negative relationship between the appraisal cost and the external failure cost. However, the research study of Chopra and Garg (2011) shows the opposite, meaning that the relationship between the appraisal cost and the external failure cost is positive.

6. Conclusions and practical implications
The measurement of CoQ and the inter-relations among the CoQ dimensions have been important research subjects for many years now. In order to further investigate the CoQ measurement within food companies, a research study was conducted in Greece. The present study contributes to the existing body of knowledge by offering deeper insights into the evaluation of CoQ in food manufacturing companies.

The study shows that the sample Greek food manufacturing companies have adequately followed the worldwide trend with regard to the measurement of CoQ. The reasons for companies not measuring the CoQ and the barriers-difficulties of companies measuring the CoQ are not significant, while the benefits derived from the CoQ measurement are significant. All the core CoQ dimensions (prevention, appraisal, internal failure and external failure cost) validated through empirical data, range within medium levels and are inter-related. More specifically, the study reveals a positive relationship between the elements of cost of conformance and between the elements of cost of non-conformance, while the relationship between the cost of conformance and the cost of non-conformance is negative.
The reasons for not measuring CoQ which are not considered important, the low level of barriers-difficulties as well as the high level of benefits of CoQ measurement, constitute significant incentives for any food company to include CoQ measurement into its daily priorities. By collecting data for evaluating the level of all the valid core dimensions of CoQ, a clear picture is provided for the existing balance between the cost of conformance (prevention and appraisal cost) and the cost of non-conformance (internal and external failure cost). This will help food companies to choose suitable strategies and allocate resources properly to improve the balance between the CoQ dimensions in favor of the cost of conformance dimensions. In so doing, the internal and external failure cost will be reduced, which in turn will make the food company more efficient and able to satisfy its customers. Thus, the foundations are laid for the food company to improve quality and competitiveness, withstand the financial crisis and survive. Companies belonging to the manufacturing sector in general can also benefit from the findings of the present study and consider CoQ measurement as the means by which they can clearly understand the pros and cons of their quality initiatives.

7. Limitations and future research recommendations
The study presented in this paper suffers from some limitations that should be taken into consideration when generalizing. First, the number of the sample food companies is not large enough, while the number of measured variables included in the model is limited, compared with all the theoretically identified CoQ elements. Second, since the data include the subjective views of only one representative from each company who is responsible for the CoQ measurement, there is a risk of receiving biased responses regarding the indicators of the CoQ. Lastly, the sample of the responding food companies is limited to companies operating in a specific country (Greece) and does not include companies operating in other European countries.

The above limitations suggest future research recommendations. Future research is recommended using a broader sample of food companies operating all over Greece as well as more CoQ elements including the hidden ones. Another avenue for future research could be to expand the study beyond the boundaries of Greece. Should a larger number of companies participate in the future research studies of CoQ, more advanced statistical methods can be applied. Moreover, using multinational data, the differences among food companies of several countries in terms of CoQ system implementation can be determined and comparisons can also be made between companies operating under conditions of economic turbulence and companies operating in technologically advanced or developed countries which do not face financial problems. Collecting subjective data from more than one company representative as well as objective data from the company files with regard to CoQ is also suggested.

Additionally, it would also be interesting to validate the underlying structure of the CoQ elements in different sub-sectors of the food sector. Given that the present research focused on the whole food manufacturing sector, it is recommended that empirical data from different food industry sub-sectors be collected in order to detect whether there are any statistically significant differences with regard to CoQ in different sub-samples within the food sector. Moreover, another proposal for future research is to confirm the findings of this study on companies that belong to the broad manufacturing and services sector as well.

References


**Corresponding author**

Evangelos Psomas can be contacted at: epsomas@cc.uoi.gr
How transformational leadership predicts employees’ affective commitment and performance

Neuza Ribeiro
CARME – Centre of Applied Research in Management and Economics, Polytechnic Institute of Leiria, Leiria, Portugal
İlhami Yücel
Erzincan Universitesi, Erzincan, Turkey, and
Daniel Gomes
Instituto Politecnico de Coimbra Escola Superior de Educacao de Coimbra, Coimbra, Portugal

Abstract

Purpose – The purpose of this paper is to examine the impact of transformational leadership (TL) on employees’ individual performance (IP) through the mediating role of affective commitment (AC). More specifically, it aims to understand how TL relates to employees’ AC, TL relates to employees’ IP, employees’ AC relates to IP and employees’ AC mediates the relationship between TL and employees’ IP.

Design/methodology/approach – A total of 476 Turkish healthcare professionals participated in this study. The mediation effect of AC in the relationship between TL and employees’ IP was tested by structural equation modeling.

Findings – The results indicate that AC mediates the relationship between TL and employees’ IP. In others words, transformational leaders promote employees’ AC which, in turn, increases their IP.

Practical implications – This study suggests that organizations should select, develop and invest in leaders who adopt a TL style because they build a climate of admiration, loyalty, respect, participation and involvement for employees which will in turn enhance their commitment and performance.

Originality/value – This study responds to calls for research studies to explore the mediating mechanism in the TL process (Judge et al., 2006), as the mediation effects explain the conditions in which TL is related to the favorable outcomes.

Keywords Affective commitment, Transformational leadership, Individual performance

1. Introduction

Transformational leadership (TL) has been considered the most influential leadership theory over the past two decades (Avolio et al., 2009; Judge and Piccolo, 2004; Sosik and Jung, 2010). TL has been defined as a set of behaviors that motivate followers to achieve performance beyond basic expectations by changing followers’ attitudes, beliefs and values (Bass, 1985; Yukl, 1999). The dynamics of TL point to followers having a strong personal identification with the leader, a shared vision for the future and the ability to work collectively for the benefit of the team (Kelloway et al., 2003), these help followers make a positive transformation in themselves (Cetin and Kinikb, 2015). Therefore, transformational leaders inspire followers to do more in their organizations than what is required (Sosik et al., 2002).

The idea that TL plays a critical role in an organization’s success is well established in leadership literature (Avolio et al., 1988; Elenkov, 2002; Wang et al., 2011; Zhu et al., 2005). A growing body of research on TL supports the suggestion that TL seems valuable, such that employees with transformational leaders demonstrate more favorable outcomes (Arnold et al., 2007; Rubin et al., 2005). For example, Limsila and Ogunlana (2008) conducted research in the construction industry in Thailand and found a positive impact of TL on
employees’ performance and organizational commitment. The effect of TL is particularly important for individuals who work in healthcare industries because these professionals often work in high pressure environments. Andrews and Dziegielewski (2005) stated that nursing staff generally prefer leaders with transformational behaviors that address followers’ individual needs. Then, the current study aims to acknowledge the relevance of TL in enhancing the affective commitment (AC) and individual performance (IP) of staff in Turkish healthcare sector.

Organizational commitment has been linked to TL across a variety of organizational settings (Abouraia and Othman, 2017; Bono and Judge, 2003; Emery and Bateman, 2007; Ismail and Yusuf, 2009; Joo et al., 2012; Malik et al., 2017; Selamat et al., 2013; Top et al., 2015). Although the empirical evidence suggests that TL is positively related to organizational commitment, when considering the impact of TL on AC and its role as a mediator variable in the relationship between TL and performance, very little evidence is available on literature.

A wealth of research exists indicating that TL is positively related to performance (Avolio et al., 2012; Liao and Chuang, 2007; Nguni et al., 2006). For example, a meta-analysis of 25 years of research based on 113 primary studies has provided strong support for positive associations between TL and IP (Wang et al., 2011). This leadership style affects organizational outcomes because transformational leaders motivate and inspire followers to achieve organizational goals, so they improve their performance. Performance beyond expectations is possible only by transforming followers’ values, attitudes and motives from a lower to a higher plane of arousal and maturity (Bass, 1985).

While the direct effects of TL on employees are well researched, the mediating role of AC in the relationship between TL and the employees’ IP is less clear. Thus, there is the need to understand the mechanism through which TL style increases IP. Possibly, those who engage in a TL style experience more support from their followers, who individually report higher levels of AC which, in turn, lead to higher levels of IP.

The current study employed AC as the mediating variable for three reasons. First is that AC is robust with strong reliability and validity (Meyer et al., 2002). Second is that among the three dimensions of organizational commitment, the affective dimension was found to correlate the strongest with individual and organizational outcomes (Meyer et al., 2002). Third, our option constitutes an answer to a call for research on the mediating mechanism in the TL process (Judge et al., 2006), as the mediation effects explain the conditions in which TL is related to the favorable outcomes.

Despite the criticisms against TL theory, we selected this construct because it has received both theoretical and empirical support more than any other leadership theory. According to Judge and Piccolo (2004), TL expresses the highest overall validity than transactional and laissez-faire leadership styles. On the other hand, TL has been the most empirically researched construct and is viewed as an effective form of leadership at the organizational, sectoral and national levels (Gyensare et al., 2016).

Moreover, concerns have been identified at the lack of context-specific research in leadership (Jordan et al., 2010; Liden and Antonakis, 2009). Therefore, developing a study of this nature in a segment of the healthcare sector in Turkey is timely and justified.

This study addresses the mentioned gaps and aims to examine the impact of TL on employees’ IP through the mediating role of AC. More specifically, it aims to understand how TL explains employees’ AC, TL explains employees’ IP, employees’ AC explains their IP and employees’ AC mediates the relationship between TL and employees’ IP. Therefore, this research intends to integrate the above constructs into a single conceptual model and it provides to the literature a more comprehensive understanding of how TL can relate to IP through AC’s mediating role.

The paper is structured as follows: the next section deals with the relevant literature and develops hypotheses. This is followed by Section 3, the research method. Section 4 presents
the main findings. The last section analyzes and discusses the findings and identifies the contributions to the literature, the implications for management, details of the study’s limitations and recommendations for future research.

2. Research background and hypothesis development

2.1 Transformational leadership

According to Bass and Riggio (2006), “TL is, at its core, about issues around the processes of transformation and change” (p. 255). In TL the relationship between leader and followers goes beyond the satisfaction of individual interests, building up a sense of common identity based on the collective ideal. Transformational leaders motivate followers to exceed their own interests on behalf of the interests of the group or organization. This occurs due to the trust, admiration, loyalty and respect that the followers feel for the leader. Transformational leaders provide deeper levels of connection and higher levels of commitment, performance and morality (Burns, 1978).

TL has been conceptualized into four key factors or dimensions (Bass and Avolio, 1990; Bass et al., 2003; Bass and Riggio, 2006; Mamede et al., 2014; Rafferty and Griffin, 2004), “the four Is” of behavior (Hoch et al., 2016): Idealized influence/charisma: reflects a leader who provides vision and sense of mission, instills pride, wins the respect and trust of followers; Inspirational motivation: the leader communicates high expectations, uses symbols to support efforts of the followers, expresses important purposes in simple ways and acts as a model of behaviors; Intellectual stimulation: the leader stimulates employees to be innovative and creative, encourages in the followers consciousness for their problems, promotes intelligence, rationality and careful problem solving; and Individualized consideration: leader gives personal attention, provides employees with feedback and delegates responsibilities to them, treats each employee individually, coaches and advises.

Even though broadly used, the conceptualization of TL in separate dimensions is problematic in the leadership literature. According to Den Hartog et al. (1997), the four-dimensional structure has not always been found. None of the individual dimensions have been clearly delineated (Mesu et al., 2015). From a theoretical perspective, the difference between idealized influence/charisma and inspirational motivation is less clear, because charismatic leaders tend to inspire people through their vision as well (Mesu et al., 2015; Rowold and Heinitz, 2007). The distinction between individualized consideration and intellectual stimulation may also be difficult because both mention employees’ development (Mesu et al., 2015).

Yukl (2006) argued that results for different components of TL measure are inconsistent. Fu et al. (2010) revealed that the dimensions are so highly inter-correlated that it is difficult to clearly determine their separate effects, even when factor analyses support their distinctiveness. Other researchers have carried out factor analyses with one-dimension scale where all dimensions loaded on one factor (Antonakis et al., 2003; Bass and Riggio, 2006). Therefore, many studies on TL have used only the composite factor rather than the four dimensions (Gyensare et al., 2016; Mesu et al., 2015; Top et al., 2013; Yucel et al., 2014; Yukl, 2006).

Prior researchers have found the positive effect of TL on employees’ attitudes and behaviors, such as job satisfaction, trust, commitment and tasks performance (Aryee et al., 2002; Avolio et al., 2004; Dhawan and Mulla, 2011; Judge and Piccolo, 2004; Sanda and Kuada, 2013) as well as negative impact on turnover intention (Dupré and Day, 2007; Gyensare et al., 2016). A TL is supposed to employ a visionary and creative style of leadership that inspires followers to make independent decisions and develop in their work (Nielsen and Munir, 2009). In sum, leadership style is a significant factor, which affects the employees’ attitudes and behaviors.
2.2 Transformational leadership and affective commitment

Organizational commitment refers to a psychological attachment of an individual toward an organization (Mathieu and Zajac, 1990). The literature indicates that of the three components of organizational commitment (affective, normative and calculative), the AC has more desirable outcomes for the organizations (Meyer et al., 2002). Mercurio (2015) suggests a conceptual framework in which AC is understood as the core essence of organizational commitment. AC can be described as the employee’s positive emotional attachment and identification with the organization (Allen and Meyer, 1990). Employees who are affectively committed to the organization want to be part of it, since they believe in its values and objectives (Allen and Meyer, 2000). Thus, organizations must be able to develop their employees’ AC (Meyer and Herscovitch, 2001).

Several studies have sought to determine the antecedents of AC (Allen and Meyer, 1996; Meyer et al., 2002; Rego et al., 2013). TL has been indicated as an important antecedent of the attachment to the organization (Shamir et al., 1993), more specifically, the literature reveals a positive relationship between TL and AC (Allen and Meyer, 1996; Bycio et al., 1995; Kark et al., 2003; Meyer et al., 2002; Podsakoff et al., 1996; Rafferty and Griffin, 2004; Yucel et al., 2014).

According to Braun et al. (2013), the identification with, and attachment to, the leader results in improved AC in the followers. Popper et al. (1992) argue that transformational leaders have an extraordinary effect on followers and their success in establishing their commitment. A transformational leader transforms and creates meaning for the employees that promotes the AC. Thus, TL appears to be particularly associated with AC (Bass and Riggio, 2006; Kane and Tremble, 2000; Meyer et al., 2002; Penley and Gould, 1988).

Social exchange theory explains the relationship between TL and organizational commitment (Blau, 1964; Gouldner, 1960; Leroy et al., 2012). When leaders seek organizational commitment from their followers, they need to engage in behaviors that go beyond the economic exchange. Transformational leaders emphasize the individual needs and personal development of their followers, encourage subordinates to do more than the expected (Bass, 1985), underline the importance of appreciating and valuing subordinates (Stone et al., 2004). As a result, followers feel trust, admiration, loyalty and respect toward transformational leaders (Yukl, 2010). In other words, these leaders transform employees by increasing motivation and commitment, and empowering them to achieve organizational goals (Yukl, 2010).

Based on the above arguments, we hypothesize:

H1. TL relates positively to AC.

2.3 TL and employees’ IP

IP is a relevant and often used outcome measure in management research and can be defined as a set of actions and behaviors that are relevant to the goals of the organization (Campbell, 1990). Thus, it is pertinent to discover the main predictors of employees’ IP. Research has traditionally tested employees’ performance as a criterion variable (Bono and Judge, 2003; Quiñones et al., 1995).

Bass (1985) stated that generally transformational leaders transform their followers toward higher performance levels. Studies that test the relationship between TL and followers’ performance have emerged (Biswas, 2014; Nielsen et al., 2009; Tse and Chiu, 2014; Vecchio et al., 2008). Researchers have examined the positive impact of TL on task performance (Aryee and Chu, 2012; Judge and Piccolo, 2004). Boerner et al. (2007) argued that TL related to follower performance and innovation. According to Rowold and Heinitz (2007), TL has a positive effect on employees’ subjective performance and on organizational profit. The influence of TL on team performance and service quality was investigated by
Lee et al. (2011). Camps and Rodriguez (2011) also found a positive relationship between TL behavior and employees’ IP.

Transformational leaders induce employees to perform beyond the minimum level of organizational expectations based on a set of leadership attitudes and behaviors: motivate and appeal to subordinates’ emotions, elicit respect from subordinates, support subordinates’ unique developmental needs and stimulate subordinates’ desire to learn and develop (Bass, 1998; Bass et al., 2003). Based on social exchange theory and the norm of reciprocity (Blau, 1964; Gouldner, 1960), these reciprocal effects between transformational leaders and their followers can be explained. The fundamental aspects of the exchange relationships between these leaders and followers are mutual benefits, trust and long-term mission achievement (Turner et al., 2002).

The transformational leaders influence followers to focus on collective interest instead of self-interest (Lussier and Achua, 2007). By doing this, they are able to stimulate followers to a higher level of performance (Yukl, 2010).

Thus, we hypothesize:

\[ H2. \] TL relates positively to employees’ IP.

2.4 AC and employees’ IP

Prior studies have argued that organizational commitment had a positive effect on job performance (Chen et al., 2006). For example, Yousef (2000) argued that organizational commitment is positively related to both job satisfaction and performance. Several researchers (Abdul Rashid et al., 2003; Chen et al., 2006; Riketta, 2002; Samad, 2005) found that committed employees are more likely to have higher work motivation, as well as higher job performance. Affectively committed employees tend to perform their jobs better and be more productive (Leroy et al., 2012; Meyer and Herscovitch, 2001; Meyer et al., 2002; Riketta, 2002).

Employees with higher levels of AC are more willing and motivated to contribute significantly toward the organization (Rego and Souto, 2004). Therefore, AC increases in-role and extra-role performance (Allen and Meyer, 1996; Jaramillo et al., 2005; Meyer et al., 2002; Riketta, 2002; Vandenberghe, 2009; Vandenberghe et al., 2004). Employees with high AC are more committed to contributing to organizational success, and they tend to improve their performance.

Thus, we hypothesize:

\[ H3. \] AC is positively related to employees’ IP.

2.5 AC mediating the relationship between TL and IP

The extraordinary commitment that transformational leaders induce in their followers may inspire their great performance (Bass and Riggio, 2006). A mechanism through which TL was related to job performance includes the followers’ positive emotion (Liang and Chi, 2013).

Transformational leaders respect their followers and are concerned with followers’ feelings and needs (Podsakoff et al., 1990). Then, they create a friendly and psychologically supportive work environment (House, 1996) and the employees’ perception of being valued and cared about by the organization influences their emotional attachment to the organization (Allen et al., 2003; Battistelli et al., 2016; Kim et al., 2016; Lee and Peccei, 2007; Lew, 2009; Sharma and Dhar, 2016). In turn, affectively committed employees tend to enhance their performance (Leroy et al., 2012; Meyer and Herscovitch, 2001; Riketta, 2002).

The study of Camps and Rodriguez (2011) found that TL increases employees’ self-perceived employability, commitment and performance. Employees who work with
transformational leaders improve the self-perception of their employability and develop their commitment to their employers who have trusted and invested in them. This increase in organizational commitment leads to higher employees’ performance (Camps and Rodriguez, 2011).

When employees perceive that an organization cares about their well-being and values their contributions (Gould-Williams, 2007; Vermeeren et al., 2011) and transformational leaders provide the necessary support to subordinates while attending to subordinates’ unique developmental needs (Bass, 1998), then the employees develop their emotional attachment to organizational goals which, in turn, result in employee’s willingness to exert effort on behalf of the organization. This proposition is based on an assumption that if leaders and organizations care for their workers, they in return will improve their performance, as a result of their AC.

Based on the above arguments, we hypothesize:

H4. AC mediates the relationship between TL and employees’ IP.

3. Method
3.1 Sample and procedures
The study comprises 476 Turkish healthcare professionals from a large hospital in a major city in eastern Turkey. Healthcare organizations have a relevant impact on society’s health and well-being but they face a multitude of problems that need to be addressed by organizational leadership. By analyzing the TL of hospital leaders and the levels of healthcare employees’ AC and their performance, the quality of patient care and hospital performance can be improved.

A survey was conducted to collect the data, meeting the aims of the study and assuring the anonymity and confidentiality of the participants’ answers. To reduce common method biases, we have fulfilled certain criteria (Podsakoff et al., 2003), notably: there was randomness in the ordering of multiple items; and we did not use scales with bipolar numerical values or assign verbal designations for the midpoints of the scales. The items were translated from English into Turkish by a first bilingual speaker and then independently back-translated into English by a second bilingual speaker (Brislin, 1980). The hospital’s HR department provided a list of all employees and their e-mail addresses. A total of 850 questionnaires were distributed, 498 were returned and 22 of them were excluded because they were incomplete. Of the 476 respondents, 57.4 percent were female and 72.5 percent were aged between 25 and 44 years (SD = 0.944; coded in the following way: 18–24; 25–34; 35–44; 45–55; ≥55). With regard to the level of education, 67 percent are graduated (SD = 0.556; coded in the following way: high school; graduate; post-graduate).

We have taken the option of using bootstrapping (Efron, 1992). Bootstrapping brings the advantage of calculating data using n replicas of our original sample, thus, helping to calculate standard error, notably, in case of distributional assumptions violation (Efron and Tibshirani, 1985). As such, using this technique provides added protection while using structural equation modeling (SEM) data analysis procedures. Given the characteristics and the advantages of using bootstrapping technique in SEM models, we have made the option of using bootstrapping while calculating our data.

3.2 Measures
The variables were operationalized through scales validated in the literature. Following Podsakoff et al.’s (2003) recommendations, in order to control for the impact of common method variance, we have performed the Harman (1967) test which suggested that the data were robust to common method variance, by pointing out that a single factor did not emerge that accounted for the majority of the covariance among the measures.
Respondents were asked to indicate the level of applicability the indicators of the different measures on a seven-point Likert scale (1: does not apply to me at all; 7: applies completely to me).

The study measures TL through 15 items proposed by Rafferty and Griffin (2004). Sample items included: “Encourages people to see changing environments as situations full of opportunities” and “Considers my personal feelings before acting.” Cronbach’s $\alpha$ was 0.93.

Based on relevant studies in the literature on TL (e.g. Aryee and Chu, 2012; Boerner et al., 2007; Fu et al., 2010; Gyensare et al., 2016; Mesu et al., 2015; Yucel et al., 2014; Yukl, 2006), this variable was considered an uni-dimensional construct.

The study measures AC through three items proposed by Rego et al. (2011). Sample items included: “I am proud to tell others that I am part of this organization” and “I feel like ‘part of the family’ at my organization.” Cronbach’s $\alpha$ was 0.84.

To measure IP, the study uses four items from Rego and Cunha (2008). Sample items included: “I am happy with the quality of my work output” and “My manager believes I am an efficient worker.” Cronbach’s $\alpha$ was 0.93.

4. Results and analysis

We have tested our study hypotheses with SEM using Amos software. For estimating the goodness-of-fit, we will report the root mean square error of approximation (RMSEA) (Steiger and Lind, 1980), comparative fit index (CFI) (Bentler, 1990), Tucker-Lewis index (TLI) (Tucker-Lewis, 1973) and $\chi^2$ values. In accordance with some authors (Browne and Cudeck, 1993; Hulland et al., 1996), we will consider a model with CFI and TLI values ranging from $\leq 0.90$ to $\leq 0.95$ combined with RMSEA values ranging from $\leq 0.10$ to 0.08 to be acceptable, and CFI and TLI values $\geq 0.95$ combined with RMSEA $\leq 0.08$ to be very good.

Table I presents the descriptive statistics and the correlations matrix of the model’s variables. Our results provide evidence that gender and age are significantly correlated ($r = 0.125; p \leq 0.01$). It is also possible to see that age and TL are also correlated with each other ($r = 0.10; p \leq 0.05$), as well as education and IP ($r = 0.116; p \leq 0.05$). The results also show that the variables of our analyzed model are significantly correlated between themselves (TL–IP ($r = 0.310; p \leq 0.01$)/TL–AC ($r = 0.656; p \leq 0.01$)/AC–IP ($r = 0.388; p \leq 0.01$)). It is also possible to verify that IP is perceived positively in this organization (mean = 5.51; SD = 1.38). As for AC, it is possible to see that workers’ AC is placed in the middle point of the scale, evidencing some degree of indifference regarding how workers are seen as having an affective bond with the organization (mean = 4.00; SD = 2.02). Finally, perceptions of TL are placed below the middle point of the scale (mean = 3.58; SD = 1.38), implying that workers may see the organization as having transformational leaders below the desirable level.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Gender</th>
<th>Age</th>
<th>Education</th>
<th>TL</th>
<th>AC</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>–</td>
<td>–</td>
<td>0.125**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>–</td>
<td>–</td>
<td>–0.058</td>
<td>0.036</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transf. leadership (TL)</td>
<td>3.58</td>
<td>1.47</td>
<td>–0.030</td>
<td>–0.10*</td>
<td>0.067</td>
<td>0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affective commitment (AC)</td>
<td>4.00</td>
<td>2.02</td>
<td>–0.005</td>
<td>–0.090</td>
<td>0.042</td>
<td>0.656**</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>Individual performance (IP)</td>
<td>5.51</td>
<td>1.38</td>
<td>–0.002</td>
<td>–0.015</td>
<td>0.116*</td>
<td>0.310**</td>
<td>0.388**</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Notes: Cronbach’s $\alpha$ reported in parenthesis. Gender: 0 for female; 1 for male. Age: 1 for 18–24; 2 for 25–34; 3 for 35–44; 4 for 45–55; 5 for $\geq 55$. Education: 1 for high school; 2 for graduate; 3 for post-graduate. *p $\leq 0.05$; **p $\leq 0.01$.
These results provide support for $H_1$–$H_3$. These results suggest that how the leadership style is developed and applied in the organization positively affects how workers develop an affective bond with the organization, and also how workers perform. In addition, the results also provide support regarding how the affective bond with the organization is positively and significantly related to employees’ performance.

Following the analysis with respect to the mediational hypothesis ($H_4$) and the goodness-of-fit of the theoretical model, the results showed acceptable fit to the data ($\chi^2(57) = 667.141, p = 0.000$; RMSEA = 0.092; CFI = 0.94; TLI = 0.93). The single-factor model revealed unacceptable fit indices ($\chi^2(54) = 1,970.923, p = 0.000$; RMSEA = 0.169; CFI = 0.77; TLI = 0.75). Table II provides the fit indices and Figure 1 presents the analyzed model.

In order to access the mediational effect if this mediation is either partial or total, we have made reference to the procedure recommended by Kenny and Judd (1984) for assessing mediational effects using SEM, and report the direct, indirect and total effects (standardized effects). Table III provides the standardized total, indirect and direct effects, for the purposes of testing the mediation hypothesis.

The results have shown that the direct effect of TL on IP is non-significant (direct effect = 0.076; $p > 0.05$), and the total effect on IP is significant (total effect = 0.327; $p \leq 0.01$).

<table>
<thead>
<tr>
<th>Path</th>
<th>Total effect (TE)</th>
<th>Indirect effect (IE)</th>
<th>Direct effect (DE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transf. lead. (TL) – Ind. perform. (IP)</td>
<td>0.327**</td>
<td>0.251**</td>
<td>0.076</td>
</tr>
</tbody>
</table>

**Note:** **$p \leq 0.01$**
as well as the indirect effect through AC (mediator) (indirect effect = 0.251; \( p \leq 0.01 \)). According to these results, \( H4 \) is supported, showing a total mediation effect of AC on the relation between TL and IP.

5. Discussion and conclusions

5.1 Discussion

The results indicate that transformational leaders induce followers’ AC and this is consistent with other studies (Allen and Meyer, 1996; Bycio et al., 1995; Kark et al., 2003; Podsakoff et al., 1996; Meyer et al., 2002; Rafferty and Griffin, 2004; Shamir et al., 1993; Yucel et al., 2014). TL causes employees to perceive that the organization supports, values and cares them and leads to attachments among the organization’s members and develops a high level of AC to the organization.

This relationship should be reciprocal and it can be explained by the social exchange theory (Leroy et al., 2012). In organizations, social exchanges are relevant because they can be decisive regarding organizational performance and IP. They can be used to explain the link between employee perceptions of workplace aspects and their subsequent attitudes and behaviors. These exchanges are characterized by co-operation and reciprocity, that is, when one partner provides a benefit to another, the act produces a sense of obligation on the part of the latter to reciprocate at some point in the future. Specifically, based on the norm of reciprocity, transformational leaders, through inspiration and high expectations they have for employees, their vision for the organization and the support they provide that fosters social support, teamwork, self-actualization and goal achievement, they thus increased levels of commitment among employees (Simosi and Xenikou, 2010). Therefore, employees who develop high-quality exchanges with their transformational leaders reciprocate by displaying positive attitudes, such as, AC.

In the same vein, TL may be an effective mechanism to improve employees’ IP as shown by our results and literature reviewed (Biswas, 2014; Camps and Rodriguez, 2011; Nielsen et al., 2009; Tse and Chiu, 2014; Vecchio et al., 2008; Yukl, 2010). Employees who work under transformational leaders understand the organizational vision and mission very clearly and consequently perform better (Avolio, 1999). Such leaders enhance employees’ expectations and recognition of their work and increase employees’ IP through TL behaviors such as individual attention, inspiration, intellectual stimulation and motivation.

This relationship can be understood as a reciprocal exchange because employees become committed to producing better job outcomes when they are valued by leaders. In other words, if the employees receive important and valuable leader’s support, driven by a sense of obligation, and based on the principle of mutual benefit, not only do they reveal AC, they also show enhanced performance to help the organization to achieve its goals.

The current study reveals a full mediation of AC in the relationship between TL and employees’ IP. It suggests that transformational leaders adopt behaviors that motivate followers to perform and identify with organizational goals and interests and have the capacity to motivate employees beyond expected levels of work performance. As a result, employees feel affectively committed and personally rewarded through work, and their performance is enhanced.

In sum, transformational leaders through charisma, inspirational motivation, intellectual stimulation and individualized consideration create an environment where employees develop a sense of organizational identification and have more a positive leader–member exchange relationship. Also employees should be more affective to the organization, more satisfied and more productive (Hendrix et al., 2015).
5.2 Limitations and future studies

This study presents several limitations. First, TL, AC and IP were measured from the same individuals using the same questionnaire at a single time. This raises the risk of introducing common method variance (Podsakoff et al., 2003). To address this concern, some preventative methods were undertaken, such as the Harman test and other procedures described earlier. Future research may explore the effect of TL on subordinate outcomes using longitudinal designs or via a multiple-source method. Second, the study included only one mediating variable, but others are plausible. For example, it is possible that TL develops well-being and engagement which, in turn, increases employees' performance. Third, moderating variables were not included. Future research may test, for example, the degree to which some personal characteristics moderate the relationships between TL and the dependent variables. Moreover, future studies, particularly experimental studies, may explore these effects under a variety of task scenarios. The different scenarios can be used to explore the effects of TL on subordinate outcomes moderated by other variables.

5.3 Theoretical contributions

From a theoretical perspective, the present study confirms the results of prior studies that found positive effects of TL on employees' AC and performance (Biswas, 2014; Camps and Rodriguez, 2011; Kark et al., 2003; Meyer et al., 2002; Nielsen et al., 2009; Rafferty and Griffin, 2004; Tse and Chiu, 2014; Vecchio et al., 2008; Yucel et al., 2014; Yukl, 2010). Therefore, the current findings support theoretical results that suggest the quality of relationships between leaders and followers may play an important role in the extent to which followers develop positive attitudes and behaviors.

This study answers the call for further research from those who have argued that more information is needed to understand the process through which TL influences employees' performance (Avolio et al., 2009). Then, the present research sought to integrate TL, AC and IP into a single conceptual model. It also constitutes an answer to a call for research on the mediating mechanism in the TL process (Judge et al., 2006), as the mediation effects explain the conditions in which TL is related to the favorable outcomes. Therefore, this study addressed significant research gaps by testing the relevant variables' mediating effects. It provides a more comprehensive understanding of how TL can relate to IP through AC's mediating role.

5.4 Implications for management

This study makes a significant practical contribution by providing advice to leaders and organizations on how to create an atmosphere of admiration, loyalty, respect, participation and involvement for employees which will, in turn, help to enhance their commitment and performance. Therefore, organizations must seek those with TL ability because they bring special assets to their organizations.

With this knowledge, organizations should select, develop and invest in leaders who adopt a TL style. These leaders communicate a compelling vision; provide symbols and emotional appeals to increase awareness of mutual goals; develop team spirit in followers; facilitate the efforts of followers to become more innovative and creative; stimulate followers to adopt new work perspectives; provide a supportive climate; listen to the individual needs of the followers; treat followers with respect; facilitate individual growth by inspiring, coaching, teaching and mentoring; and delegate to help followers achieve their tasks and grow through individual challenges.

In sum, developing positive leadership styles (such as TL) and improving employee attitudes (such as AC) are important strategies to promote employees' performance. The results of this research were expected to be beneficial to the management of healthcare services in showing that the TL style used by supervisors might engender a sense of AC among healthcare staff that would lead to a performance improvement.
References


**Further reading**


**Corresponding author**

Neuza Ribeiro can be contacted at: neuza.ribeiro@ipleiria.pt

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldsight.com
Productivity paradox?  
The impact of office redesign on employee productivity

Erlinda N. Yunus and Erni Ernawati
Department of Operations Management, Sekolah Tinggi Manajemen PPM, jakarta, Indonesia

Abstract
Purpose – The purpose of this paper is twofold: first, to investigate the relationship between office redesign and employee productivity; and second to highlight the impact of privacy on work productivity across different generations.

Design/methodology/approach – This study examines open-office policy more comprehensively by integrating socio-behavioral and physical aspects of the office, and by using a mixed-method approach that incorporates most significant change, factor analysis and hierarchical regression analysis. Using a census method, the respondents were all consultants and trainers in an educational institution who were experiencing office design changes from a combi, cellular-like office to a more open, non-territorial office.

Findings – Three variables emerged as impacts of office redesign perceived by respondents: friendship, collaboration and privacy. Collaboration and privacy exert a positive influence on work productivity, while friendship does not. The relationship between privacy and work productivity is stronger for the Generation Y than for senior employees, namely, the Baby Boomers and Generation X.

Research limitations/implications – This study examines the impacts of office redesign in one organization. Future studies should advance the findings by empirically testing the theoretical model in broader contexts. Future studies could also enrich the literature by bringing cultural aspects into the discussion and comparing Asian-based and European or Western-based findings.

Practical implications – For Gen Y employees who prefer freedom, mobility and flexibility to personalization in their workplace, the open office could be a better solution for organizations that aim for both work productivity and efficiency.

Keywords Privacy, Friendship, Collaboration, Productivity, Open office, Generation Y

1. Introduction
Productivity is an important aspect of an organization. Generally, the concept of productivity indicates the ratio of output to input (Pritchard, 1992 in Pritchard et al., 2008; de Been et al., 2017). At the individual level, work productivity is the employees’ capacity to generate products or services needed to achieve organization goals. A high work productivity would improve organization performance (Huselid, 1995; Prakash et al., 2017) such that organizations will seek the best ways to enhance their employees’ productivity to generate positive and sustainable impacts.

Office design is one of aspects that are found to directly impact work productivity. The relationship between office design and work productivity has been acknowledged since the 1930s (e.g. a study by Pennock, 1930; Mayo et al., 1939). Initially, an office was designed with cubicles, known as a cellular-type office; however, since the 1950s, open-plan designs were installed and became increasingly popular through the 1970s due to their efficiency (Brennan et al., 2002; Brunia et al., 2016). Open offices eliminate separating walls, lower the
partitions between desks and hence require less space. DTZ Research Data showed that the cost of office space is approximately $20,000 per workstation per annum in London and Hong Kong (Kuljanin, 2014). Even in Indonesia, a developing country with a rapid-growth economy, the office cost of its two major cities (i.e. Jakarta and Surabaya) could reach $5,000 and $2,000, respectively, per workstation per annum. Not to mention that Jakarta has had the highest growth of office cost in the world for two years in a row since 2013, i.e., a 34 percent increase (Kuljanin, 2014).

Consequently, efficiency becomes the keyword for transforming a cellular design to an open-plan office (Van der Voordt, 2003; Laihonen et al., 2012; Kim et al., 2016). On average, through an open office, space allocated per workstation (in meters squared or m²) is reduced up to 50 percent (Roelofsen, 2008; Kim et al., 2016) and firms save up to 80 percent per annum on costs related to office leasing (James, 2016). Moreover, since the 1990s, firms could opt for flexi-desking or desk-sharing concepts, where employees do not occupy one particular desk, but instead share it with other colleagues when not in the office (Al Horr et al., 2016; Kim et al., 2016). Nowadays, more than 70 percent of American and European firms apply open-office and flexi-desk concepts (GSA, 2011). IBM was able to save office costs up to $100m annually by applying the concept of flexi-desk in its US and European offices, whereas the Sabre Corporation, a technology company based in Texas, USA, implemented the same program and was able to save up to $10m in the last three years (GSA, 2011).

Nevertheless, alterations of office design may also impact employees’ satisfaction and productivity. On the one hand, open offices consume less space and cut costs; on the other hand, employees may feel noise, distractions (Roelofsen, 2008; McElroy and Morrow, 2010; Kim and de Dear, 2013) and less privacy (Ding, 2008; Kim and de Dear, 2013; de Been and Beijer, 2014) in such settings. An intensive study by Davis et al. (2011) indicated that open offices reduce employees’ attention, productivity, creativity and satisfaction.

It comes as no surprise, then, that previous studies have emphasized the negative relationship between open offices and work productivity (Oldham and Rotchford, 1983; Brenman et al., 2002; Veitch et al., 2007; McElroy and Morrow, 2010; de Been and Beijer, 2014). However, further studies are still required to seek employees’ perceptions related to office redesign and to provide a better understanding of the link between office design and employees’ productivity. Furthermore, as the open-office controversy still persists (Brunia et al., 2016), further studies need to shed light upon the different preferences among generations regarding office designs, considering that the open-office concept is frequently deemed more positive for younger generations who do not fuss over a private workstation like their senior counterparts (McElroy and Morrow, 2010; Konnikova, 2014). Open offices are perceived to increase communication and interaction among employees (Elsbach, 2003; Bruniia et al., 2016) and to develop autonomy and flexibility (Kim et al., 2016), which are important factors for the young generations.

This study deeply explores significant changes experienced by employees after an office redesign. The research questions that guide the study are as follows:

**RQ1.** What significant changes do employees experience after an office redesign?

**RQ2.** Do the changes exert a positive influence on work productivity?

**RQ3.** Would the changes be different for younger generations than older generations?

The study contributes to the literature by deepening our understanding of the impacts of office redesign on employees; critically investigating open-office policy from the view of socio-behavioral and physical aspects, particularly among different employees’ generations; and using more thorough mixed-method research that combines qualitative-quantitative techniques of the most significant change (MSC), factor analysis and hierarchical regression analysis.
Moreover, the current study is different from previous studies in that this study presents a different cultural context than the predominantly western open-office studies of the past. Given the nature of a mixed-method study, the exploratory research preceded the in-depth literature review; thus, this paper is organized as follows. The next section, Section 2, describes the qualitative study stage – both the methodology and the findings – which resulted in changes experienced by the employees as predictors of employee productivity. Section 3 synthesizes the extant literature and formulates the hypotheses, followed by Section 4, which describes the quantitative study stage. Section 5 discusses the findings in detail and their implications. This paper concludes with a brief summary and description of potential future research.

2. The qualitative study
The current study investigates the significant changes experienced by employees after an office redesign that encompasses layout, office desk design and decorations. The studied office was formerly organized in a combi style, which, although it was already based on an open-plan design, enclosed every two or four workstations with high partitions and thus resembled "cubicles." Each cubicle was used by 2–4 employees, who each had their own desk and bookshelf. Other fully closed rooms were available for formal meetings. The management decided to change this old office plan into a more open office with non-territorial desks (also known as a flex office), where high partitions were replaced by very low partitions (approximately 30 cm) that divided the workstations. No particular desk is assigned to any one employee, as the desks are shared and used as needed. The bookshelves were removed and replaced by lockers. The ratio between employees and workstations is now 3:1.

The redesign was deemed quite drastic and thus triggered resistance. This situation was worthy of an in-depth study.

2.1 Unit of analysis and the study context
The unit of analysis of this study is at the individual level. Since an office redesign is expected to result in significant changes for employees, the target of this study was employees who were directly exposed to the office changes, specifically, from the previous cubicle office design to a more open and flexible design.

The research was conducted in an educational institution that has been established for 50 years in Jakarta, Indonesia. The institution (hereafter abbreviated as INST) is a foundation with a mission of disseminating management education throughout Indonesia, an emerging country in Southeast Asia. INST has three strategic business units: a business school, a short course or training provider and a consulting firm.

The top management of INST recently applied a flex-office policy, replacing the cubicle-type office with a more open, non-territorial office. The cubicles, which used to be workstations for employees, were substituted with low-partition desks. The employees were allowed to occupy any desk since the policy released desk possession and instituted flexible desk ownership. This design was only implemented in one working area of the INST, that is, the area for trainers and consultants.

2.2 Most significant change
The first stage of the current study was a qualitative study to explore the experiences and the changes felt by the employees after the implementation of an open-office design. At this stage, we applied the MSC technique. The MSC is a method for monitoring a program or policy and is commonly utilized in the socio-economics sector for assessing or evaluating results and impacts of a program (Davies and Dart, 2005; Wilder and Walpole, 2008). The MSC systematically collects pieces of evidence of real changes which are often
neglected or ignored in general monitoring techniques (Dart, 2000); therefore, the method was deemed appropriate in the context of a more flexible office. The MSC is able to accurately identify significant changes because it explores experiences of the employees using in-depth interviews, which were presented in the form of stories.

The MSC’s processes, as advised by Davies and Dart (2005), are summarized into seven stages:

1. Performing an initial orientation by identifying objects of the study: the researchers obtained an approval from the INST and visited the flex-office area. A preparation was conducted, including defining stakeholders who would provide information regarding the significant changes.

2. Defining the domain of change or change areas.

3. Collecting stories of significant changes from the selected informants; the interview was conducted once for each informant for 30–60 min. For this study, the researchers gathered five stories from five informants who were the target of the study. These five informants were selected through diligent consideration so that all aspects were covered: work tenures (senior and junior), positions in the organization (managerial and non-managerial), as well as authorities in the organization (decision maker and non-decision maker).

4. Verifying the stories: the verification process was applied to each of the informants’ stories. Since the study was conducted by two researchers, the researchers exchanged roles at the verification stage; if one story was composed by one researcher, the other researcher verified it. Each verification lasted for about the same length of time as the initial interview.

5. Selecting the most significant story using a focus group discussion (FGD) technique: the FGD was conducted with eight participants who were different from the informants. Those who were invited to join the FGD were chosen based on similar considerations (as previously discussed) so that these participants represented each aspect of the target population. The first step of the FGD involved the reading of all stories by the researcher. The next step was a discussion among the FGD participants to define which story was the most representative in outlining the flex-office experiences. Finally, the FGD participants agreed on the significant impacts of the flex-office policy.

6. Providing feedback from the FGD to the previous informants, particularly regarding the selected story and the reasons for the selection of that story.

7. Composing the operational concept: the selected story was then transformed into a collection of variables denoting significant changes precipitated by office redesign. The researchers developed indicators for each variable for the quantitative study. This final step was a preparation for collecting the overall data. A questionnaire was standardized to all employees in the open-office area.

The profiles of the 5 informants, 8 FGD participants and 36 respondents (or the entire population) are provided in Tables A1 and AII. The number of female and male employees was relatively balanced (i.e. 52.8 and 47.2 percent, respectively). The majority (more than 60 percent) of employees had worked in the office for less than 5 years, while 22.9 percent of employees had worked for 20 years or more. Regarding age, 34.3 percent of the employees were of a younger generation (less than 30 years old), 28.6 percent were 30 to less than 40 years old, 17.1 percent were between 40 to less than 50 years old, and the rest (20 percent) were 50 years old or more. Before joining the current organization, the majority of
employees (83.3 percent) have joined other companies; most of them worked in an open office with dedicated desks.

Five stories were summarized from interviewing the informants and were then compiled by the researchers. Each story was given a title. The stories are as follows:

1. A story about A, one of the directors at INST, entitled “Mending a Neglected Family.”
2. A story about B, a senior trainer/consultant at INST, entitled “I Need a Place that is My Place.”
3. A story about C, a young department head at INST, entitled “It Is More Spacious and Not Crowded Anymore.”
4. A story about D, a junior trainer/consultant employee at INST, entitled “Now It Is Not Messy and Stressful.”
5. A story about E, a young administrative staff member, entitled “People Becoming Positive and Happy.”

2.3 Variables emerged from the MSC

The uniqueness of the MSC technique is that it values each story as a representation of the informant’s true experiences that has its own meaning. Therefore, the MSC technique suggests that researchers select one story that mostly represents the changes in the office dynamic. The FGD was one step that must be conducted in the MSC technique in order to define the MSCs experienced by the member of the population in the study. Based on the FGD, the story of C was perceived as the most representative of the feelings of the FGD participants. Two themes emerged from this story: collaboration and friendship.

However, the FGD participants disagreed that this one story could portray the overall experiences they felt, and thus the consensus reached among the participants was to add another theme that was deemed pertinent as one of the significant changes: privacy. The FGD participants felt that the previous design gave more privacy than the new one since it provided cubicles for employees. The new design, on the other hand, has lower partitions and places the workstations side by side to create opportunities for discussion and interaction. The new situation creates a more crowded atmosphere and reduces concentration, hence the lesser sense of privacy. Therefore, this study examined privacy from the psychological standpoint, or, to use a term suggested by Sundstrom et al. (1980), “psychological privacy.”

The participants thus settled on three variables as the outcomes of the office redesign (i.e. MSCs), that is, privacy, collaboration and friendship. These three themes became the all-encompassing variables depicting the changes of the office redesign. Drawing from previous studies, privacy and collaboration were abundantly discussed as factors emerging from the open-office design. Meanwhile, friendship is a new variable in the context of open offices which provides a unique contribution to the literature.

The privacy issues (in terms of distractions, interactions and lowered concentration) were quite obvious and mostly raised by the senior FGD participants (i.e. Generation X or Baby Boomers); however, it was not the main concern of the young FGD participants (Generation Y; those who were less than 30 years old or were born in and after the 1980s). This matter will be further tested.

The researcher prepared a survey that included three variables as the predictors (privacy, collaboration and friendship) and one variable as the criterion (productivity). The assessment toward the instrument was conducted through validity and reliability checks using exploratory factor analysis and Cronbach’s α, which will be explained further in the Quantitative Study section.
3. Hypotheses formulation

3.1 Office design and productivity

Extant studies categorized office designs into three types: individual and shared room, combi office and flex office (de Been and Beijer, 2014). This categorization was based on architectural features (i.e. layout and design) and functional features (i.e. the functions of the work office). The individual and shared room category indicates offices where a particular working space is designated for each employee. The combi office refers to a particular office provided for each employee, with an additional space for special activities that require high concentration, as well as for formal and informal meetings. Meanwhile, the flex office refers to the absence of dedicated space for each employee; in other words, it is utilized for all employees’ activities (de Been and Beijer, 2014).

There have been a number of studies identifying the impact of the office environment on employee welfare, satisfaction and productivity (e.g. DeCroon et al., 2005; Bodin-Danielsson and Bodin, 2008; de Been and Beijer, 2014). These findings are pertinent since satisfaction with the physical environment is apparently related to work satisfaction and productivity (de Been and Beijer, 2014; de Been et al., 2017). The office environment’s ambiance and layout also influenced employees’ behaviors and productivity, boosted learning (Kasuganti, 2018) and aided knowledge sharing among employees (Haynes et al., 2017).

Open offices gained attention because companies found that they could improve communication and collaboration through the release of cubicles, which encouraged employee interaction (Brunia et al., 2016). There are two basic approaches that are used to recognize how an open-plan office impacts employees: social relations approaches and socio-technical approaches. The social relations approach argues that whenever interior partitions and barriers are abandoned to create an open-plan office, social relations among employees will be nourished, which has a positive impact on the employees’ motivation and satisfaction (Bach, 1965; Zeitlin, 1969).

Meanwhile, those who support the socio-technical approach declare that physical boundaries influence work experience in two ways. First, boundaries may change the work area into a more private workstation. When an area is bounded by a partition, employees’ privacy is institutionalized. This design supports personal conversation and information sharing. Furthermore, the bounded area reduces external disruptions (Gyllenhammar, 1977). Second, a bounded work area defines the characteristics of work processes undertaken by employees (Cummings and Griggs, 1976). When a work area is surrounded by partitions, employees are able to recognize their own tools and equipment, as well as the technology available for them to accomplish their tasks.

Studies have found several factors that are affected by the implementation of open offices, such as privacy, noise and distractions (Roelofsen, 2008; Brunia et al., 2016), as well as communication and social interactions (de Been and Beijer, 2014; Haynes et al., 2017). The benefits of the open office often clash with its negative impacts and create a counterproductive situation for companies. Therefore, three variables (i.e. privacy, collaboration and friendship) – which potentially contradict one another – will be examined in detail.

3.2 The relationship between privacy and work productivity

Studies indicate that there is an increasing priority for privacy at work (Oldham, 1988; Sundstrom et al., 1980). Privacy could refer to architectural privacy, which is naturally provided by the acoustic and physical aspects of an office environment, or psychological privacy, which concerns an employee’s sense of control over their given environment (Sundstrom et al., 1980). In an open office, privacy is significantly reduced when partitions that divide workstations are lowered or altogether eliminated and the desks are placed adjacent to one another. This more open environment encourages more interactions among employees, yet also generates more noise (Roelofsen, 2008).
Employees who feel dissatisfied with their work environment and privacy do not perform well (de Been et al. 2017). Working in a combi office or flex office, where employees can move to and choose to work at various office areas, apparently lowers employees’ productivity, concentration and privacy, as compared to those who work in an individual and shared room. Yet compared to those who are in an individual and shared room, in combi offices (not in flex offices), employees are more satisfied because they are able to communicate with others (de Been and Beijer, 2014).

It seems that people prefer a workstation that enables them to keep their privacy rather than one with less privacy (Hedge, 1982; Brunia and Hartjes-Gosselink, 2009). In an open office, employees are more stimulated by the close relationships they form with those who sit near them (Oldham, 1988); nevertheless, it is more difficult for an employee to avoid contact with his/her colleagues and to maintain the level of privacy he/she needs to perform his/her job.

The ability to work individually without disruption is an important factor in office design. More than 50 percent of employees who work in open-plan offices revealed that they were disturbed by noise during their work, particularly noise coming from phones, employees’ chats, air conditioners and automatic machines in the office (Sundstrom et al., 1994). The difficulty of concentrating in an open-plan office was caused by the increasing occasions of communication and disruption among employees (Banbury and Berry, 2005). These disruptions were a negative factor for productivity (Haynes, 2008; de Been et al., 2017).

On that note, the first hypothesis argues that privacy would positively impact employees’ productivity since the ability to control disruptions and have more concentration at work would be directly related to the ability to perform the assigned tasks well:

\[ H1. \text{ Privacy is positively related to work productivity.} \]

3.3 The relationship between collaboration and work productivity

The physical environment of an office should support employee activities, including optimal communication, concentration and mood (de Been et al., 2017). Studies have found that the concept of the open office provides the most conducive environment for knowledge creation, as it allows employees to interact and collaborate in a spontaneous manner. An open office would stimulate interactions and informal conversations, which could lead to both information sharing and learning (de Been et al., 2017; Kasuganti, 2018). Increased communication and interaction could therefore lead to more effective collaboration as well (Heerwagen et al., 2004).

De Been and Beijer (2014) further argued that open offices, either the combi (that is based on an open or half-open office with an assigned workstation for each employee) or the flex (similar to the combi office but with no assigned workstation) increases the frequency of interactions among employees more than the individual and shared room offices. The interaction that could progress to collaboration among coworkers plays a role in their productivity (Brill and Weidemann, 2001; Haynes et al., 2017).

Therefore, this study further conjectured that collaboration positively impacts work productivity:

\[ H2. \text{ Collaboration is positively related to work productivity.} \]

3.4 The relationship between friendship and work productivity

Previous studies affirmed that the physical conditions of an organization would eventually externalize into employees’ work performance. Thus, transforming the work environment as a medium for an organizational development was suggested by numerous studies (e.g. Allen, et al., 2005; de Been et al., 2017; Kasuganti, 2018). In their study, Allen...
et al. (2005) evaluated a number of cases in the British Government and proposed that the layout of office workstations could be optimized to improve employees’ collaboration and openness, hence enhancing both organizational and individual performances. The increased collaboration and openness among employees would trigger friendship and stimulate work satisfaction. The more open layout also increased friendship opportunities (Rasila and Rothe, 2012). Accordingly, our third hypothesis incorporates friendship, which also improves employees’ productivity:

\[ H_3. \text{Friendship is positively related to work productivity.} \]

3.5 The impact of work generation type on the relationship between privacy and productivity
The diversity of employees’ generations could serve as one determinant in office design, as open offices are perceived more positively by younger generations who are not concerned with personalization at their office and who possess the capability to work with high disruptions as compared to the older generations (McElroy and Morrow, 2010; Rasila and Rothe, 2012; Konnikova, 2014). An open office also enhances communication and interaction (Elsbach, 2003) and improves employees’ autonomy and flexibility (Kim et al., 2016), which are important factors for younger generations at work. Young generations do not feel that distractions are a nuisance (McElroy and Morrow, 2010) and consider the drawbacks of the open office as “fair trade-offs” for the improved communication and knowledge sharing they would receive (Rasila and Rothe, 2012).

Therefore, the final hypothesis is concerned with employees’ generations as moderating factors, which will alter the relationship between privacy and work productivity:

\[ H_4. \text{The relationship between privacy and work productivity is moderated by employees’ generations, such that the younger the generation, the higher the ability to control distractions in a lowered privacy situation, and the more the situation improves their performance.} \]

3.6 Research framework
Based on the aforementioned literature and the findings of the qualitative stage of this study, we propose a research framework as depicted in Figure 1.

4. Quantitative study
Subsequent to the qualitative stage, the second stage of the current study involved collecting quantitative data using a survey. This study conducted the survey after all variables were formulated into measurable indicators based on previous literature. This study measured friendship through three indicators – mutual trust, commitment, and...
shared interests and values (Nielsen et al., 2000; Berman et al., 2002); privacy through three indicators – visual disruptions and interruptions, speech privacy, and control over accessibility (Brill, 1985 in Ding, 2008; McElroy and Morrow, 2010; Leadon, 2015); and collaboration through three indicators – knowledge transfer, coordinated work, and ideation and creative development (McElroy and Morrow, 2010; Leadon, 2015). Productivity was measured using two indicators: performance and time management (Brennan et al., 2002; Al Horr et al., 2016). Each of the indicators owned a minimum of three items. The scoring was on a differential semantic scale, ranging between −3 to +3.

A complete instrument was composed and formed into a self-administered questionnaire that consisted of three sections: the respondent profile, the measurement of the MSCs and the measurement of work productivity. The population of Stage 2 was all employees who experienced the flex-office policy. They were 36 trainers and consultants. This study employed a census method and each respondent filled out the questionnaire independently.

4.1 Exploratory factor analysis (EFA)
Before doing the hierarchical regression analysis, factor analysis was carried out to ensure that factors from the initial stage were valid to proceed to the stage of examining the impact. The steps of factor analysis are as follows:

1. assessing factors through a data cleaning and KMO Bartlett analysis in order to obtain consistent KMO Bartlett values;
2. performing factorization and rotation to define new and more parsimonious factors that represent each variable; and
3. validating each factor to ensure that they can be generalized into the population and giving a suitable new name to each factor.

New variables (or previously, factors) generated from factor analysis were further examined using hierarchical regression analysis.

The researcher performed several tests before ensuring the conformity of the EFA, that is, using Bartlett and KMO tests for each variable (friendship, privacy, collaboration and productivity). The Bartlett test for four variables was significant \( p < 0.01 \), which indicated that the data did not generate an identity matrix and therefore could be accepted for factor analysis (Hair et al., 2010). The KMO statistics for friendship, privacy, collaboration and productivity were 0.809, 0.806, 0.779 and 0.717, respectively; this means that they were accepted since all were above 0.50 (Hair et al., 2010).

The orthogonal matrix (or varimax) was opted to facilitate the interpretation of the results. The EFA for friendship resulted in three components as expected, where the loading for all items exceeded 0.40 (valid). The items of privacy, which initially were categorized into three components, did not load highly on the respected component; therefore, one iteration was applied by reducing items one-by-one such that each item was included exclusively to only one component. The component was named “Distraction control” in order to accurately describe the measurement of privacy.

The same procedure was employed for other variables. The results of EFA were further checked for reliability using Cronbach’s \( \alpha \) to ensure the consistency of each component or variable. The summary of the EFA can be seen in Table I.

4.2 The hierarchical regression analysis
The descriptive statistics of the variables are displayed in Table II.

To investigate the contribution of each variable, the analysis was accomplished with a hierarchy model. Privacy, as the most frequently examined variable in the context of the
open office, was entered into Model 1, followed by collaboration into Model 2. Friendship, which was a new variable in the open-office study, was included in Model 3. Finally, in order to examine the last hypothesis related to the different generations as a moderator, the interaction between privacy and generation was entered into Model 4. The results of the overall models are displayed in Table III.

The first hypothesis of this study conjectured that privacy would have a positive impact on work productivity. The first model was significant ($F = 12.587; p < 0.01$) and privacy explained 28.2 percent of the variance of productivity. This result indicated that the regression coefficient is significant ($p < 0.01$) and positive ($\beta = 0.346$) and supported $H_1$.

The second model was examined for the second hypothesis, where collaboration was predicted to have a positive impact on work productivity. This model was also significant

---

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Components</th>
<th>No. of Items</th>
<th>Cronbach’s $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendship</td>
<td>“Nonexclusive workplace relations that involve mutual trust, commitment, reciprocal liking and shared interests or values” (Berman et al., 2002)</td>
<td>Mutual trust</td>
<td>4</td>
<td>0.898</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commitment</td>
<td>3</td>
<td>0.847</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shared interests and values</td>
<td>5</td>
<td>0.901</td>
</tr>
<tr>
<td>Privacy</td>
<td>“Personal control over interactions and/or communications with others” (Brill, 1985 in Ding, 2008)</td>
<td>Distraction control</td>
<td>7</td>
<td>0.894</td>
</tr>
<tr>
<td>Collaboration</td>
<td>“The degree to which an organization has a collaborative culture” (McElroy and Morrow, 2010)</td>
<td>Knowledge transfer</td>
<td>5</td>
<td>0.886</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ideation and creative development</td>
<td>5</td>
<td>0.893</td>
</tr>
<tr>
<td>Productivity</td>
<td>“The ratio of output to input” (Oseld and Bartlett, 1999 in Al Horr et al., 2016)</td>
<td>Productivity</td>
<td>5</td>
<td>0.763</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Components</th>
<th>No. of Items</th>
<th>Cronbach’s $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy</td>
<td></td>
<td>Distraction control</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Knowledge transfer</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ideation and creative development</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mutual trust</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commitment</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shared interests and values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaboration</td>
<td></td>
<td>Knowledge transfer</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ideation and creative development</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mutual trust</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commitment</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shared interests and values</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friendship</td>
<td></td>
<td>Mutual trust</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commitment</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shared interests and values</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productivity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Privacy</th>
<th>Collaboration</th>
<th>Friendship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction control</td>
<td>Knowledge transfer</td>
<td>Ideation and creative development</td>
</tr>
<tr>
<td>0.498**</td>
<td>0.430*</td>
<td>0.735**</td>
</tr>
<tr>
<td>0.612**</td>
<td>0.524**</td>
<td>0.571**</td>
</tr>
<tr>
<td>Mean</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>$-0.2857$</td>
<td>$1.31206$</td>
<td></td>
</tr>
<tr>
<td>$0.8686$</td>
<td>$1.1509$</td>
<td></td>
</tr>
<tr>
<td>$0.3647$</td>
<td>$1.14359$</td>
<td></td>
</tr>
<tr>
<td>$0.5714$</td>
<td>$1.37543$</td>
<td></td>
</tr>
<tr>
<td>$1.2286$</td>
<td>$1.02827$</td>
<td></td>
</tr>
<tr>
<td>$0.6914$</td>
<td>$1.22555$</td>
<td></td>
</tr>
<tr>
<td>$0.5371$</td>
<td>$0.94778$</td>
<td></td>
</tr>
</tbody>
</table>

Notes: **Correlation is significant at the 0.05 and 0.01 levels, respectively (two-tailed)
and both variables (i.e. privacy and collaboration) were able to explain 43.3 percent of productivity variance. Collaboration itself contributed 15.1 percent and was a significant contribution ($p < 0.05$). However, assessing both indicators of collaboration more deeply, the researcher found that only ideation and creative development had a significant impact on productivity ($\beta = 0.349; p < 0.05$), indicating partial support for $H2$.

The third hypothesis examined the impact of friendship on productivity. Model 3 was also significant ($F = 3.976; p < 0.01$), explaining 46.9 percent of productivity variance. However, adding this friendship variable did not contribute a significant value to explaining the variance of productivity. Furthermore, the regression coefficient for all three indicators of friendship was not significant and therefore this result failed to provide evidence to support $H3$.

Finally, the fourth hypothesis presumed that there was a difference of strength in the relationship between privacy and productivity among generations. Based on the results of the previous study, which indicated that the gap only happened between Generation X and Y (and did not occur between Generation X and the Baby Boomers), this study split the data into two working generations, namely, the senior generation (Baby Boomers and Generation X, or those who are 30 years old or more) and the younger generation (Generation Y or those who are less than 30 years old). The hierarchical regression analysis showed that Model 4 was significant ($F = 5.030; p < 0.01$) and it was able to explain 57.5 percent of productivity variance. The interaction between privacy and productivity contributed an additional 10.6 percent as compared to the previous model (significant at $p$-value 0.05). As expected, this moderating variable was significant ($\beta = -0.496; p < 0.01$), which indicated that the strength between the capability to control interruption or distraction (privacy) and productivity was stronger for the younger generation (Gen Y) if compared to that of the senior generation (Baby Boomers and Gen X). This result provided a support for $H4$.

The difference between privacy and productivity for the younger generation and the senior generation can be seen in Figure 2. The results of this study will be elaborated further in the next section.

<table>
<thead>
<tr>
<th>DV: productivity</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy</td>
<td>0.346**</td>
<td>0.236*</td>
<td>0.139</td>
<td>0.956**</td>
</tr>
<tr>
<td>Collaboration</td>
<td>-0.075</td>
<td>0.628</td>
<td>-0.106</td>
<td>0.556</td>
</tr>
<tr>
<td>Knowledge transfer</td>
<td>0.349*</td>
<td>0.354*</td>
<td>0.297</td>
<td>0.051</td>
</tr>
<tr>
<td>Ideation and creative development</td>
<td>0.956**</td>
<td>0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutual trust</td>
<td>0.204</td>
<td>0.204</td>
<td>0.116</td>
<td>0.435</td>
</tr>
<tr>
<td>Commitment</td>
<td>0.045</td>
<td>0.803</td>
<td>0.022</td>
<td>0.891</td>
</tr>
<tr>
<td>Shared interests and values</td>
<td>-0.139</td>
<td>0.393</td>
<td>0.019</td>
<td>0.907</td>
</tr>
<tr>
<td>Privacy $\times$ generation</td>
<td>-0.496**</td>
<td>0.017</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III. Results of hierarchical regression analysis

Notes: $\beta$ = unstandardized coefficient. *$p < 0.05$; **$p < 0.01$
5. Discussion and implications

5.1 Theoretical implications

A flexible office was applied at INST to replace the previous office design in the pursuit of efficiency. This flex office replaced the former cellular-like design and physically resulted in a more open, non-territorial office. As quoted from A, one of the directors of INST and the initiator of the policy, a flex office was expected to lower office costs. This was justifiable since the flex office utilizes a smaller workspace and hence increases the capacity of employees’ workstations. Besides, without a dedicated desk, where every employee is allowed to use any unoccupied desks, the desk utilization would increase.

Aside from the efficiency aspect, as explained by A, a change of mindset was another target of the flex-office policy. The old mindset, namely, that each employee must “own” a workstation, had to be replaced by a new perspective: that every employee could work anywhere within the INST campus. This new mindset enables the employee to work flexibly. This change of culture is one of his concerns:

I hope that a culture of ‘shame’ will raise, i.e., employees will feel ashamed when they stay at their desk[s] for the whole day because that means they do not earn anything to achieve their target performance [that is, doing a consulting project or teaching] (A).

Another objective of the flex-office policy was the culture of cleanliness, or the 5-S culture (after Seiri, Seiton, Seiso, Seiketsu and Shitsuke in the Japanese culture), as all employees are required to clear their belongings before leaving to maintain the desk clean. Consequently, personalization of any workstations is not allowed.

Van der Voordt (2003) stated that generally, the objectives of implementing flexible working spaces are: to improve effectiveness, including productivity; to increase pleasure at work to enhance employee satisfaction; to promote a positive image of modern and professional organization; to attract and maintain “rare” human resources and customers; to develop flexibility; and for cost efficiency. These objectives are aligned with the goals expected by INST management.

The open-office policy, nevertheless, has its own drawbacks. Studies have shown that employees’ internal motivation and satisfaction toward their job and colleagues decrease after an office redesign, particularly from a cellular office into an open office. Employees complained about having less identity at work (Brunia and Hartjes-Gosselink, 2009), as well...
as about having fewer opportunities to build friendship (Oldham and Brass, 1979; Berman et al., 2002). These facts became realities at INST.

INST employees felt that their privacy, degree of collaboration and level of friendship significantly changed in the new office design. One of the informants, B, experienced less privacy when the new design was implemented. The level of privacy was reduced since the new workstations were more open and desks were set practically side by side (as compared to the previous office, where high-partitioned cubicles were installed to surround every workstation). Now, people can see each other and see what others are doing and thus are easily distracted. B felt that she could not fully concentrate on her work since there were many disruptions and distractions. “I need a place that is my place,” B said. Along the same line, R, one of the FGD participants, shared that the open office hindered his creative works due to the noise. T, another FGD participant, raised the same concern:
D, on the other hand, had a different view. She felt happy with the flex-office design since she felt she had more colleagues and friends in the workplace. The flex office encouraged her to be more open-minded about her environment: “[Previously] I felt restrained and stressful sitting inside the cubicle.”

Furthermore, in this study, there is a contrast between young employees and their seniors. Even though the open office decreases attention (Roelofsen, 2008; Davis et al., 2011; Brunia et al., 2016), younger Gen Y employees are more capable of controlling frequent distractions or interruptions related to the current workstation design. Gen Y is used to working in energetic and dynamic environments (Kasalová et al., 2015) so that a physical environment as active and noisy as an open office would not automatically decrease their capacity to control work interruptions and disruption; they still feel comfortable and are able to focus compared to their seniors. Based on the interviews with the informants who were young employees, they were able to adjust to the situation and they could continue working at their workstation using earphones to block the noise. Indeed, as a previous study found, Gen Y considers that the open office offers benefits that compensate for the privacy and noise issues (Rasila and Rothe, 2012). This attitude was different from that of senior employees who were not used to working in such conditions. The seniors felt uncomfortable and could not work well.

Collaboration is a positive change felt by the employees and it offers a real impact toward work productivity, particularly collaboration in developing ideas and creativity. By implementing the open office, employees are eager to develop ideas with the colleagues sitting next to them. Finally, they were able to increase their work productivity, as mentioned by C:

The positive side is there, so we are able to get to know our colleagues from other different functional teams [therefore it is easier to have collaboration].

The results of the study indicate that positive collaboration occurred partially, which is different from the initial hypothesis: knowledge transfer does not have a positive impact on productivity. However, this is understood when the statements of one informant, C, are explored:

In my opinion, there is no change, the knowledge transfer happens when we would like others to know […] it happened before [i.e. before the flex office policy]! Surely it was OK […].

Therefore, the intentions and frequency for sharing knowledge with colleagues around the work area, the level of shared knowledge and the collaboration among colleagues do not directly enhance the employees’ work productivity. In other words, employee productivity is not influenced by intentions, frequency and level of shared knowledge among the employees, but is affected by employees’ ideation and creative development. As mentioned in previous studies, the open-plan office facilitates communication and interaction among employee members. Interaction and communication support new ideas and enhance employee creativity so they contribute to work productivity.

Friendship is one aspect that is changing after the open office was implemented. However, even though employees felt changes because of the policy, the change of friendship does not influence work productivity. This result is interesting because even though one informant stated that she lost her best friend because of the open office, less friendship is not a barrier to achieving work performance. In other words, despite having fewer close friends as a result of removing the cubicles (Oldham and Brass, 1979), at the same time, an open office creates opportunities to have more friends, provides more resources to support work performance and decreases barriers in achieving success (Berman et al., 2002; Leadon, 2015).

5.2 Managerial implications
The results of the study indicate that collaboration and privacy have a positive impact on work productivity. Referring to the results, managers could optimize the implementation of
an open-office policy by omitting barriers. On the one hand, an open office responds to companies’ needs for efficiency, but on the other hand, employees’ satisfaction and productivity could become main concerns for the companies. Based on the results of the study, managers need to ensure that they provide conducive environments for their employees so that collaborations may flourish.

Managers must understand how different generations navigate the relationship between privacy and productivity. Managers must be capable of detecting and accommodating the different needs and requirements of those generations. The ability to control interruptions or distractions is stronger for young employees (Gen Y) than for their seniors (Baby Boomers and Gen X), so an open office could be a better platform for the younger generation to increase collaboration and coordination at work. In knowledge-based organizations, maintaining concentration is as important as collaboration and communication among the employees. This is a challenge for an open office: managers must be able to bridge the need for working individually with the need to work as a team.

In the future, as the number of Gen Y employees grows, management can implement flexible office designs so that efficiency will increase. Because Gen Y values freedom, mobility and flexibility, providing special rooms to accommodate activities that require high levels of concentration and privacy is no longer necessary. Employees will carry out their own work activities independently as well as in teams in the same work area; therefore, organizations can achieve a higher level of office space utilization and reduce associated costs without having to sacrifice the productivity of all employees.

6. Conclusion and further study

6.1 Conclusion

The study focuses on changes in office design from a combi, cellular-like office to a more open, non-territorial office. The results of the qualitative study at the initial stage show that three main variables are the most significant impacts of the office redesign: privacy, collaboration and friendship. Regarding the privacy variable, this study uses Sundstrom et al.’s (1980) definition of “psychological privacy.” Privacy and collaboration have a positive impact on work productivity; however, the study does not obtain adequate evidence to show the significant impact of friendship. Furthermore, the impact of privacy on productivity is greater for Generation Y than for the Baby Boomers and Generation X. In other words, the results of this study further reinforce the results of previous research that illustrates the differences in the impact of privacy on productivity because of differences in the generation of the workforce within an organization.

The results of the study enrich the operations management discipline, which deepens our understanding of the impact of office redesign on productivity. This study also contributes to the literature by critically investigating open-office policy from the view of socio-behavioral and physical office aspects, particularly among different employees’ generations. The research design – which combined qualitative and quantitative methods – is one of the study’s major strengths because the emerging variables are actual and comprehensive. The study’s new instrument has a valid and reliable psychometric property, which contributes to future research related to office design.

6.2 Limitations of the study and further research

Some limitations of the study can be anticipated and this will inform the scope of further research. First, the study only takes place in one organization. Even though the qualitative study conducted at the initial stage was in-depth, the results are limited to the condition of the INST organization. Further research can broaden this scope by surveying respondents of different organizations so that generalization is increased.
Second, the study focuses on the conditions of an office redesign three months after the policy was applied. On the one hand, the survey period provides advantages since the respondents directly feel the impact of new office design; on the other hand, the study period may be a transition period, and respondents may still be too sensitive to the change so that they give answers stronger or deeper than expected. Further research can combine the results through a longitudinal study which focuses on the impact for a certain period of time—for example, three months after the implementation, six months or one year after the application, or an even longer period of time.

Finally, even though the study also evaluates the impact of privacy on productivity among generations, the researcher surveyed only 36 employees (who are all members of the population). The total number of young employees (Gen Y) is only 12 people; meanwhile, senior employees (Baby Boomers and Gen X) numbered 23 individuals (one person did not mention his/her age so his/her generation cannot be identified). Further study is still needed on Gen Y, considering that young employees will be more dominant in the organization in the future.

References
Cummings, T.G. and Griggs, W.H. (1976), "Worker reactions to autonomous work groups: conditions for functioning, differential effects, and individual differences", Organization and Administrative Sciences, Vol. 7 No. 4, pp. 87-100.


Gyllenhammar, P.G. (1977), People at Work, Addison-Wesley, Reading, MA.


<table>
<thead>
<tr>
<th>No.</th>
<th>Stakeholders</th>
<th>Initial</th>
<th>Selecting criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informants</td>
<td>1 Director of human capital, the initiator of the open-office policy</td>
<td>A</td>
<td>Policymaker</td>
</tr>
<tr>
<td></td>
<td>2 Senior trainer/consultant</td>
<td>B</td>
<td>Employee, senior</td>
</tr>
<tr>
<td></td>
<td>3 Trainer/consultant, head of department</td>
<td>C</td>
<td>Employee, at managerial level</td>
</tr>
<tr>
<td></td>
<td>4 Junior trainer/consultant</td>
<td>D</td>
<td>Employee, less than 3 years working at INST</td>
</tr>
<tr>
<td></td>
<td>5 Administration staff</td>
<td>E</td>
<td>Employee, provides administrative assistants to trainers/consultants</td>
</tr>
<tr>
<td>FGD participants</td>
<td>1 Senior trainer/consultant</td>
<td>R</td>
<td>Employee, senior</td>
</tr>
<tr>
<td></td>
<td>2 Senior trainer/consultant, head of department</td>
<td>S</td>
<td>Employee, senior, at managerial level</td>
</tr>
<tr>
<td></td>
<td>3 Junior trainer/consultant</td>
<td>T</td>
<td>Employee, 3–5 years working at INST</td>
</tr>
<tr>
<td></td>
<td>4 Junior trainer/consultant</td>
<td>U</td>
<td>Employee, 3–5 years working at INST</td>
</tr>
<tr>
<td></td>
<td>5 Junior trainer/consultant, head of department</td>
<td>V</td>
<td>Employee, 3–5 years working at INST, at managerial level</td>
</tr>
<tr>
<td></td>
<td>6 Junior trainer/consultant</td>
<td>W</td>
<td>Employee, less than 3 years working at INST</td>
</tr>
<tr>
<td></td>
<td>7 Administration staff</td>
<td>X</td>
<td>Employee, provides administrative assistants to trainers/consultants</td>
</tr>
<tr>
<td></td>
<td>8 Administration staff</td>
<td>Y</td>
<td>Employee, provides administrative assistants to trainers/consultants</td>
</tr>
</tbody>
</table>

Table AII. Survey respondents (all employees experiencing the open office)

<table>
<thead>
<tr>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>17</td>
</tr>
<tr>
<td>Female</td>
<td>19</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
</tr>
<tr>
<td>&lt; 30</td>
<td>12</td>
</tr>
<tr>
<td>30– &lt; 40</td>
<td>10</td>
</tr>
<tr>
<td>40– &lt; 50</td>
<td>6</td>
</tr>
<tr>
<td>50 or more</td>
<td>7</td>
</tr>
<tr>
<td>Does not respond</td>
<td>1</td>
</tr>
<tr>
<td>Tenure (years)</td>
<td></td>
</tr>
<tr>
<td>&lt; 5</td>
<td>22</td>
</tr>
<tr>
<td>5– &lt; 10</td>
<td>2</td>
</tr>
<tr>
<td>10– &lt; 20</td>
<td>3</td>
</tr>
<tr>
<td>20 or more</td>
<td>8</td>
</tr>
<tr>
<td>Does not respond</td>
<td>1</td>
</tr>
<tr>
<td>Previous working experience</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>30</td>
</tr>
<tr>
<td>No</td>
<td>6</td>
</tr>
</tbody>
</table>

Table AIII. Office design in previous office (those who have working experience)

<table>
<thead>
<tr>
<th>Survey respondents</th>
<th>Office design in previous office (those who have working experience)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular office</td>
<td>10</td>
<td>34.5</td>
</tr>
<tr>
<td>Open office, dedicated desk</td>
<td>14</td>
<td>48.3</td>
</tr>
<tr>
<td>Open office, flexible desk</td>
<td>1</td>
<td>3.4</td>
</tr>
<tr>
<td>Others</td>
<td>4</td>
<td>13.8</td>
</tr>
</tbody>
</table>
Appendix 2. Interview questions – qualitative stage

(1) The background of the informant at INST (working duration, position, age, area of expertise, main job at INST).

(2) The changes felt after the enactment of the open office.

(3) The observed changes from the working environment after the enactment of the open office.

(4) The most significant changes that are perceived or observed.

(5) The hopes for successful implementation of the open office.
### Appendix 3

#### Productivity
- **Perf1**: The ease of fulfillment of earning-load (difficult – easy)
- **Perf2**: The quality of work (poor – good)
- **Time1**: The completion of daily To-Do List as planned (not completed – completed)
- **Time2**: The speed of completion of each job (on average) (slow – fast)
- **Time3**: The availability of time to complete each job (on average) (low – high)

#### Privacy
- **VDI1**: For me, conversations with coworkers in my surrounding (disturbing – not disturbing)
- **VDI2**: For me, coworkers who suddenly come (an unplanned visit) (disturbing – not disturbing)
- **VDI3**: For me, a sudden invitation to chat from a coworker next to me (disturbing – not disturbing)
- **SP1**: For me, having a private conversation at work (not comfortable – comfortable)
- **SP2**: For me, having a discussion about work (not comfortable – comfortable)
- **SP3**: For me, singing at work just to build a relaxed atmosphere (as mood booster) (not comfortable – comfortable)
- **COA3**: Frequency of coworkers came over while I was working (low – high)

#### Collaboration

**Knowledge transfer**
- **KT1**: For me, the desire to share knowledge with coworkers next to (or near) me (low – high)
- **KT2**: For me, the frequency of knowledge sharing with coworkers next to (or near) me (low – high)
- **KT3**: For me, the contribution of the knowledge-sharing process to the completion of my work (low – high)
- **KT4**: The level of knowledge I share with coworkers next to (or near) me (limited – deep)
- **CW1**: Teamwork with coworkers next to (or near) me (difficult – easy)

**Ideation and creative development**
- **ICD1**: For me, the desire to develop ideas and creativity with coworkers next to (or near) me (low – high)
- **ICD2**: The quality of new ideas formed from the interaction process with coworkers next to (or near) me (poor – good)
- **ICD3**: The number of new ideas formed from the interaction process with coworkers next to (or near) me (null – high)
- **ICD4**: The implementation of the ideas and creativity that I gained from coworkers next to (or near) me (conceptual – operational)
- **ICD5**: The contribution of idea implementation that I gained from coworkers next to (or near) me (limited to unit – organizational)

#### Friendship

**Mutual trust**
- **MT1**: For me, the level of mutual trust with coworkers next to (or near) me (weak – strong)
- **MT2**: For me, the depth level of personal stuffs shared with coworkers next to (or near) me (limited – deep)
- **MT3**: For me, the desire to share about personal matters with coworkers next to (or near) me (low – high)
- **MT4**: For me, the frequency of sharing personal matters with coworkers next to (or near) me (low – high)

**Commitment**
- **CO1**: For me, the desire to help each other with coworkers next to (or near) me (low – high)
- **CO2**: For me, the frequency of helping each other with my coworkers next to (or near) me (low – high)
- **CO3**: For me, the level of fulfillment of what has been agreed (things outside of work) with coworkers next to (or near) me (low – high)

**Shared Interests and values**
- **SIV1**: For me, discussing or talking about out-of-work things with coworkers next to (or near) me (boring – interesting)
- **SIV2**: For me, knowing the ins and outs of my coworker’s private life next to (or near) me (boring – interesting)
- **SIV3**: For me, knowing the point of view or values of life of coworkers next to (or near) me (boring – interesting)
- **SIV4**: For me, the frequency to share or discuss about the perspective or values of life with coworkers next to (or near) me (low – high)
- **SIV5**: My knowledge of my coworker’s private life next to (or near) me (limited – deep)

---

**Table AIII.**

Survey items (score is between −3 and 3)
About the authors
Erlinda N. Yunus is Assistant Professor of Operations Management at the PPM School of Management. She received PhD degree in Supply Chain Management from Southern Illinois University, Carbondale, IL, USA. She has published her research in many journals, particularly in the areas of supply chain management, inter-firm collaboration and innovation. Erlinda N. Yunus is the corresponding author and can be contacted at: erl@ppm-manajemen.ac.id

Erni Ernawati is Lecturer of Operations Management at the PPM School of Management. She is studying Doctoral degree at Bogor Agricultural University, Bogor, Indonesia, in the field of Supply Chain Management.

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Can artificial neural networks predict lawyers’ performance rankings?

Susana Almeida Lopes and Maria Eduarda Duarte
Faculty of Psychology, University of Lisbon, Lisbon, Portugal, and
João Almeida Lopes
Department of Pharmaceutics and Pharmaceutical Technology,
Faculty of Pharmacy, University of Lisbon, Lisbon, Portugal

Abstract

Purpose – The purpose of this paper is to propose a predictive model that could replace lawyers’ annual performance rankings and inform talent management (TM) in law firms.

Design/methodology/approach – Eight years of performance rankings of a sample of 140 lawyers from one law firm are used. Artificial neural networks (ANNs) are used to model and simulate performance rankings over time. Multivariate regression analysis is used to compare with the non-linear networks.

Findings – With a lag of one year, performance ranking changes are predicted by the networks with an accuracy of 71 percent, over performing regression analysis by 15 percent. With a lag of two years, accuracy is reduced by 4 percent.

Research limitations/implications – This study contributes to the literature of TM in law firms and to predictive research. Generalizability would require replication with broader samples.

Practical implications – Neural networks enable extended intervals for performance rankings. Reducing the time and effort spent benefits partners and lawyers alike, who can instead devote time to in-depth feedback. Strategic planning, early identification of the most talented and avenues for tailored careers become open.

Originality/value – This study pioneers the use of ANNs in law firm TM. The method surpasses traditional static study of performance through its use of non-linear simulation and prediction modeling.

Keywords Knowledge workers, Talent management, Performance appraisal, Neural networks, Career, Law firm

Paper type Research paper

1. Introduction

The new millennium brought to law firms an unusual drop in demand, lower rates and weaker client loyalty (Muir et al., 2004; Stumpf, 2007). Law firms, which largely employ knowledge workers (Drucker, 1959), have identified talent as a differentiator during recessions (Cappelli, 2000; Gallardo-Gallardo et al., 2013; Ready et al., 2010). Talent management (TM) is used to attract, retain and develop the most talented lawyers, who are believed to add value to a firm’s performance, creating competitive advantage and by depleting the reserves of competitors (Felin and Hesterly, 2007; Gallardo-Gallardo et al., 2013).

In law firms, talent is equated with high performance (Brittain, 2005). The most talented lawyers are those who rank at the top of their firms in terms of performance. Thus, lawyers who exceptionally outperform their peers are frequently called the best in their class, the A Players, and the most talented (e.g. Ready et al., 2010; Silzer and Dowell, 2010; Smart, 2005; Ulrich and Smallwood, 2012).

Such an approach to talent, which seeks to identify high performers for career advancement, fits the career model for law firms. In a tournament system that has its origins in the mid-nineteenth century in the USA, from admission into a firm to attaining partnership, each lawyer’s performance is ranked annually against peers (Pinnington, 2011). High-performing lawyers are entitled to career advancement, and average and low-performing lawyers remain at the same professional level or are advised to leave the firm.
Recent evidence supports the use of performance rankings to build a high-performing workforce, encouraging low performers to improve or to leave and releasing employees’ best self to support the organization’s success (Högland, 2012; Netessine and Yakubovich, 2012). Notwithstanding, the annual ranking of the performance of a firm’s lawyers is a highly bureaucratic endeavor that both partners and lawyers must undergo. Reliable prediction of the most talented would be of great value for practitioners, but this possibility remains a mirage. Prediction of performance rankings could allow more time to elapse between performance rankings. The time freed from this greatly disliked practice (Aguinis et al., 2012) could be transferred to talent development and supporting career advancement.

This paper aims to fill this gap by proposing a way to predict performance rankings. Either statistical or intelligent methods can be used for prediction. In this paper, we propose the use of artificial neural networks (ANNs), an intelligent method of prediction that can, under certain conditions, outperform statistical methods.

The paper is organized as follows. First, it overviews TM in law firms, performance appraisal and competency frameworks. Second, it examines the potential use of ANNs to address performance rankings. Third, an ANN is proposed that was trained on eight years of lawyers’ performance rankings to forecast performance rankings. For comparison purposes, a multivariate regression analysis is performed using the same data. Finally, the potential use of ANNs for TM in law firms is discussed.

2. Background
The economic slowdown in the first decade of the new millennium exposed firms to the experience of a drop in demand, followed by decreases in revenues and profits (Muir et al., 2004). Even many highly prestigious firms have laid off workers in the USA, UK and other European countries. Several firms merged, whereas others closed entirely. The annual double-figure increases in revenue and profits resulting from annual fee increases are no longer. Clients, who are also facing tumultuous times, are managing smaller budgets and are pressuring providers, including law firms, for innovative service at lower fees.

The career model of the law firm was built around a hierarchical pyramid of partners, associates and trainees. It implied a continuous annual growth in profits, as well as in numbers of lawyers, for continuing career advancement and the creation of new partners each year (Galanter and Palay, 1990, 1994). Although the number of lawyers reaching partnership had always been few, the highest-performing lawyers could nevertheless expect this ultimate happy ending. This assumption of a reliable career path is now being defied. Although high-performing lawyers may not have a strong enough business case for making partner, their firms may also not be able to afford the loss of their talent (Mottershead, 2010).

Such challenges require a solution from law firms. The acknowledgment of the disproportionate contribution of the most talented to a firm’s performance (Felin and Hesterly, 2007; Gallardo-Gallardo et al., 2013) has embedded TM in the legal profession. Managing up-and-coming talent has been found to be vital in the support of firms’ adaptation to the new normal (Davis, 2009), and talent is listed among the most critical assets in the sustainability of a firm (Boudreau and Ramstad, 2005). Large and/or international law firms in the USA and UK, drivers of change in the legal market, have created or changed their human resource (HR) department designations for TM (Mottershead, 2010), sparking a new trend. The identification, development and retention of the most talented lawyers have become a priority.

2.1 Performance appraisal in law firms
Appraisals are the cornerstone of TM. They enable differentiation of talent and support up the career model. Since the mid-nineteenth century, lawyers have been ranked against their peers during the whole course of their work in the firm, from admission to partnership
Partners annually appraise lawyers in the same cohort (i.e. those with the same qualification year). The highest-performing lawyers advance to the next rung on the career ladder and receive the largest bonuses, while average and low-performing lawyers remain at the same professional level or are invited to leave, their replacements being vetted through the recruitment of trainees from elite law schools (Welch and Welch, 2005). This tournament was first introduced by the American law firm Cravath to create a continuous renewal of the workforce, supporting excellence and meritocratic policies.

Most firms have followed a TM approach that is focused on identifying and retaining high-performing lawyers. Accordingly, different percentages of the workforce are expected to perform at different levels (Welch and Welch, 2005). The professionalized HR departments have introduced additional and sophisticated practices to differentiate lawyers according to performance, using appraisal systems based on relative comparison, such as forced-distribution ranking systems. The distributions used are frequently adapted from the Gaussian curve, and lawyers are ranked into performance levels (e.g. 20 percent–70 percent–10 percent), talent levels, or the well-known designation of A, B and C players (Collins, 2001; Guest et al., 2004; Ready et al., 2010; Welch and Welch, 2005). Like the original tournament, force-ranking systems are instituted to improve the potential of the workforce (Scullen et al., 2005).

Ranking systems are considered to have greater validity than other appraisal methods (e.g. Balzer and Sulsky, 1992; Chattopadhayay and Ghosh, 2012; Goffin et al., 2009; Heneman, 1986; Nathan and Alexander, 1988; Wagner and Goffin, 1997). They are acknowledged to offer better prospects for differentiating individuals’ performance, avoiding the frequent rating biases that prevent marked differentiation between individuals’ performance, such as leniency (a tendency to over-evaluate performance, first described by Ford in 1931) and the halo effect (described by Thorndike in 1920, which reflects exaggerated correlations among ratings of disparate criteria) (McBriarty, 1988; Stewart and Nandkeolyar, 2006). Rankings find support in natural social-comparison processes (Wagner and Goffin, 1997) that underpin decision-making processes, such as appraisals. Scullion et al. (2000) confirmed that it is easier for managers to identify the contribution of each individual through comparison with peers.

Of course, ranking systems are not without their critics. On the contrary, they are consistently the target of fierce criticism, which, for instance, alleges discouragement of collaboration and communication as unintended consequences (Pfeffer and Sutton, 2006). The over evaluation of average performers integrated in low-performing teams and the under-evaluation of high performers integrated in high-performing teams have also been reported (O’Boyle and Aguinis, 2012).

Criticism is not exclusively brought to bear at ranking systems. Performance appraisal is both the most widely used (Guest et al., 2004) and the most disliked TM practice (Aguinis et al., 2012). Its cost in effort and wasted time, as well as the negative impact on team cohesion, are commonly criticized (Lawler et al., 2012). Most individuals believe that they perform above average (Sharot et al., 2011), so rankings lead to dissatisfaction and feelings of injustice among those ranked at average and lower levels, as well as to difficult conversations with partners, who would prefer to avoid managing negative impacts on interpersonal relationships (Bol, 2011).

Performance appraisal is, however, a powerful TM practice (Chattopadhayay and Ghosh, 2012; Judges and Ferris, 1993; Murphy and Cleveland, 1995), which is required to effectively manage talent (Lawler et al., 2012). It is the annual rite of appraising that “triggers dread and apprehension in the most experienced, battle-hardened manager” (Roberts and Pregitzer, 2007, p. 15). In response of criticism of annual rankings, Allen & Overy, Hogan Lovells and Slaughter and May, which are three of the largest law firms in London and role models for other firms, recently announced the replacement of yearly rankings with other forms of feedback (Simmons, 2017). The time and effort spent on making rankings to identify talent may have been leaving scarce energy for talent development.
2.2 Competency frameworks in law firms

To support TM, law firms have implemented competency frameworks that are drawn from high-performers’ profiles (Mottershead, 2010; Polden, 2012), which describe the range of skills required for career success. Hard (i.e. legal) and soft skills (i.e. managerial) are given behavioral descriptions. Legal knowledge is a primary requisite (Bock and Berman, 2011; Polden, 2012) at the base of such hard skills as oral advocacy, drafting, analysis and problem solving. Lawyers are also required to display soft skills related to teamwork, efficiency management, client relationships, business development and coaching (Bock and Berman, 2011; Mottershead, 2010; Polden, 2012; Stumpf, 2007).

Competency frameworks form criteria for TM, from recruitment to appraisals. The development of lawyers requires in-depth feedback, taking all the skills of the competency framework into full consideration. On the contrary, performance rankings result from a comparison between lawyers’ overall performance (Lopes, 2016).

3. Performance rankings prediction

Prediction is a critical form of knowledge about indeterminate or anticipated events. It is essential for making decisions in the present that will have impact in the future (Jantan et al., 2009). Within organizations, performance prediction is vital for forecasting purposes and central to forming TM strategies (Cascio and Aguinis, 2011; Hinds et al., 2000; Mehrabad et al., 2011; Sonnentag and Frese, 2012). However, few studies exist on performance prediction. Prediction models intended to support practitioners by talent forecasting have thus far been a mirage, and law firms are no exception. Law firms attempt to identify the most talented trainees directly out of law school. Recruitment is supported by assessment tools, such as ability tests and personality questionnaires, which have demonstrated validity to predict performance over time (Tziner et al., 1993). Following their admission into a firm, however, it is a struggle to identify high performers. Time and effort are spent in creating annual rankings of lawyers according to performance differentiation, jeopardizing investments in feedback and development.

Multivariate models, such as ordinary least squares, that estimate parameters in a linear regression mode are the most common approach to prediction in TM seen in the few studies available. However, linear models fail to uncover non-linear patterns. For data that do not fit parametric assumptions (e.g. rankings of performance), noisy and missing data (in consequence, e.g., of turnover over time, which is common in law firms), and with longitudinal samples including over 100 cases, such this study, linear models are not appropriate (Klimasauskas, 1991; Scarborough and Somers, 2006). ANNs may be a suitable option for overcoming the identified constraints, as they have demonstrated superiority to regression analysis for the purposes of prediction in comparability studies (e.g. Caudill, 1991; Mehrabad et al., 2011; Noorossana et al., 2009).

ANNs have been misused in TM (Chandrasekar et al., 2015; Scarborough and Somers, 2006; Wong et al., 2000). Their rare application has been in the field of classification and for confirmatory purposes. For instance, ANNs have been successfully used to uncover
non-linear relationships between satisfaction and performance (Somers, 2001), between satisfaction and commitment (Chandrasekar et al., 2015) and between tenure and turnover (Seitz et al., 2000); which linear models failed to explain (Chandrasekar et al., 2015; Huang, 2012; Scarborough and Somers, 2006). For prediction purposes, their application has been limited. Employee selection for recruitment purposes (e.g. Mathuriya and Bansal, 2012) and turnover modeling (e.g. Sexton et al., 2005) are the most common applications found in the literature.

To our knowledge, ANNs have never addressed performance rankings, either in law firms or in other settings. Schmidt et al. (1988) identified trends in high and low performers over time, pointing the way to an avenue for research in prediction that has not yet been pursued, although 30 years have gone by. We propose an ANN for prediction of performance rankings over time to fill this gap.

3.1 Performance rankings predictors

ANNs are machine learning methods and, like the brain, learn from experience. However, they do not precisely mimic biological neural networks. They are mathematical, data-driven processes, highly dependent on the nature and quality of their data for the learning they exhibit. Therefore, in the case under discussion, ANNs used to predict performance rankings, the learning mechanism is no less highly dependent on the inputs, the predictors (Jin and Gupta, 1999).

This work proposes an ANN that learns trends from past performance rankings to predict performance rankings over time. Past performance is known to be the most powerful predictor of future performance (Sturman et al., 2005; Sturman, 2007). The first predictor incorporated into the model is past performance rankings. There are recent studies that show performance trends over time (e.g. Berrah et al., 2006; Devaraj et al., 2004; Hua Tan et al., 2004; Unahabhokha et al., 2007) and profuse literature demonstrating that performance is sufficiently stable to be predicted (Hofmann et al., 1992, 1993; Somer et al. and Fresc, 2012; Stewart and Nandkeolyar, 2006; Thoresen et al., 2004).

Performance includes both stability and dynamism (i.e. lack of stability). This is because performance is underlined by attributes that remain stable over a lifetime, such as cognitive ability and personality, and it is influenced by knowledge, experience (Schmidt et al., 1986; Sturman, 2003) and motivation (Kanfer, 1992), which shift and lead to performance dynamism. Studies have shown that the predictive validity of measures of performance decreases over time, due to the dynamic dimension of performance (Austin et al., 1989; Barrett et al., 1989; Ployhart and Hakel, 1998; Rambo et al., 1983), but the correlation over any period has been found to remain positive, pointing to a stable dimension (Ackerman, 1987; Henry and Hulin, 1987; Murphy, 1989). Considering meta-analytic results for appraisals, Sturman et al. (2005) showed performance stability over a one-year period, ranging correlations from 0.85 to 0.67. Alessandri and Borgomi (2015) also found a large degree of performance stability over a four-year period.

Knowledge, experience and motivation vary according to the phase of a lawyer’s career. Murphy’s (1989) and Kanfer and Ackerman’s (1989) models show that performance follows a steep learning curve during the learning phase. Baltes and Baltes’s (1990) theory of selection, optimization and compensation corroborates the idea that younger individuals devote more resources to their work at the beginning of their careers, then entering a maintenance phase, during which their learning curve becomes shallower. This also applies to newcomers in a law firm. Through their careers, lawyers are in either a learning, developmental or growth phase (junior lawyers and newcomers) or a maintenance phase (middle and senior lawyers, and those with greater tenure). Two variables that are related to learning phase are included in the model proposed in this work: professional level and tenure. These variables, related to chronological time, are of relevance in longitudinal
studies (Harris et al., 2006) and are frequently integrated in studies of performance (Ackerman, 1992; Farrell and McDaniel, 2001; Tesluk and Jacobs, 1998). Lawyers are sorted into professional levels according to their years of experience following passing the bar exam, it is related to experience and age. Tenure relates to experience and age because many lawyers develop their careers within the confines of one law firm. However, recently, an increasing number of lawyers have begun to make career transitions between firms.

The fourth and last variable to be learned by the ANN is billable hours. Each lawyer has an annual target of working hours to be billed to clients. Billable hours represent the most common fee arrangement in law firms. Timesheets are used to charge clients for time spent on different matters, broken down into short time intervals, with amounts per hour defined according to the seniority of the given lawyer (Campbell et al., 2012). The number of billed hours accumulated by a lawyer contributes directly to the financial performance of the firm. To incentivize billing hours, accomplishments are a frequent criterion for bonuses (Campbell et al., 2012; Mottershead, 2010). Lopes et al. (2015) found a marked positive correlation between the number of billable hours and appraisal ratings.

In this study, we use an ANN as an exploratory tool, following Scarborough and Somers’s (2006) proposal. Instead of using an ANN in a confirmatory way to confirm a linear hypothesis, the full range of possible relationships among the four imputed variables is explored. A 70 percent rate of correct prediction is fair performance by an ANN, according to the literature (Adefowoju and Ososian, 2004; Emuoyibofarhe et al., 2003; Oladokun et al., 2008). We expect our model to attain that degree of accuracy. As has been found, we expect prediction accuracy to decrease as the simulation extends over a longer period of time (Austin et al., 1989; Barrett et al., 1985, 1989; Hagan et al., 2014; Ployhart and Hakel, 1998; Rambo et al., 1983).

4. Methods
4.1 Setting and data
In 2016, data were drawn from a large Portuguese law firm. Variables for individual differences and performance rankings were collected from the administrative records of the firm. All 140 lawyers appraised between 2008 and 2015 were included in the study (Table I). In 2008, the competency framework used by the firm to conduct the appraisals was revised.
Between 2008 and 2015 the new competency framework was used for the appraisals and performance rankings. Although additional partners did join the partnership throughout the eight-year period of study, the head of each practice, who was responsible for appraisals within that practice, did not change.

4.2 Measures
4.2.1 Performance rankings. The performance rankings in this firm result from appraisals. A two-step approach is followed, first appraising and then ranking lawyers’ performance. This procedure is common in law firms and organizations in various industries that invest in TM (Welch and Welch, 2005). First, each year, including the eight years of this study, the performance of each lawyer is rated by the partners, using a competency framework, including hard skills (i.e. knowledge and solutions, communication and drafting and client orientation) and soft skills (i.e. business development, firm focus, leadership, resource management and achievement focus). The ratings for each skill are calculated by averaging sub-items using a five-point, behavioral-observation rating scale (Christ and Boice, 2009), anchored by behavior frequency. The overall appraisal ratings for each lawyer are computed by averaging all ratings of evaluated skills, in each of the eight years.

Second, the overall appraisal ratings of the lawyers are ranked according to the professional level (i.e. each lawyer’s performance was compared against peers at the same professional level: junior, middle and senior). Based on a pre-defined distribution adapted from the Gaussian curve (5 percent–25 percent–50 percent–20 percent), lawyers are placed, for each of the eight years, into four performance groups (1–4): 1 = low performance, 2 = average performance, 3 = high performance and 4 = very high performance.

4.2.2 Billable hours. Billable hours are the number of hours worked and billed to clients by each lawyer. Lawyers have an annual target for billable hours. The percentage of accomplishment of that target for each of the eight years by each lawyer is integrated into the model.

4.2.3 Professional level. The professional level ranges from junior = 1, to middle = 2 and senior = 3 levels. In the law firm, career progression is linked to both experience and performance. The number of years at each professional level varied for different individuals.

4.2.4 Organizational tenure. Tenure is calculated based on the date of admission to the firm, for each of the eight years of performance rankings.

4.3 Proposed predictive model
In this paper, an ANN termed a multilayer feedforward neural network is selected to assemble a mathematical model to predict of performance rankings. This type of ANN is characterized by a series of layers composed of nodes (or neurons). Because each layer’s nodes are closely connected with those of the neighboring layers (the equivalent of brain synapses) information propagates through the network, generating outputs (Hagan et al., 2014). This type of ANN is used to map a relation between two sets of data. One set of inputs is received and then translated into sets of corresponding outputs. Signs only flow in one direction (see Jin and Gupta, 1999).

ANNs operate in two stages: training and simulation. An ANN is trained using multiple examples (known data) of pairs of inputs/outputs (Figure 1(a)) and learns through experience. The training of an ANN is the process of adjusting its parameters to the empirical data given using a procedure and algorithm to make its predictions are as accurate as possible. This training is an iterative process that proceeds until one or more criteria are met (e.g. global error or maximum number of iterations). More details on data propagation and the training of feedforward ANNs can be found in Beale et al. (1992). After the end of the training process, the network can make predictions. This is known as
simulation mode. Our simulation begins from a known situation (Year 0). The performance ranking prediction for Year 1 is then used to estimate the Year 2 performance ranking, and so on. Simulation works in a closed-loop fashion (Figure 1(b)).

The application of an ANN to the prediction of lawyers’ performance rankings using past performance rankings, professional level, tenure and billable hours requires a feedforward autoregressive non-linear dynamic process with exogenous inputs network. The autoregressive aspect of this process means that the prediction of future performance rankings depends on actual performance ranking. It must be non-linear because mapping between the inputs and performance ranking is complex and it is not possible to represent it using a combination of linear relations. It is dynamic because it represents a process that evolves over time. Finally, it accepts data from exogenous sources (in this case, professional level, tenure and billable hours) (Gupta et al., 2004; Pearlmutter, 1990). The ANN must be supervised, meaning that the output should compare with the known correct values during training (Scarborough and Somers, 2006). A schematic representation is provided in Figure 1.

Our ANN combines professional level, tenure and billable hours with the most recent performance ranking and then predicts the subsequent performance ranking. It considers the initial performance ranking (1–4) as an input and performance ranking at each of the lag times as output. The occurrences for all lag times for all eight years of performance rankings are input. Network inputs are delayed to perform simulations with multiple time lags. For instance, different time periods (from one to six years) are taken into account when the ANN is built. The rarity of occurrences for the seven-year lag time precludes the performance of the simulation for that period. Because there are fewer occurrences and longer time, prediction accuracy is expected to decrease as the simulation extends over longer periods of time (Hagan et al., 2014). Maximal accuracy is expected for the one-year lag time, and 496 occurrences over the eight years of performance rankings are considered.

The model is calibrated using data for approximately 70 percent of the lawyers in the database, and it is independently tested on the remaining 30 percent. An early-stopping
training procedure is implemented to avoid overfitting. The networks architecture is standard three layer, with one input and one hidden (i.e. hyperbolic tangent function) and one output layer (i.e. linear function), according to Caudill’s (1991) and Klimasauskas’s (1991) recommendation. Both authors found that most problems can be solved with ANNs using three layers. The number of nodes in the hidden layer is optimized, using cross-validation. All calculations are performed with the neural network toolbox for MATLAB, version 8.6 (Beale et al., 1992).

ANNs and multivariate models, which estimate parameters in a linear regression mode, can both be used for predictive purposes. Following a commonly reported practice (Somers, 2001) and a reviewer’s suggestion, we compare the results obtained from the ANN with results obtained from a multiple linear regression analysis (MLR), which is the most common approach to prediction in the field of HR. Here, a linear autoregressive with exogenous inputs is used to compare with a non-linear ANN. The MLR model is built with single inputs (no interactions) for comparison purposes. The same variables (as used for the ANN) are input into the model: performance ranking in Year K, professional level, tenure and billable hours, and the output was change in rankings (ranking [Year K+dK]− ranking [Year K]), to mimic the ANN inputs. Delay dK varies between a one- and six-year lag time. For each data set (for each dK) cases are randomly split 70 percent/30 percent for training and testing. The model coefficients are estimated by ordinary least squares using the training set. The calibrated model is applied to the testing set. The presented results correspond to the predictions of the testing set.

Because performance comprises stability and dynamism, two testing conditions (both for ANNs and MLR) were set: one including all conditions (performance ranking maintenance and change, from one year to the next), and a dynamic condition including only performance ranking change from one year to the next).

5. Results
Prediction results for the ANN and the MLR are presented in Table II. Loss of predictive accuracy is always expected between training and testing (Scarborough and Somers, 2006), but for the networks developed in this work the loss is non-significant.

For a lag time of one year the ANN yields a 71 percent rate of correct predictions in conditions where performance ranking change, and it predicts about three out of four cases correctly (73 percent) when all conditions are included. With a two-year lag time, the prediction levels were 67 and 69 percent, respectively. Thus, the model was found optimal, as attested in the literature (Oladokun et al., 2008). The accuracy of our network exceeds,

<table>
<thead>
<tr>
<th>Prediction lag time (years)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input variable ID</td>
<td>Variable N occurrences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Tenure</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>−</td>
</tr>
<tr>
<td>2</td>
<td>Professional level</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>−</td>
</tr>
<tr>
<td>3</td>
<td>Billable hours</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>Performance ranking (year 0)</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>ANN Correct predictions (training/validation) (%)</td>
<td>74</td>
<td>71</td>
<td>66</td>
<td>65</td>
<td>64</td>
<td>58</td>
</tr>
<tr>
<td>ANN correct predictions all conditions (simulation) (%)</td>
<td>73</td>
<td>69</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>ANN correct predictions when ranking changes (simulation) (%)</td>
<td>71</td>
<td>67</td>
<td>62</td>
<td>63</td>
<td>63</td>
<td>57</td>
</tr>
<tr>
<td>MLR correct predictions all conditions (%)</td>
<td>72</td>
<td>66</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>MLR correct predictions when ranking changes (%)</td>
<td>56</td>
<td>41</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
</tbody>
</table>

**Notes:** Categorical symbols were assigned to the observed sensitivity, 0 poor relevance to +++ highly relevance; − non-relevant
for example, Chandrasekar et al.’s (2015) 64.1 percent network accuracy (in the context of classifying employees according to satisfaction).

As expected, the accuracy level decreases over time. With a three-year lag time, the prediction level of the network decreases to about 62 percent when ranking changes and to 56 percent when all conditions are considered, which is a non-significant result, similar to random chance (for which one can expect a 50 percent accuracy).

The initial performance rankings is the most relevant input for the network at all lag times, which is in line with the literature (Sturman et al., 2005; Sturman, 2007). The other three predictors, in terms of decreasing importance, are billable hours, tenure and professional level.

For comparative purposes, an MLR was performed. For all conditions, the difference of results from that of the ANN is non-significant. But under the dynamic condition the MLR is unable to find an accurate prediction. With a one-year lag time, for instance, the results do not significantly outperform random prediction (56 percent).

6. Discussion

Law firms emphasize TM to ensure long-term sustainability, but this goal remains far from being achieved. More sophisticated practices are in use, such as assessment tools for recruitment purposes, competency frameworks and forced-distribution ranking systems, but the annual rite of ranking lawyers to identify high performers and support career decision making persists. Appraisal and rankings, which are the least popular TM processes (Aguinis et al., 2012), are repeated year after year, monopolizing time and effort and, sometimes, damaging good relationships between partners and lawyers.

In line with the law firm career model, the priority of identifying the most talented lawyers is dominant. The performance ranking of lawyers is unduly valued, over and above their ratings in each of the skills of the competency framework. Therefore, feedback that could support lawyers preparing for their increasingly daunting tasks in the new economic reality may be precluded. According to Nelson (1981), partners are the firm finders (business developers), minders (managers) and grinders (producers). Their time is scarce. Prediction could allow for greater spare time to be reinvested in providing feedback and supporting career development.

This study pioneers performance prediction for TM in law firms. Prediction is commonplace in engineering but infrequent in HR field (Jantan et al., 2009). Predictive models are much less accurate in the behavioral science than they are in engineering because human behavior is difficult to measure reliably (Scarborough and Somers, 2006). This may underline the lack of attention that has come from the field and the lack of predictive models to support practitioners. We intend to fill the gap by proposing an ANN that learns from input data (namely, initial performance ranking, billable hours, professional level and tenure) and predicts performance rankings over time.

ANNs are a model of choice when parametric assumptions are not met, when noisy data exist, or longitudinal data with over than 100 cases form the set, as is the case in this work (Klimasauskas, 1991). In fact, ANNs overcome the limitations of MLR and generate more accurate results (see Collins and Clark, 1993; Sharda and Patil, 1992; Somers, 2001, for reviews). MLR were able to predict when all conditions (performance ranking maintenance and change) were considered, but did not overcome a random prediction in the condition of ranking change. This may result from more accuracy predicting stability over time that future studies should verify.

The relevance of tenure and professional level to the model supports a likely different trend of performance during the learning phase, as suggested by the learning theories of Murphy (1989), and Kanfer and Ackerman (1989). The motivation to thrive early in one’s career (Baltes and Baltes, 1990) might also play a role in the improvement of performance.
among junior lawyers. Billable hours outweigh even professional level and tenure in importance for performance ranking predictions. Increases in performance ranking were shown to be linked with increases in billable hours, and the reverse was also true. This evidence is in accordance with Lopes et al.’s (2015) findings, which were related to a strong correlation between the number of billable hours and appraisal ratings. High-performing lawyers produce more billable hours because clients and partners solicit them more often.

Implications for implementation in law firms can be drawn. We suggest that predictive models, such as the model proposed in this paper, could have prevented situations such as the complete abandonment of performance rankings by Allen & Overy, Hogan Lovells and Slaughter and May. This is a recent trend, initiated by some audit and consulting firms that do not stress rankings for the identification of the most talented individuals, although this is required to thrive in fast-changing markets (Ashton and Morton, 2005; Buckingham and Vosburgh, 2001; Dries, 2009; Sengupta, 2012).

We argue that appraisals and performance rankings should continue in law firms. First, performance rankings are linked to a meritocratic system that sends a message of quality to the clients and the market, creating a competence allure (Greenwood, 2003). Second, rankings support the generally accepted career model, which has been successful in driving profitability. Competition among lawyers to reach the top boosts their motivation to produce large numbers of billable hours without the necessity of complex managerial and control processes (Galanter and Palay, 1990, 1994). Third, this widely applied HR practice (Guest et al., 2004) enables the differentiation of high, average and low performers, which is critical for managing talent. Fourth, ranking systems have increasingly demonstrated their greater validity than other appraisal methods for differentiation purposes (Chattopadhayay and Ghosh, 2012). Fifth, communicating rankings to lawyers helps to create a culture of transparency, as benchmarking become possible. Sixth and last, the model adjustment allows for the prediction of up to two-year lag times. Thus, regular rankings are still needed.

But for two years, instead of spending time completing predictable performance rankings, partners might benefit from additional time spent developing lawyers. The promotion of lawyers’ awareness, through feedback in relation of each appraisal criterion, as well as preparing lawyers for additional challenges along their career paths, is of much greater importance than ranking performance annually.

An important application of an ANN would be in the field of strategic planning. HR practitioners could benefit from forecasting talent to better allocate resources. The earlier signposting of lawyers whose performance rankings are likely to change would allow a closer career examination, for example.

One final application rests in the possibility to identify the most talented, highest-performing lawyers early in their careers. This major possibility has the potential to influence the career model. A talented lawyer can be identified early and might benefit from support for development, including skills required for future managerial roles.

6.1 Limitations and future research
A first limitation stems from the data set, which was gathered exclusively from only one firm, which precludes any generalization of results because of common method variance, which influences contextual factors in measures that cause systematic covariation (Podsakoff et al., 2003). Thus, replication study with additional firms is necessary for any generalization to be meaningful. Broader samples are also required to address the second limitation: high turnover in the sample. Turnover is greatest for knowledge workers, such as lawyers, among all types of professions (Somaya and Williamson, 2008), causing range restrictions for the analysis of performance over time (Goodman and Blum, 1996; Schmidt and Hunter, 2004; Sturman and Trevor, 2001). For this reason, correlations among variables might be reduced in our results (Sackett and Yang, 2003; Schmidt and Hunter, 2004).
Analysis is conducted for all occurrences at lag times, which allows for overcoming the frequent limitations of the analysis of only cases with complete data. However, as lag times increase, the occurrences that fed the network decrease, from 496 when the lag was 1 year to 32 when it was 7, precluding, for instance, the analysis of the final lag. Network training becomes progressively less effective and errors increase, not just because prediction was more difficult over a longer horizon but because there are fewer occurrences to train the network.

During the period of analysis, lawyers advance in their careers. One input in the predictive model is the professional level, which revealed itself to be an important predictor. As a reviewer stressed, some validity issues are raised by career advancement; we note these and future studies should address them.

Another limitation results from the biases affecting performance rankings (Bol, 2011), which are well-known but impossible to control in longitudinal studies. The predictive model is trained to predict performance rankings, and it learned the rater biases, which were replicated. A post-evaluation of the ANN by the raters was out of the scope of this work. The analysis of time and effort that ANN may reduce is must be pursued in future research. Additionally, the satisfaction of lawyers, partners and HR practitioners is important for future evaluations.

Because no widely accepted theory for the design of networks is available, decisions on training, the number of hidden layers and nodes and training adjustments for increasing accuracy must be conducted by trial and error. Thus, different and better networks can be designed (Naik and Ragothaman, 2004).

One topic that is worth investigating relates to additional predictors that may increase the accuracy of the predictive model. ANNs that consider different professional levels and tenures should also be explored. This might allow for more accurate predictions over time. Firms, in this scenario, would not need to wait for the full evolution of a lawyer’s career to identify a tournament winner. New career architectures following different performance ranking trends are a final topic for future research. TM requires that talent be managed for the long term (Boudreau and Ramstad, 2005), and ANNs allow “looking at long-standing problems in new ways” (Scarborough and Somers, 2006, p. 46).

6.2 Conclusion
It is time consuming and troublesome to rank all lawyers against peers. This burden of this practice, however, can be alleviated by the predictive use of ANNs. This paper employed methods beyond the traditional static study of performance, including non-linear modeling for prediction. The study tested an ANN’s prediction of performance rankings that is adjusted until two-year lag time. The superiority of the ANN over an MLR model was tested and confirmed. The time freed can be invested in strategic planning, lawyers’ feedback to raise awareness and talent development for readiness in the turmoil of the new millennium.

References


Further reading


About the authors

Susana Almeida Lopes received the Psychology Degree from Coimbra University, with specializations in Work and Organizational Psychology, and Clinical Psychology (2000); the PhD Degree in Organizational Psychology from Lisbon University (2017); and post graduations on Applied Quantitative Methods from ISCTE (2002), Law Firm Management from Porto Business School (2010), and on Strategic HR Practices from Cornell University (2010). SAL is currently a member of the Research Center for the Psychological Science of the Lisbon University. She started her career at SHL, where she advanced from Trainee to Managing Consultant. She is currently the Organizational Director at VdA, where she addresses strategic HR planning, talent management, competency model, appraisal system, career management, reward, development and recruitment. She directs the HR analytics research projects at VdA Legal Analytics Bureau. SAL is a member of the board of the Order of Psychologist in Portugal, since 2016. She was awarded several times for talent management innovative projects and ranked among the most innovative individuals by Financial Times in 2014. Susana Almeida Lopes is the corresponding author and can be contacted at: sal@vda.pt

Maria Eduarda Duarte is Full Professor at the University of Lisbon, Faculty of Psychology, where she directs the Master Course in Psychology of Human Resources, Work and Organizations. Her professional interests include career psychology theory and research, with special emphasis on issues relevant to adults and the world of work. Her publications and presentations have encompassed topics on adult’s career problems, testing and assessment and counseling process. She was Chair of the Portuguese Psychological Society (2005–2018); she also served on editorial boards for some Portuguese, European and Iberia-American journals. She was the Director of the National Institute of Guidance (2009–2014). She is President of Counseling Division (2014–2018), IAAP. She is President of UNESCO CHAIR – Lifelong guidance and counseling, University of Wroclaw. She is Fellow Award – IAAP (2014), and ESVDC award 2015. She is also National Defence Adviser, since 2007.

João Almeida Lopes received the BSc/MSc Degree in Chemical Engineering in 1997 at the Technical University of Lisbon, and the PhD Degree in Pharmaceutical Engineering in 2001. Until 2003 he was a Postdoctoral Researcher at the same University and Consultant for the pharmaceutical industry.
In 2006, he joined the University of Porto as Assistant Researcher in the areas of green chemistry, chemometrics and process spectroscopy. He created the Chemometrics and Process Analytical Technology research group being its coordinator since then. He founded Worldmetrics Solutions, a consulting company providing technology solutions for high-throughput industrial processes monitoring. He has been Assistant Professor of Pharmaceutics at the Faculty of Pharmacy, University of Lisbon since 2014. JAL published more than 100 scientific articles, coordinated more than 40 R&D projects and supervised circa 90 students and researchers (MSc, PhD, Postdoc levels). He was the President of the Portuguese Association of Researchers in Science and Technology (2012–2014) and currently General Assembly President, Vice-President of the Portuguese Society of Pharmaceutical Sciences (from 2017), appointed President of the Analytical Chemistry division of the Portuguese Society of Chemistry (from 2019), member of the steering committee of the QbD/PAT network of the European Federation for Pharmaceutical Sciences (from 2011) and the Voice of Researchers initiative of the European Commission (2010–2016). He is Scientific Consultant in Process Systems Engineering, Process Analytics and Pharmaceutical Development for several industrial companies.
The impact of corporate finance decisions on market value in emerging markets

Guler Aras
Business Administration, Center for Finance, Governance and Sustainability (CFGS), Yildiz Technical University, Istanbul, Turkey, and
Filiz Mutlu Yildirim
Business Administration, Yildiz Technical University, Istanbul, Turkey

Abstract

Purpose – The purpose of this paper is to analyze the conformity of the impact of corporate finance decisions on market value with the basic theoretical approaches in the two emerging economies, which show great similarities in terms of the economic structures, and to examine the results obtained by determining how these decisions affect market value comparatively.

Design/methodology/approach – In this study, the effect of corporate finance decisions on market value is tested empirically with panel data analysis method by using data of 274 real sector firms traded in BIST and 249 firms in BOVESPA industry index, between 2010 and 2014.

Findings – The analysis results show that the increase in the borrowing level of firms operating in both countries reduces the market value; the increase in the level of profitability in the firms has a positive effect on the market value. In addition, it is possible to say that the effect is different in terms of investment decisions for Turkish and Brazilian firms.

Research limitations/implications – The limitations of the study are that the non-financial sectors between 2010 and 2014 in Turkey and Brazil including the company data, and the companies with missing years are removed from the data set. The findings show that more effective and balanced management of the variables by the financial manager affects these rates and have an impact on increasing the market value.

Originality/value – In this study, a fundamental subject in finance is addressed by analytical methodology and comparative tests for countries are conducted.

Keywords Turkey, Brazil, Market value, Panel data analysis, Corporate finance decisions

Paper type Research paper

1. Introduction

Market value, which is one of the most important indicators of corporate profitability and success, is in interaction with corporate finance decisions. Capital structure and investment decisions are among the main areas of corporate finance. Reducing the cost of capital by examining the capital structure and maximizing the market value consequently, the formation of effective asset composition plays a decisive role in the future performance of firms. In this context, corporates’ desire to increase their profitability, to make their operations sustainable and to protect their competitive advantages, increased the importance given to the market value. This situation necessitated the investigation of the impact of corporate finance decisions on market value. When the literature is examined, a large number of studies have been detected which examine the relationship between market value and capital structure and investment decisions.

Basic theoretical approaches that explain the relationship between capital structure and market value are the net income approach, the net operating income approach, the traditional approach and the Modigliani and Miller approach. According to the net income approach, the cost of capital can be reduced, and the market value can be increased while the cost of capital is fixed in the net operating income approach, and it does not have an effect on the market value (Durand, 1952). The traditional approach has a single optimal
capital structure. According to this approach, market value increases until reaching an optimal capital structure, and decreases as it moves away from this level. According to the Modigliani and Miller approach, capital structure has no effect on market value under certain assumptions (Modigliani and Miller, 1958).

A number of studies have been carried out to address the relationship between capital structure and financial ratios and market values, with reference to the basic theoretical approaches of capital structure (Birgili and Düzer, 2010; Altan and Arkan, 2011; De Lima et al., 2012; Loncan and Caldeira, 2014; Ayrçay and Türk, 2014; Inci, 2014). Studies that test the relationship between investment decisions and financial performance and market value are frequently found in the literature. Shin and Soenen (1998) are among the first names examining the relationship between working capital and financial performance. This relation is also tested by Deloof (2003), Gill et al. (2010), Ching et al. (2011), Karaduman et al. (2011), Vural et al. (2012), Kendirli and Konak (2014) and Keskin and Gökalp (2016). In addition, De Almeida and Eid (2014) and Ata and Bugan (2016) have conducted studies to examine the relationship between working capital and market value.

Financial liberalization, financial innovations, technological developments, financial crises, increased direct investments, increased mergers and acquisitions in the financial sector that the globalization in financial markets brings with it have brought financial markets closer together and enabled countries to be comparable on the international area. The aim of the study in this direction is to test country differentiation in the influence of corporate finance decisions on the market value in developing countries for Turkish and Brazilian firms and to examine the results comparatively by determining how these decisions affect market value. The reason for investigating these two countries is that both countries are included in the group of developing countries and have similar economic characteristics, making them comparable.

As well as the developing markets both in the academic host for investors wishing to achieve greater returns than possible; in this case, the most important developments in the market makes comparison of Turkey and Brazil necessary. Characterized by low growth rates, the international economic environment has retained the ability of both countries to rapidly and profoundly influence local economies (Welch, 2013). In a report released in 2005, Fitch drew attention to the similarity of Turkey and Brazil. According to this report, both countries, as a legacy of the past, need the confidence of the markets to sustain their public and external debt. The need for a high primary budget surplus to be able to sustain this confidence, strong monetary policies to ensure the credibility of central banks and structural reforms needed to further develop the market economy are needed (Fitch Ratings, 2005). Turkey and Brazil, developing countries in their economic structures and monitoring improvement in the economy, slipped from the economic crisis of the past which gave birth to the idea that they resemble each other. This opens up the importance of the need to analyze the economies of both countries through economic and financial indicators and make both countries comparable.

To sum up the information about the economic structures of the countries that have discussed, the Turkish economy experienced tough times with the crises of the 1990s and 2001, but it has taken firm steps despite uncertainties that emerged after the crises, with policies and structural reforms implemented. Following the global financial crisis and the slowdown in 2008 and the contraction in 2009, it has increased its growth rates from 0.66 and −4.83 percent to 9.16 percent in 2010. Brazil is expected to be the one of the world’s strongest economies in the 2050s, along with other BRIC countries (Brazil, Russia, India and China), thanks to such features as being the main provider state for South American countries and the amount of population they have (Deik, 2014). Brazil is also affected by the global crisis in 2008, and in 2010, it achieved a 7.57 percent growth thanks to its policies. Macroeconomic indicators of Turkey and Brazil are comparatively included. Unemployment
rates were obtained from Turkish Statistical Institute and other rates were obtained from the World Bank (www.tuik.gov.tr/; www.worldbank.org/).

In this study, the effect of corporate finance decisions on market value is tested empirically with panel data analysis method by using data of 274 real sector firms traded in BIST and 249 firms in BOVESPA industry index, between 2010 and 2014. The analysis results show the increase in the borrowing level of firms, operating in both countries reduces the market value; and the increase in the level of profitability in the firms has a positive effect on the market value. Furthermore, it is possible to say that the effect is different, in terms of investment decisions for Turkish and Brazilian firms.

The study is structured into literature review, methodology, research findings and conclusion sections after the introduction section. In the literature review, information about the basic theoretical approaches and the studies done with reference to these approaches is provided. In the methodology section, the purpose and the method of the study, the data set used, constraints, variables, modeling and statistical analysis are included. The research findings obtained after the analyses performed are evaluated in the conclusion section, and the studies that can be done later are suggested after specifying the constraints of the study.

2. Literature review
In the literature, a number of studies have been carried out to determine the effect of corporate finance decisions on market value, and in these studies, the relationship between capital structure and investment decisions and market value has been examined. The reason for investigating these two countries is that Turkey and Brazil are both developing countries and are also in the same class with having similar economic structures. This situation brings similarities in the studies done in the literature. The main theoretical approaches that examine the relationship between capital structure and market value are the net income approach, the net operating income approach, the traditional approach and the Modigliani–Miller approach.

Durand (1952) defined the net income approach in his work. The basic assumption of the net income approach is that the share of debt within the capital share increases, while the cost of debt and equity remains constant. According to this approach, since the cost of equity is higher than the cost of debt, increasing the share of debt in the capital structure reduces the cost of capital and increases the market value. Durand also included the net operating income approach in his study. The basic assumption is that as the share of debt in the capital share increases, the cost of equity rises despite the fixed cost of debt. In this approach, the cost of capital is fixed and changes in the capital structure have no effect on market value. The traditional approach has a single optimal capital structure and borrowing to a certain extent reduces the cost of capital and then raises it from this level. Modigliani and Miller (1958) argue that market value is not influenced by capital structure under some assumptions.

Modigliani and Miller (1963) included the corporation tax in the model they created in 1958. With the tax savings applied on debts, the debt increase in capital structure has become more advantageous compared to the increase in equity. Miller (1977) and Haugen and Senbet (1978) have added income tax to the model as well as corporation tax. The study suggests that the tax savings brought by the debts will cause difficulties in fulfilling the obligations in the future, which may lead to the cost of bankruptcy. Jensen and Meckling (1976) and Fama (1980) have studied agency cost; Myers and Majluf (1984) have studied the cost of asymmetric information.

Birgili and Düzer (2010), Aydınçay and Türk (2014), Altan and Arkan (2011) and İnci (2014) investigated the effects of capital structure and financial ratios on market value for Turkish companies. Birgili and Düzer (2010) investigated the financial ratios and firm value relationships in the study conducted with companies that traded in IMKB, covering
2001–2006 by using panel data analysis method. The analysis have reached the conclusion that firms’ liquidity situation, financial structure and stock market performance have more effect on firm value, but there is no statistically significant relationship between the activity rates and profitability ratios and firm value. Likewise, as Ayrçay and Türk (2014) stated, the relationship between financial ratios and firm value is tested with the data of the companies traded in BIST between 2004 and 2011. According to the findings, there is a statistically significant relationship between acid-test ratio, asset turnover rate, MV/BV ratio and financial leverage ratios and firm value, but no significant relationship between borrowing rate, and asset profitability ratio and firm value has been observed. Altan and Arkan (2011) determined that firm values re influenced by firms’ financial structures in their studies that they used company data registered in IMKB industry index.

İnci (2014) examined the impact of financial management decisions on profitability and market value in a study of industrial companies at BIST by using panel data analysis method. As a result of the study covering the years 2003–2012, a meaningful association has been found that the value and profitability of the companies are negatively correlated with the capital structure and net fixed asset structure, and positively associated with the cash dividend distribution, turnover rates, asset growth and liquidity ratios. Burca (2008) tested the existence, direction and strength of the relationship between capital structure and stock value in IMKB, BOVESPA, NYSE and FTSE between 1996 and 2006, by using correlation analysis, portfolio analysis and panel regression analysis methods. According to the findings, there is a positive relationship between the capital structure and firm value for all the stock markets covered by the investigation. Dinçergök (2010) comparatively examined firm-specific and country-specific factors affecting capital structure for Turkey, Argentina, Brazil, Indonesia and Mexico. Demirdöven (2013) tested the validity of capital structure theories for Turkey and Brazil and investigated similar and different aspects of the results.

Mesquita and Lara (2003) tested the existence of a relationship between capital structure and profitability using the OLS method using Brazilian company data. As a result of the analysis made, positive relation between profitability and long-term debt and negative relation with short term debt are found. De Lima et al. (2012) examined the impact of capital structure on market value by using the top 1,000 company data in Brazil in 2010. Brito et al. (2007) carried out studies to define the determinants of capital structure for the largest companies in Brazil, Forte et al. (2013) carried out studies with small- and medium-sized corporations. Bastos and Nakamura (2009) carried out a study with publicly traded companies in Brazil, Mexico and Chile. In this study covering 2001–2006, they investigated capital structure determinants by using panel data analysis method. As a result of the research, it has been proved that the liquidity, profitability, market value book value ratio and size are the most specific factors for the capital structure of all three country firms.

Thippayana (2014), Kumar et al. (2017), Sofat and Singh (2017) and Sheikh and Qureshi (2017) examined determinants of capital structure. Apostol (2017) also examined optimal capital structure. Similarly, Vo and Ellis (2017), Ezirim et al. (2017) and Ha and Tai (2017) investigated the relationship between capital structure and firm value.

In Shin and Soenen’s (1998) study, the relationship between working capital and profitability is examined by using data of US firms between 1975 and 1994. The result of the analysis showed that there is a strong correlation between the working capital and profitability in the negative direction. Gill et al. (2010) investigated the relationship between working capital management and profitability on American firms. Ogundipe et al. (2012) examined the impact of working capital management on firm performance and market value in Nigeria, while Bandara (2015) examined the impact of the working capital management policy on market value on Sri Lanka firms. Palombini and Nakamura (2012)
examined the key factors of working capital management in the Brazilian market using panel data analysis and came to the conclusion that the level, size and growth rate of debt can affect corporations’ working capital management. Loncan and Caldeira (2014) used panel data analysis method to examine the relationship between capital structure, cash amount and firm value. Working on Brazilian companies, Ching et al. (2011) studied the relationship between working capital and profitability. Likewise, De Almeida and Eid (2014) also utilized Brazilian companies. They examined the relationship between working capital and firm value.

Luo and Hachiya (2005) found that the cash level effects market value in the negative direction in their study in which they used the data from Japanese firms that were traded on the Tokyo Stock Exchange. Similarly, Lee and Lee (2009) found a negative relationship between cash level and firm value in their study of Asian countries including Malaysia, Philippines, Indonesia, Singapore and Thailand between 2001 and 2005. In studies conducted on Turkish firms, Karaduman et al. (2011) and Keskin and Gokalp (2016) investigated the effect of the working capital management on the performance of the firms; Çakır and Küşükkaplan (2012) investigated the effect of working capital elements of firms on the profitability and market value of firms; and Küşükkaplan (2013) investigated the relationship between the internal variables of firms and the market value. Similarly, Ata and Buğan (2016) also examined the relationship between the working capital and firm value. In this study, the panel data analysis method is used.

Altaf and Shah (2017) and Zhang et al. (2017) conducted research on working capital management. Moreover, Singh et al. (2017), Panda and Nanda (2018), Rao and Rao (2017), Shrivastava et al. (2017) and Kasozi (2017) examined the relationship between working capital and profitability. Koumanakos (2008) tested the hypothesis that efficient inventory management leads to an improvement in a firm’s financial performance. In addition, Shah et al. (2017) investigated the most effective and significant factors influencing working capital policy. Furthermore, Aras et al. (2017) examined the impact of investment and financing decisions on market value in Turkey. A recent study which is conducted by Saona and San Martin (2018) investigated determinants of firm value in Latin America. They utilized public firms from Argentina, Brazil, Chile, Colombia, Mexico and Peru for the 1997–2013 period. As a result, ownership concentration, capital structure and dividend policy are determined as significant drivers of the market value of the firm.

3. Methodology
3.1 The purpose of the study and methods
The purpose of this study is to test the country differentiation of impact of the corporate finance decisions on market value in developing countries for Turkish and Brazilian firms and to examine the results obtained by determining how these decisions affect market value comparatively.

Panel data analysis method is used in the study. The panel data have both time and space dimensions, and the same cross-sectional units are observed over time (Gujarat and Porter, 2014). The reasons for using panel data analysis in the study are that it has more number of observations and sample randomness than time series and horizontal section data combined, it allows to control the effect of variables that are unmeasurable or not included in the model, it reveals the dynamic relation and econometric analyses are more effective than time series and horizontal section data (Hsiao, 2006). The usage of both horizontal section and time series dimensions in the panel data makes it possible to investigate complex models more effectively, not only in cross-section or only in time series data (Gujarat and Porter, 2014). The availability of panel data actually simplifies the computation and inference; multiple observations may allow a researcher to make different transformations to identify an otherwise unidentified model (Hsiao, 2006).
3.2 Data set used in the study and constraints
The study covers exchange rates and inflation rates of both countries and the data of 249 companies in BOVESPA and 274 companies in BIST from non-financial sectors that are regularly accessible to their data between 2010 and 2014. The reason for using non-financial sector is to investigate the impact of corporate finance decisions on the market value more efficiently by removing financial sectors from the data set, since the balance sheet structures of financial sectors differ. The homogeneity of the selected data set will lead to more qualified and effective results. Therefore, financial companies have been removed from the data set. Turkey and Brazil data are obtained from Bloomberg Professional Service, inflation rates and exchange rates are obtained from the World Bank. Extreme values are removed from the data.

The limitations of the study are that in the non-financial sectors between 2010 and 2014 in Turkey and Brazil, including the company data, the companies with missing years are removed from the data set due to the use of balanced panel data and extreme values are removed from the data set.

3.3 Variables used in the study
The study consists of three models. The logarithm of the historical market capitalization value is used as a dependent variable throughout the models. The historical market capitalization value is one of the indicators of the market value and obtained by multiplying the stock market value by the total number of shares. It is aimed to obtain more healthy results by decreasing the differences between values by logarithmic transformation.

The independent variables used are the total debt/total equity ratio in Model 1; in Model 2, accounts payable turnover rate, accounts receivable turnover rate, asset turnover rate, inventory turnover rate, asset reinvestment ratio, net fixed assets turnover rate, cash ratio; In Model 3, total debt/total equity ratio, accounts payable turnover rate, accounts receivable turnover rate, asset turnover rate, inventory turnover rate, asset reinvestment ratio, net fixed asset turnover rate, cash ratio, exchange rate and inflation rate are used. Return on assets ratio is used as control variable in all three models.

The studies in the literature have been done to determine the independent variables that explain the dependent variable in the most appropriate way. In addition, the explanatory power of the model which is frequently used in the literature has been researched. Moreover, independent variables and control variable which are included in the study have a high correlation with the dependent variable and low correlation among themselves. There are abbreviations and definitions of the variables used in Table I.

3.4 Model of the study
Models used in the study are structured as follows.

Model 1: capital structure decisions:
\[
LHM_{cap_{it}} = a + b_1 TDTE_{it} + b_2 ROA_{it} + e_{it}.
\]

Model 2: investment decisions:
\[
LHM_{cap_{it}} = a + b_1 APT_{it} + b_2 ART_{it} + b_3 AT_{it} + b_4 IT_{it} + b_5 ARR_{it} + b_6 NFAT_{it} + b_7 CR_{it} + b_8 ROA_{it} + e_{it}.
\]

Model 3: capital structure decisions, investment decisions and macroeconomic variables:
\[
LHM_{cap_{it}} = a + b_1 TDTE_{it} + b_2 APT_{it} + b_3 ART_{it} + b_4 AT_{it} + b_5 IT_{it} + b_6 ARR_{it} + b_7 NFAT_{it} + b_8 CR_{it} + b_9 ER_{it} + b_{10} IR_{it} + b_{11} ROA_{it} + e_{it}.
\]
\[
i = 1, \ldots, N \text{ and } t = 1, \ldots, T,
\]
where \( i \) represents the number of corporations, \( t \) represents the number of periods, \( a \) is constant term and \( e \) is error term.

### 3.5 Statistical analysis

The data used in the study are purified from extreme values and 617 observation values for Turkey and 349 observation values for Brazil have been reached. The average values of descriptive statistics for Turkish and Brazilian companies are summarized in Table II.

In the analysis, the existence of the problem of collinearity caused by the high correlation between independent variables and the control variable is investigated by applying a correlation analysis. As a result of the study, it is determined that there is no problem of collinearity and the correlation analyzes for Turkish and Brazilian companies are given in Tables III and IV.

In the literature, panel data models are generally applied in two different methods: random effect models and fixed effect models. In the random effect models, error components and explanatory variables are unrelated, while in the fixed effect models,
there is a relation between them (Gujarati and Porter, 2014). Accordingly, the Hausman (1978) test is used to select random and fixed panel data models.

The Hausman test results obtained for Turkey and Brazil are given in Table V. The Hausman test for Turkey and Brazil has a \( p \)-value of less than 0.05 at the 5 percent significance level: the hypothesis is rejected: “There is no relationship between error components and independent variables.” In this case, there is a relation between the error components and the independent variables, so endogeneity problem is determined. It is suggested in the literature to use fixed effect models where the problem of endogeneity exists (Greene, 2003), so in this study, the fixed effect model is chosen.

In the random effect model, it is assumed that the constant term of a single unit is randomly drawn from a much larger mass with constant mean value. The singular constant

<table>
<thead>
<tr>
<th>Models</th>
<th>Hausman test values</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \chi^2 )</td>
<td>Turkey</td>
<td>Brazil</td>
<td>Turkey</td>
</tr>
<tr>
<td></td>
<td>( p )-value</td>
<td>51.60</td>
<td>46.77</td>
<td>56.92</td>
</tr>
<tr>
<td>Note:</td>
<td>( **p &lt; 0.05 )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
term is expressed as a deviation from this constant mean value. The advantage of the random effect model over the fixed effect model is that it is conservative in the degree of freedom because N (number of cross-sections) does not need to estimate the cross-section constant. The random effect model is suitable where the constant term of each section (random) is not related to the explanatory variables. Another advantage of the random effect model is that it allows to model the fixed variables for a certain subject. This cannot be done in a fixed effect model because all these variables are linearly related to the constant term that is specific to the subject (Gujarati and Porter, 2014). The fixed effect model examines the unobserved effects of the horizontal cross-sectional units as an unchanging parameter that remains constant over time.

After the selection of the panel data model to be applied, the presence of changing variance and autocorrelation are investigated. The changing variance is that the variances of the error terms in the population regression function are not constant and autocorrelation is the existence of the relationship between consecutive error terms. Changing variance and autocorrelation in the model leads to misleading results and the inability to make effective estimates (Gujarati and Porter, 2014). In the study, modified Wald test is used for testing the presence of changing variance; the Wooldridge test is used to test the presence of autocorrelation. The modified Wald test and Wooldridge test results obtained for Turkey and Brazil are given in Tables VI and VII.

The modified Wald test for Turkey and Brazil has a \( p \)-value of less than 0.05 at the 5 percent significance level: the hypothesis is rejected: “There is no changing variance.” The Wooldridge test for Turkey and Brazil has a \( p \)-value of less than 0.05 at the 5 percent significance level: the hypothesis is rejected: “There is no autocorrelation.” In this case, changing variance and autocorrelation are determined. The variance and autocorrelation problems obtained as a result of the analysis are solved by including the cluster estimator. Panel data analysis results for Turkey and Brazil are given in Table VIII.

4. Research findings
Models constructed within the scope of the analysis are tested with panel data analysis method. Findings obtained from the examination are presented in this section.
Model 1 shows the effect of firms’ capital structure decisions on the market value. ROA and Cons in Turkey have a significant relationship in positive direction with the dependent variables.

<table>
<thead>
<tr>
<th>Models</th>
<th>Modified Wald test values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Turkey</td>
<td>Brazil</td>
</tr>
<tr>
<td>( \chi^2 )</td>
<td>3.8e+31</td>
</tr>
<tr>
<td>( p )-value</td>
<td>0.0000***</td>
</tr>
</tbody>
</table>

**Note:** ***\( p < 0.05 \)**

<table>
<thead>
<tr>
<th>Models</th>
<th>Wooldridge test values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Turkey</td>
<td>Brazil</td>
</tr>
<tr>
<td>( F )</td>
<td>25.100</td>
</tr>
<tr>
<td>( p )-value</td>
<td>0.0000***</td>
</tr>
</tbody>
</table>

**Note:** ***\( p < 0.05 \)***
<table>
<thead>
<tr>
<th>Models</th>
<th>Independent variables and control variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Turkey</td>
<td>Brazil</td>
<td>Turkey</td>
<td>Brazil</td>
</tr>
<tr>
<td>TD/TE</td>
<td>−0.0470 (0.0702)</td>
<td>−0.1455 (0.0559)**</td>
<td>−0.1330 (0.0687)*</td>
<td>−0.2014 (0.0878)**</td>
</tr>
<tr>
<td>APT</td>
<td>−0.0243 (0.0125)*</td>
<td>−0.0236 (0.0134)*</td>
<td>−0.0072 (0.0121)</td>
<td>−0.0257 (0.0113)**</td>
</tr>
<tr>
<td>ART</td>
<td>0.0020 (0.0080)</td>
<td>0.0654 (0.0319)**</td>
<td>0.0154 (0.0112)</td>
<td>0.0028 (0.0235)*</td>
</tr>
<tr>
<td>AT</td>
<td>−0.1800 (0.1512)</td>
<td>0.5085 −0.4190</td>
<td>−0.3614 (0.1500)**</td>
<td>−0.0085 (0.5200)</td>
</tr>
<tr>
<td>IT</td>
<td>−0.0019 (0.0110)</td>
<td>−0.0018 (0.0007)***</td>
<td>0.0048 (0.0156)</td>
<td>−0.0024 (0.0006)***</td>
</tr>
<tr>
<td>AARR</td>
<td>0.0225 (0.0106)**</td>
<td>0.0328 (0.0233)</td>
<td>0.0182 (0.0116)</td>
<td>0.0280 (0.0237)</td>
</tr>
<tr>
<td>NFAT</td>
<td>0.0086 (0.0025)***</td>
<td>−0.0241 (0.0132)*</td>
<td>0.0067 (0.0021)***</td>
<td>−0.0228 (0.0132)*</td>
</tr>
<tr>
<td>CR</td>
<td>0.0427 (0.0463)</td>
<td>0.0416 (0.0998)</td>
<td>0.1066 (0.0356)*</td>
<td>0.0203 (0.0893)</td>
</tr>
<tr>
<td>ER</td>
<td>0.3071 (0.0938)***</td>
<td>0.2207 (0.1467)</td>
<td>0.0865 (0.0124)***</td>
<td>−0.0842 (0.0227)**</td>
</tr>
<tr>
<td>IR</td>
<td>0.0174 (0.0044)***</td>
<td>0.0383 (0.0072)***</td>
<td>0.0212 (0.0040)***</td>
<td>0.0434 (0.0091)***</td>
</tr>
<tr>
<td>ROA</td>
<td>0.1747 (0.0518)***</td>
<td>14.3523 (0.0615)***</td>
<td>12.7162 (0.1360)***</td>
<td>13.7936 (0.2878)***</td>
</tr>
<tr>
<td>F-test</td>
<td>12.75***</td>
<td>14.3523 (0.0615)***</td>
<td>12.7162 (0.1360)***</td>
<td>14.5933 (0.2915)***</td>
</tr>
<tr>
<td>R²</td>
<td>0.0471</td>
<td>0.1667</td>
<td>0.0582</td>
<td>0.2965</td>
</tr>
</tbody>
</table>

Number of obs. 874 793 701 390 617 349

Notes: *p < 0.10; **p < 0.05; ***p < 0.01
variable at 1 percent level, whereas there is no significant relationship between TDTE and the dependent variable. In Model 1, the research findings in Turkey are parallel with Modigliani and Miller’s (1958) approach, and capital structure decisions have no meaningful relationship with market value and do not affect market value by changing capital structure. In Brazil, there is a meaningful relationship with TDTE at the 5 percent significance level in the negative direction and with ROA and Cons at the 1 percent significance level in the positive direction.

Model 2 shows the effect of firms’ investment decisions on the market value. In Turkey, ART, AT, IT and CR variables have no significant effect on market value, whereas NFAT, ROA and Cons have a significant relationship in positive direction with the dependent variable at 1 percent level, ARR has a significant relationship at 5 percent level in the positive direction and APT has a meaningful relationship at 10 percent level in negative direction. In Brazil, high profitability of firms have a positive effect on market value, while AT, ARR and CR variables have no meaningful effect on market value. In Brazil, at 1 percent level ROA and Cons have a significant relationship in the positive direction and IT in the negative direction, ART has a significant relationship in the positive direction at 5 percent level and, at 10 percent level, APT and NFAT have a significant relationship in the negative direction with the dependent variable.

Model 3 shows the effects of firms’ capital structure decisions, investment decisions and macroeconomic variables on the market value. The independent variables and control variable in Model 3 explains the dependent variable in the ratios of 24 percent in Turkey and 30 percent in Brazil. As a result of the F-test, the model is significant at the 1 percent level. In Model 3, for Turkey; NFAT, ER, IR, ROA and Cons have a significant relationship at 1 percent level in positive direction, AT has a significant relationship at 5 percent level in negative direction, at 10 percent level TDTE has a significant relationship between negative direction and CR in the positive direction. There is no significant relationship between APT, ART, IT, ARR and dependent variable. In Brazil, at 1 percent level ROA, Cons have a significant relationship in the positive direction and IT in the negative direction, TDTE, APT and IR have a significant relationship at 5 percent level in the negative direction, and at 10 percent level, ART has a significant relationship between the positive direction and NFAT in the negative direction. There is no significant relationship between AT, APT, CR, ER and the dependent variable.

When compared to our study, there are similar and different study findings in the literature. Masulis (1983) and Welch (2004) found that the level of borrowing is positively correlated with firm value, Mumtaz et al. (2013) found that the level of borrowing is negatively correlated with firm value, while Hull (1999) defends the existence of an optimal capital structure. Ayırçay and Türk (2014) found a negative correlation between firm value and asset turnover rate, while Mansoori and Muhammad (2012) found significant correlation in positive direction. Raheman et al. (2007) found a negative correlation between firm value and inventory turnover rate, while İnci (2004) found a positive correlation. Birgili and Düzer (2010) found a significant positive correlation between cash ratio and market value. On the other hand, Luo and Hachiya (2005) and Lee and Lee (2009) have found a negative relationship between cash level and market value.

Studies on examining the relationship between macroeconomic indicators and stock prices frequently exist in the literature. The exchange rate and inflation variables are often used in these studies. In the analyses made, Solnik (1987) and Maysami et al. (2004) found a positive correlation between stock price and exchange rate, and Soenen and Hennig (1988) found a negative correlation. In studies that tested the relationship between stock return and inflation, Fama (1981), Geske and Roll (1983), Chopin and Zhong (2000) and Kim (2003) found a statistically significant negative correlation between stock returns and inflation.
5. Conclusion

This study is conducted to test the country differentiation in the influence of corporate finance decisions on the market value in the developing countries for Turkish and Brazilian firms and to examine the results comparatively by determining how these decisions affect the market value.

In Model 1, the findings of the study show that the significant effect of the return on assets ratio on market value in positive direction is common in Turkey and Brazil. On the other hand, the effect of total debt/total equity ratio on market value differs in Turkey and Brazil. In Model 2 applied in the study, the return on assets ratio’s positive and the accounts payable turnover rate’s negative significant effect on the market value are common in Turkey and Brazil. Country differentiation is detected when the accounts receivable turnover rate, the inventory turnover rate, the asset reinvestment ratio and the net fixed asset turnover rate’s effect on the market value is investigated, and no significant relation is found between the asset turnover rate and the cash ratio and the market value in both countries.

In Model 3, return on assets ratio’s positive, and total debt/total equity ratio’s significant negative effects on market values are common in Turkey and Brazil. In Turkey, 37.11 percent of the assets in corporate are financed by liabilities and 62.89 percent by equity, and In Brazil, 45.05 percent of assets in corporate are financed by liabilities and 54.95 percent by equity. Country differentiation is detected when the accounts payable turnover rate, the accounts receivable turnover rate, the asset turnover rate, the inventory turnover rate, the net fixed asset turnover rate, the cash ratio, the exchange rate and the inflation rate’s effects on the market value are investigated, and no significant relationship is found between the asset reinvestment ratio and the market value in both countries. In this case, it is possible for the financial manager to increase the market value through more efficient and balanced management of the variables.

In financing policy decisions, it is aimed to maximize market value by determining the lowest cost fund source at a certain risk level. Fund needs can be met by two sources: debt and equity. The most important issue to be decided here is the distribution of debt and equity. The analysis results show that an increase in the borrowing level of firms operating in both countries reduces the market value. In this case, finance managers can increase market value by focusing on financing through equity. Financial managers who avoid risk because they do not have compulsory payments such as interest and principal payments in the financing via equity can generally prefer low borrowing level.

It is possible to say that the effect is different in terms of investment policy decisions for Turkish and Brazilian firms. In Turkey, findings of the study show that the net fixed asset turnover rate and the cash ratio have a significant relationship in the positive direction and the asset turnover rate in the negative direction with the market value. The positive impact on the market value of the net fixed asset turnover rate indicates that these fixed assets are used effectively and managed properly. As the liquidity increases, the debt payment capacity increases and the financial risk decreases, which also increases the market value. However, high liquidity can be a sign of idle usage of sources. For this reason, the financial manager should be able to provide an optimum balance between effective use and risk, taking into account the social and economic conditions, the business sector and the sector averages. The asset turnover rate measures the sales generated by each unit asset and how effectively the firm uses all its assets. According to the study results, the effect of the market value of the asset turnover rate is negative in Turkey and, in this case, it is advisable for finance managers to increase the asset level in the direction of sales.

In Brazil, the findings of the study show that the accounts payable turnover rate, the inventory turnover rate and the net fixed asset turnover rate have a significant relationship between the negative direction and the accounts receivable turnover rate in positive
direction with the market value. Companies can pay their debts in approximately every 37 days and collect their receivables in about every 55 days in Brazil. In such a case, it may be advisable to increase the debt repayment period and reduce the collection period. The high inventory turnover rate indicates the effectiveness of inventory management. However, according to the results of the study, there is a negative effect on the market value. This may be due to the additional cost created by the issuance of small orders to replenish inventories in the event of a low stock level and to the pauses in sales if additional inventories cannot be provided. In this case, it may be advisable for finance managers to increase their stock levels, but the storage cost and the risk of idle investment should not be overlooked if kept too high. In addition to the results, decreases in the net fixed asset turnover rate have a positive impact on the market value.

Ratios of company structures that can be reached from the descriptive statistics table and macroeconomic indicators of Turkey and Brazil are among the main reasons for the country differentiation. Macroeconomic indicators of Turkey and Brazil are given in Tables AI–AIII.

It can be said that in the direction of the obtained results, the effect of the corporate finance decisions on the market value shows similarity and difference for the Turkish and Brazilian firms. The findings obtained in the study are parallel with the literature. The findings of previous studies by Birgili and Düzer (2010) and Ayrıçay and Türk (2014) also showed that the ratios used in financial analysis have an effect on firm value. Altan and Arkan (2011) found that the financial structure has effects on firm value, and Ata and Buğan (2016) reached the conclusion that the working capital management effectiveness has an effect on the firm value.

The limitations of the study are that the non-financial sectors between 2010 and 2014 in Turkey and Brazil include the company data, the companies with missing years are removed from the data set due to the use of balanced panel data and extreme values are removed from the data set. In future studies, sectoral differentiation of the effect of the corporate finance decisions on the market value can be investigated by separating sectors in the data set used.

References


Corporate finance decisions

1973


**Further reading**

Turkish Statistical Institute, available at: www.tuik.gov.tr/


**Appendix**

<table>
<thead>
<tr>
<th>Macroeconomic indicators</th>
<th>Inflation (annual %)</th>
<th>Dollar exchange rate</th>
<th>GDP growth (annual %)</th>
<th>GDP ($bn )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Turkey</td>
<td>Brazil</td>
<td>Turkey</td>
<td>Brazil</td>
</tr>
<tr>
<td>2005</td>
<td>10.14</td>
<td>6.87</td>
<td>1.34</td>
<td>2.43</td>
</tr>
<tr>
<td>2006</td>
<td>9.60</td>
<td>4.18</td>
<td>1.43</td>
<td>2.18</td>
</tr>
<tr>
<td>2007</td>
<td>8.76</td>
<td>3.64</td>
<td>1.30</td>
<td>1.95</td>
</tr>
<tr>
<td>2008</td>
<td>10.44</td>
<td>5.86</td>
<td>1.30</td>
<td>1.83</td>
</tr>
<tr>
<td>2009</td>
<td>6.25</td>
<td>4.89</td>
<td>1.55</td>
<td>2.00</td>
</tr>
<tr>
<td>2010</td>
<td>8.57</td>
<td>5.04</td>
<td>1.50</td>
<td>1.76</td>
</tr>
<tr>
<td>2011</td>
<td>6.47</td>
<td>6.64</td>
<td>1.68</td>
<td>1.67</td>
</tr>
<tr>
<td>2012</td>
<td>8.89</td>
<td>5.40</td>
<td>1.80</td>
<td>1.95</td>
</tr>
<tr>
<td>2013</td>
<td>7.49</td>
<td>6.20</td>
<td>1.90</td>
<td>2.16</td>
</tr>
<tr>
<td>2014</td>
<td>8.85</td>
<td>6.33</td>
<td>2.19</td>
<td>2.35</td>
</tr>
</tbody>
</table>

**Table AI.** Macroeconomic indicators of Turkey and Brazil 1

<table>
<thead>
<tr>
<th>Macroeconomic indicators</th>
<th>GDP per capita ($)</th>
<th>Unemployment rate (%)</th>
<th>Total reserves ($bn )</th>
<th>Deposit interest rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Turkey</td>
<td>Brazil</td>
<td>Turkey</td>
<td>Brazil</td>
</tr>
<tr>
<td>2005</td>
<td>7,117.3</td>
<td>4,733.2</td>
<td>10.6</td>
<td>9.8</td>
</tr>
<tr>
<td>2006</td>
<td>7,727.3</td>
<td>5,809.1</td>
<td>10.2</td>
<td>10.0</td>
</tr>
<tr>
<td>2007</td>
<td>9,309.5</td>
<td>7,241.1</td>
<td>10.3</td>
<td>9.3</td>
</tr>
<tr>
<td>2008</td>
<td>10,382.3</td>
<td>8,700.6</td>
<td>11.0</td>
<td>7.9</td>
</tr>
<tr>
<td>2009</td>
<td>8,623.9</td>
<td>8,462.4</td>
<td>14.0</td>
<td>8.1</td>
</tr>
<tr>
<td>2010</td>
<td>10,111.5</td>
<td>11,124.1</td>
<td>11.9</td>
<td>6.7</td>
</tr>
<tr>
<td>2011</td>
<td>10,584.2</td>
<td>10,042.2</td>
<td>9.8</td>
<td>6.0</td>
</tr>
<tr>
<td>2012</td>
<td>10,646.0</td>
<td>11,922.7</td>
<td>9.2</td>
<td>5.5</td>
</tr>
<tr>
<td>2013</td>
<td>10,975.1</td>
<td>11,711.7</td>
<td>9.7</td>
<td>5.4</td>
</tr>
<tr>
<td>2014</td>
<td>10,529.6</td>
<td>11,384.6</td>
<td>9.9</td>
<td>4.9</td>
</tr>
</tbody>
</table>

**Table AII.** Macroeconomic indicators of Turkey and Brazil 2
About the authors

Guler Aras is Professor of Finance and Accounting at Yildiz Technical University (YTU), Istanbul, Turkey. She is Founding Director of the Center for Finance, Governance and Sustainability (CFGS) at YTU. She served as Visiting Professor at Georgetown University McDonough School of Business, Center for Financial Markets and Policy. She is Editor of Routledge Book Series Finance, Governance and Sustainability: Challenges to Theory and Practice and Gower Book Series Corporate Social Responsibility. Professor Aras has published more than 20 books and has contributed over 200 articles in academic, business and professional journals and magazines in finance, banking, capital market, sustainability, corporate social responsibility, corporate governance and integrated reporting. Guler Aras is the corresponding author and can be contacted at: aras@yildiz.edu.tr

Filiz Mutlu Yildirim is PhD Candidate at Yildiz Technical University (YTU), Istanbul, Turkey. Her PhD research is in integrated reporting.

Table AIII. Macroeconomic indicators of Turkey and Brazil 3

<table>
<thead>
<tr>
<th>Years/countries</th>
<th>Macroeconomic indicators</th>
<th>Current account balance (% of GDP)</th>
<th>Gross domestic savings (% of GDP)</th>
<th>Foreign direct investment, net inflows (% of GDP)</th>
<th>Foreign direct investment, net outflows (% of GDP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkey</td>
<td>Brazil</td>
<td>Turkey</td>
<td>Brazil</td>
<td>Turkey</td>
<td>Brazil</td>
</tr>
<tr>
<td>2005</td>
<td>4.44</td>
<td>1.57</td>
<td>16.49</td>
<td>20.83</td>
<td>2.08</td>
</tr>
<tr>
<td>2006</td>
<td>6.00</td>
<td>1.23</td>
<td>17.14</td>
<td>20.75</td>
<td>3.80</td>
</tr>
<tr>
<td>2007</td>
<td>5.84</td>
<td>0.11</td>
<td>15.91</td>
<td>21.35</td>
<td>3.41</td>
</tr>
<tr>
<td>2008</td>
<td>5.50</td>
<td>1.66</td>
<td>17.35</td>
<td>21.61</td>
<td>2.72</td>
</tr>
<tr>
<td>2009</td>
<td>1.95</td>
<td>1.46</td>
<td>13.83</td>
<td>18.55</td>
<td>1.40</td>
</tr>
<tr>
<td>2010</td>
<td>6.20</td>
<td>2.14</td>
<td>13.97</td>
<td>20.76</td>
<td>1.24</td>
</tr>
<tr>
<td>2011</td>
<td>9.68</td>
<td>2.01</td>
<td>14.89</td>
<td>21.04</td>
<td>2.09</td>
</tr>
<tr>
<td>2012</td>
<td>6.15</td>
<td>2.25</td>
<td>14.97</td>
<td>18.99</td>
<td>1.68</td>
</tr>
<tr>
<td>2013</td>
<td>7.85</td>
<td>3.39</td>
<td>14.10</td>
<td>18.35</td>
<td>1.51</td>
</tr>
<tr>
<td>2014</td>
<td>5.82</td>
<td>4.43</td>
<td>15.93</td>
<td>17.32</td>
<td>1.57</td>
</tr>
</tbody>
</table>
Retail supply chain responsiveness
Towards a retail-specific framework and a future research agenda

Erik Sandberg
Department of Management and Engineering, Linköping University, Linköping, Sweden, and
Hamid Jafari
Department of Industrial Engineering and Management, Jönköping University, Jönköping, Sweden

Abstract
Purpose – The purpose of this paper is to review existing research on retail supply chain responsiveness, develop categories to be included in a retail-specific responsiveness framework, and identify future research areas within the scope of retail supply chain responsiveness.

Design/methodology/approach – This paper presents an inductive systematic literature review of 46 academic, peer-reviewed articles. Based around the two major review questions on retailers’ role in the creation of supply chain responsiveness and future research areas, an inductive, qualitative, content analysis was conducted. Further analysis was conducted by using the software NVivo 11.

Findings – Existing research are grouped into a framework of four categories that together span the existing research. The categories are labelled supply chain orchestration, market orientation, supply chain operations and supply management. Two to three subthemes in each category are presented. Thereafter, promising future research areas are outlined, covering methodological issues, theoretical underpinnings, inclusion of context variables and outcomes of retail supply chain responsiveness.

Research limitations/implications – The conducted systematic literature review has been limited to academic, peer-reviewed articles.

Practical implications – The findings of the paper constitute a promising initial step towards a retail-specific framework on retail supply chain responsiveness.

Originality/value – The paper questions the comprehensiveness of established models in responsiveness, and argues that existing “general” literature on supply chain responsiveness gives little guidance and structure to retailers’ specific role and involvement in supply chain responsiveness. In particular, the paper focuses on the retailers’ role for creation of supply chain responsiveness, which has not previously been addressed in research.

Keywords Supply chain, Flexibility, Responsiveness, Agility, Retail

Paper type Literature review

1. Introduction
Today’s competitive business landscape requires addressing changes in market and customer requirements promptly and properly. Much academic effort in supply chain management (SCM) has been devoted to addressing the challenges caused by competition, environmental uncertainty and dynamism, as well as meeting extremely demanding customer requirements (Gligor, 2016; Kozenkova et al., 2015; Lau, 2012). To tackle these challenges, responsiveness has become a major research topic in logistics and SCM, and operations management literature, and considered as a key concept which significantly contributes to sustainable competitive advantage (Reichhart and Holweg, 2007; Singh, 2015; Hoyt et al., 2007). Responsiveness corresponds to “the ability to respond and adapt time-effectively based on the ability to ‘read’ and understand actual market signals” (Catalan and Kotzab, 2003, p. 677) from some sort of external stimuli – mainly customer and market demand (Reichhart and Holweg, 2007).

Very similar to that of flexibility, the majority of the body of the literature on responsiveness pertains to the manufacturing setting since manufacturing is considered as
the main source for creation of responsiveness in the supply chain. Other supply chain members such as retailers, distributors or service providers are generally neglected in literature (Jafari, 2015; Chaudhuri, 2008; Sohel Rana et al., 2015) and as a result, retailer’s role for the creation of responsiveness has not yet been explored in detail. Thus, a fundamental starting point for this research is that although manufacturing still holds a key role in creating responsiveness, the other supply chain members deserve more research attention, given the overall developments of supply chains.

What is evident in the structure of supply chains over the course of the recent decades is that the distribution and downstream side, especially retailers, have come to the forefront of attention as they have obtained an unprecedented power (Geylani et al., 2007; Sandberg, 2013). The unique position of retailers in supply chains in confronting shoppers and suppliers, in the context of extreme environmental uncertainty, calls for revisiting conventional management approaches (Reynolds et al., 2007; Ganesan et al., 2009), and responsiveness of retailers is no exception. The fact is, retailers face individual shoppers and consumers who take higher quality, on-shelf availability, wide assortments, speedy delivery, after-sales services, fun shopping experience and lower prices for granted. Internet and mobile services such as price-comparison apps and social media reviews have made it easier for shoppers to cross-shop in a transparent omni-channel setting. Accordingly, retailers need to be even more responsive to handle for instance shopper reactions to new merchandise, early seasonal weather changes and last-minute promotions to meet quarterly sales goals (Thomas et al., 2010). Altogether, this development calls for improved capabilities of retailers to create responsiveness in their supply chain.

Despite the relevance and importance of responsiveness, and its subcategory retail supply chain responsiveness that is the focus of this research, there is still confusion about the actual meaning and operationalization of responsiveness that goes beyond superficial definitions (Holweg, 2005; Hoyt et al., 2007). Furthermore, several scholars have questioned the existing research on responsiveness for being confined to specific methods. For instance, Seth and Panigrahi (2015, p. 1071) state that supply chain responsiveness is “much more than inventory management as wrongly exaggerated by modelling-based studies”. The same type of concerns seems to be relevant for several similar concepts such as flexibility, agility and adaptability (Reichhart and Holweg, 2007; Jafari, 2015).

The purpose of this paper is to review existing research on retail supply chain responsiveness and develop categories to be included in a retail-specific responsiveness framework. In addition, a future research agenda on retail supply chain responsiveness is outlined. We question the comprehensiveness of established literature on responsiveness and argue that existing “general” literature on supply chain responsiveness gives little guidance and structure to retailers’ specific role and involvement in supply chain responsiveness. In particular, we focus on the role of retailers in the creation of supply chain responsiveness. In light of the increasingly demanding and dynamic end customer requirements, responsiveness that stems from the retailers is an important and vital complement to responsiveness that is created further upstream in the supply chain, i.e. at the manufacturers. In terms of contribution, our literature review results in outlining four categories of responsiveness that could function as a first step towards a retail-specific responsiveness framework. In contrast to other existing general frameworks, such a framework brings forward aspects other than manufacturing as important for creating responsiveness. For practitioners, this could inform and improve decision making in the retail supply chain environment.

The remainder of this paper starts with a presentation of our methodology approach which has been an inductive systematic literature review (Denyer and Tranfield, 2009; Tranfield et al., 2003). Thereafter, a descriptive quantitative overview of the reviewed articles and a thematic analysis of their content are provided. Finally, future research topics and some final remarks conclude the paper.
2. Methodology

In order to avoid preconceptions and provide an independent, exhaustive review of existing research in the area of retail supply chain responsiveness, this study has adopted a systematic literature review approach. Given the confusion and relatively sparse existence of relevant research, in line with our exploratory purpose, the review can be described as inductive. The review process can be described in accordance with Denyer and Tranfield’s (2009) approach, following the five steps of question formulation, locating studies, study selection and evaluation, analysis and synthesis and reporting and using the results.

2.1 Step 1: question formulation

A first important step of this research was, based on the purpose of the research, to develop more precise inquiries that could enable the necessary focus and direction of the systematic literature review process. As a major review question, we have focussed on the retailers’ responsiveness creation in the supply chain. More specifically, this typically incorporates strategies, role as well as specific tools available for creation of supply chain responsiveness. A secondary review question has been to identify future research topics suggested in the reviewed articles. Beside these two review questions, the authors also decided early in the research process to identify in each reviewed article the following issues:

- formal publication information: author(s); year of publication; title; volume/no; and journal; and
- content-related information: purpose; level of analysis (firm level or supply chain level); paper type (empirical or theoretical); methodology approach; retail sector(s) discussed; theoretical basis for the paper; general findings/conclusions; and suggested future research topics.

In parallel to the review question formulation phase, another important part at this stage of research was a scoping study (Tranfield et al., 2003), where five articles (Holweg, 2005; Hoyt et al., 2007; Jafari, 2015; Reichhart and Holweg, 2007; Singh, 2015) presenting general models or frameworks of supply chain responsiveness (or flexibility) were reviewed by the authors. The purpose of this scoping study was to better understand the larger research field of overarching (i.e. including all supply chain members) supply chain responsiveness. An overall understanding of the academic definitions of responsiveness and related terms – such as flexibility and agility – were also attained, which provided this research with valuable input for constructing the search strings.

2.2 Step 2: locating studies

Based on the scoping study, three search strings were constructed with “responsiveness”, “flexibility”, respectively, “agil*”, each term combined with “retail”. Due to the scarcity and fragmentation of existing research on retail supply chain responsiveness it was decided to keep the search strings open and inclusive, with relatively few limiters. However, in order to ensure the quality of the articles, being “academic research” was used a major limiter; as a result, white papers, consultancy reports and books have consciously been outside the search scope for this study. The academic, full text, peer-reviewed journal articles have been targeted in the two databases of Business Source Premier (Ebscohost) and Scopus. Together, these two databases were considered to have valid coverage of the supply chain literature, including journals from, e.g., Taylor & Francis, Emerald, Elsevier, Inderscience and Springer.

2.3 Step 3: study selection criteria

Figure 1 shows the selection process from 677 hits to the 46 articles that were read in full.
After article selection and deletion of duplications (98 articles) the abstracts of the remaining 579 articles were all scanned as a means to identify the articles clearly outside the scope of our research.

To ensure reliability and validity, all abstracts were reviewed by both authors. The exclusion or inclusion of the articles was discussed and the precise selection criteria could be developed and refined. During this process another 500 articles could be removed (from 579 articles to 79). Major reasons for removing articles were: lack of focus on retail supply chain responsiveness (i.e. the key terms had been used clearly in another context and given another meaning), poor retail focus, i.e., the article discussed for instance the manufacturer’s flexibility rather than the retailer’s, and specific focus on the banking sector in which customer responsiveness appears to be a major research topic as well.

The relatively large amount of deleted articles was expected due to the broad search strings. The construction of more well-defined search strings including terms such as “supply chain” or “logistics” would certainly have, on the one hand, narrowed down the initial amount of hits, but, on the other hand, increased the risk for excluding relevant articles. In all, 579 articles were considered as a reasonable amount of articles possible to review at an abstract level.

The 79 articles were then read through in full text independently by both researchers. Another 33 articles were judged to be outside the scope of this research, mainly due to their lack of focus on retail supply chain responsiveness. As a result, the selection process provided the researchers with 46 relevant articles. To quantitatively measure the level of agreement of the two involved researchers, the measure of Cohen’s κ (Lombard et al., 2002) was calculated to 0.80 during the last, critical review round. Landis and Koch (1977 in Seuring and Gold, 2012) suggest that if the value of Cohen’s κ is less than 0.60 the discrepancy among the reviewers might be too high and the data material be recoded.

2.4 Step 4: analysis and synthesis

During the last step of the selection process, the articles were read through carefully and their contents were individually scrutinized and broken down into constituent parts that were inserted in a data extraction sheet in Excel. This coding process was first conducted individually by each researcher, and thereafter the authors’ two excel sheets were merged into one after joint discussion, following what Seuring and Gold (2012) label “discursive alignment of interpretation”. As such, during this process of merging the data into one single extraction sheet, reliability and validity were ensured (Seuring and Gold, 2012).

Overall, a content analysis approach as elaborated by Seuring and Gold (2012) was thereafter taken in the analysis. Given the fragmented and ill-defined research area targeted, i.e. retail supply chain responsiveness, we argue that content analysis is a suitable tool for strengthening and developing a more solid understanding of the area. A content analysis “represents an effective tool for analysing a sample of research documents in a systematic and rule-governed way” (Seuring and Gold, 2012, p. 546), and aims at an objective identification of the content in a data set (such as selected articles). It overlaps with the
concept of thematic analysis, which is mainly a qualitative method for uncovering different “themes” within a data set (Fugard and Potts, 2015; Jonsson and Tolstoy, 2014).

An inductive approach was applied during the content analysis. First, as a means to structure the reviewed articles, an overall categorisation into four groups depending on their supply chain focus was made, in the article labelled supply chain orchestration, supply chain operations (internal matters related to responsiveness), market orientation and supply management. Based on this overall categorisation, the contents of each category were thereafter further elaborated in a clustering process, inspired by Owen’s (1984) criteria of recurrence (of the same message with different words), repetition of words and terminology and forcefulness, i.e., how much a certain message was emphasised in an article. In particular, the latter criterion guided the analysis in this study. Major findings related to the review questions were highlighted in each article, and through an iterative process, the major messages were condensed beyond simple word or phrase counting, and related to the findings of other articles in the category. Later on, as a final step in the content analysis process, some further coding was conducted using the software NVivo 11 to categorise retail supply chain responsiveness in existing research, as well as to provide some insights on directions for future research from methodological and theoretical perspectives. Using qualitative data analysis software tools – such as NVivo – are increasingly being used in systematic reviews to increase to overall validity of such studies (Bazeley and Jackson, 2013).

Regarding the second review question pertaining to the future research areas, in the content analysis process, all potential areas for future research highlighted by the reviewed articles were compiled. Then, in an iterative process – and based on our own general judgement on what appears to be more relevant or worthy of further research – these areas were clustered into four major categories using NVivo. These categories are discussed in detail in Section 4. Therefore, a combination of objective and subjective approaches was used for coding and classification of the future research areas.

A descriptive analysis focussing on the quantitative aspects such as year of publication, methodology used, etc., was also conducted using Microsoft Excel. Next section starts with a brief overview of this descriptive analysis, before the content analysis is presented.

3. Findings
3.1 Overview of the reviewed articles
The results of the study show that articles addressing responsiveness in a retailing context have appeared in a myriad of academic journals. Half of the published articles are mere instances in the respective journals they have appeared in. As shown in Table 1, International Journal of Retail & Distribution Management has been the most popular forum with over 13 per cent of the articles published appearing in it.

As depicted in Figure 2, apparently, an increasing interest in responsiveness issues in retailing has been shown by researchers. The trend shows that from 1999, when the first

<table>
<thead>
<tr>
<th>Journal</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>International Journal of Retail &amp; Distribution Management</td>
<td>6</td>
</tr>
<tr>
<td>The International Journal of Logistics Management</td>
<td>4</td>
</tr>
<tr>
<td>International Journal of Production Economics</td>
<td>3</td>
</tr>
<tr>
<td>Industrial Marketing Management; Journal of Fashion Marketing and Management: An International Journal; Journal of Retailing; Production Planning &amp; Control; The International Review of Retail, Distribution and Consumer Research</td>
<td>2</td>
</tr>
<tr>
<td>Others (23 different journals)</td>
<td>1</td>
</tr>
</tbody>
</table>

| Table I. Journals of studied articles and frequency of appearance      |
|-----------------------------------------------------------------------|-----------|
| 1981                                                                  | 1         |
article was published, until 2005, on average only one article was published per year. However, this number raised to five articles in 2006, and a record number of six articles in 2014. Almost 70 per cent of the reviewed articles can be labelled as “empirical”, from which the majority apply a case study approach, see Table II. Interviews seem to be the most popular form of primary data collection. Interestingly, purely theoretical articles appear to be relatively unpopular, among which just over 14 per cent are of literature review type. Even those two literature review articles are carried out and reported in a generic fashion and lack a systematic character.

In terms of theoretical grounding, in total, 61 instances were found in which the scholars had addressed one or multiple specific bodies of literature or theories, see Table III. However, these literature fields or theories were in most cases vaguely applied and sparsely defined.

In conclusion, the general overview of the reviewed articles reveals that just as for the more general literature on responsiveness, also retail-specific research in the area is fragmented and lacks clear theoretical rigidity, origin and focus. A striking indicator of this is that there are relatively few articles with clear definitions of the targeted topic.

<table>
<thead>
<tr>
<th>Paper type (Total)</th>
<th>Method (Total)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical (32)</td>
<td>Multiple</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Interviews</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Secondary data</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Observation</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Simulation</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Content analysis</td>
<td>1</td>
</tr>
<tr>
<td>Survey (12)</td>
<td>Structural equation modelling</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Descriptive statistics</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Regression analysis</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>$t$-test</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Clustering</td>
<td>1</td>
</tr>
<tr>
<td>Theoretical (14)</td>
<td>Conceptual modelling</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Fuzzy logic</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Hypothetical scenarios</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>1</td>
</tr>
<tr>
<td>Literature review (2)</td>
<td>Generic</td>
<td>2</td>
</tr>
<tr>
<td>Mathematical modelling (7)</td>
<td>Optimisation</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Simulation</td>
<td>2</td>
</tr>
</tbody>
</table>

Table II. Method in the reviewed articles
Of the reviewed 46 articles, we have only found precise, explicitly stated definitions of (in the article) used key terms such as responsiveness, flexibility or agility in 16 articles. In the other articles, these terms are only vaguely described and at a very overall level.

3.2 Categories of retail supply chain responsiveness
The reviewed articles can be grouped into four loosely defined categories, each covering two to three main topics, see Figure 3.

3.2.1 Supply chain orchestration. A first category of papers concerns the retailers’ involvement in and efforts for the creation of responsiveness at a supply chain level. It is claimed that retailers must align an agile strategy in their supply chains, i.e., their entire supply chains must be coordinated and managed towards an agile strategy (Sohel Rana et al., 2015; Storey et al., 2005; Randall et al., 2011). Christopher et al. (2004) points out that in an agile supply chain, there is often a need for network arrangements along the supply chain, with a strong focal firm acting as an “orchestrator” of the supply chain. According to the article, large, international fashion retailers may be such firms, taking the role as champions of the supply chain, and hence achieve responsiveness in the supply chain.
In a similar vein, Randall et al. (2011) reports that retailers identify a future need to invest for further agility improvements, in processes, infrastructure, technology as well as people.

A recurrent message in this theme of literature, and related to the suggested network approach, is that responsiveness is created by the retailers mainly through relationship management upstream towards suppliers as well as downstream towards consumers. The replenishment process, in which the matching of supply and demand is conducted, is considered crucial. D’Avolio et al. (2015) focusses on this process, and argues that retailers’ partnerships and collaborations towards suppliers as well as customers concerning replenishment are fundamental for a supply chain’s ability to react timely to changes in customer desires.

For successful achievement of responsiveness towards end customers, it is important that all parties of the supply chain are coordinated and organised towards joint goals. For instance, Barnes and Lea-Greenwood (2010) investigate the concept of fast fashion and its resulting responsiveness at a store level, and demonstrate that at this stage of the supply chain, there is still a lot to do for the retailers in order to achieve supply chain responsiveness. The combination of rigid, centralised control and lack of resources at store level, results in a retail environment that is ineffective at the point of sale, and hampers the entire supply chain’s ability to respond to end customers’ demands. In order to avoid such obstacles, responsiveness at a supply chain level may be created by the launch of supply chain-wide SCM-based concepts such as quick response, Just-In-Time (JIT), efficient consumer response, collaborative forecasting, planning and replenishment, etc. (Bhardwaj and Fairhurst, 2010; Storey et al., 2005). Common for these practices is that they are expected to enhance responsiveness towards end customers through integration efforts focussing on collaboration, information sharing and trust among supply chain members.

3.2.2 Market orientation. Another group of articles are focussing on the market-oriented capabilities needed for creating supply chain responsiveness. Being the last member of the supply chain with a position towards end consumers, the retailers’ market-oriented capabilities may be of vital importance for the entire respective supply chains. Thus, retailers’ market orientation is a crucial starting point for proper supply chain responsiveness, and for such an orientation, an appropriate set of market sensing capabilities by the retailer is required. Important retailing strategies such as customer segmentation, pricing, etc., originate from these capabilities (Lau, 2012).

From the textile sector, Masson et al. (2007) acknowledge that the sophisticated efforts conducted by the retailers to identify emerging fashion trends are of significant importance for agile supply chain practices. In their paper, carefully specified market segments were identified and trends for each individual segment were outlined and monitored. Similarly, Christopher et al. (2004) conclude that apart from carefully monitoring point-of-sales data, examining emerging fashion trends and future customer demands is crucial.

Atapattu and Sedera (2014) point out that market sensing capabilities need to be well aligned with responding capabilities in order to achieve customer satisfaction. Thus, for retailers, a major task is to match these two types of capabilities with each other in order to achieve responsiveness. The need for alignment of capabilities is exemplified by Helo and Luomala (2011), who present a conceptual framework that illustrates the potential areas where consumer data can be integrated in the logistics planning process. To have knowledge about consumer behaviour, and have the ability to incorporate this knowledge into the logistics planning process, may facilitate proper market responsiveness activities.

Within the wider scope of research on market sensing capabilities, internationalisation and adaptation to local market requirements are discussed in a number of articles reviewed. A particular focus is given to the importance of understanding cultural peculiarities and the ability to adapt to these circumstances. A proper alignment to local culture is hence key for
the retailer’s supply chain responsiveness (Swoboda et al., 2014; Boulay et al., 2016). Okongwu and Santosa (2008) present some best practices that enable supply chains in France and Indonesia to maintain high level of efficiency, flexibility and responsiveness and demonstrate that cultural market differences results in different strategies for success. In a similar vein, Yu and Ramanathan (2012) investigate foreign retailers’ operations in the Chinese customer market and stress the need for adopting to local culture such as customer shopping habits. Franchising could be an option for increased local responsiveness according to Boulay et al. (2016). An important ingredient for successful adoption to local circumstances may be the ability of service innovation. For such service innovation, the retailers’ entrepreneurial proclivity is an important capability. The entrepreneurial proclivity and the resulting ability for market responsiveness are driven, in particular, by knowledge about the supplier-side, e.g., breadth of product offerings, pricing of products and existing and potential suppliers (Griffith et al., 2006). From a similar service innovation perspective, the Chinese domestic retailers’ successful response to foreign multinational companies’ entrance on the Chinese customer market are studied by Shi and Au-Yeung (2015). The authors argue that the Chinese retailers are smaller, independent and have a simpler organisational structure (compared to multinational, foreign companies). This makes them faster in responding to the rivals’ movement and China’s constantly changing retail market. Another strategy to meet competition from foreign competitors on the domestic market is to implement principles of a high reliability organisation. This is defined as organisations that “performs successfully in highly volatile environments by adopting flexible practices and the continuous pursuit of improvement and learning in its operations” (Ciravegna and Brenes, 2016, p. 4500). Ciravegna and Brenes (2016) demonstrates, based on such principles, how Selectos, the leading food retailer in El Salvador, succeeded to compete when WalMart entered their domestic market.

3.2.3 Supply chain operations. Another group of articles concerns a wide range of retail supply chain operations aimed at improving responsiveness. In common for this research is the focus on logistics planning and distribution as a facilitator and driver for improved supply chain responsiveness, e.g., replenishment of stores (Leung et al., 2003), IT tools, such as a dynamic picking system (Gong and de Koster, 2008), sharing of supply information with demand in a timely manner at retail outlets (Jain et al., 2009), centralised vs decentralised ordering processes (Ganesan et al., 2007) and design of ordering policies (order quantities) to enhance responsiveness (meet customer demand quantities) (Yu et al., 2011).

Although focusing on separate, narrow operational concerns, together, these articles demonstrate a wide range of tools and areas for retailers to improve their responsiveness towards customers. Inventory levels and associated information flows (ordering, safety stock policies, etc.) are typically in focus (i.e. responsiveness at this operational level is mainly considered to be enhanced through high product availability, which in turn requires adequate information sharing). From the reviewed literature, there still seem to be potential for further improvements in information sharing in retail supply chains. Masson et al. (2007) indicated that even if information sharing and related technologies were relatively well developed close to the customers in their study, they were less developed upstream in the studied supply chain. These shortages represent a future improvement area in which retailers may play an important role.

A frequently discussed aspect of information sharing in the reviewed articles is the use of RFID technology as an enabler for, among others, improved supply chain responsiveness. Major advantages associated with the use of RFID in the retail environment are improved cost performance, as well as responsiveness-related performance such as reduction of lead times, improvement of accuracy, on-shelf availability (Azvedo and Carvalho, 2012),
increased visibility, quicker flow of information, improved information sharing and easier and quicker identification of products (Azevedo and Ferreira, 2009). RFID is also expected to improve decision making since it provides management with information earlier (faster), and with higher quality (accuracy, relevancy, completeness, etc.) (Sellitto et al., 2007). A disadvantage of particular importance is still the costs associated with the technology (Azevedo and Carvalho, 2012; Azevedo and Ferreira, 2009) and lack of know-how in the organisation (Azevedo and Ferreira, 2009).

3.2.4 Supply management. A fourth theme in the reviewed literature is responsiveness as a part of the larger area of supply management, including topics such as implications of global sourcing and supplier relationship management. Global sourcing is often discussed as a necessary development for retailers, in which longer lead times, cultural differences and other barriers might jeopardise responsiveness of the retailers. The major trade-off related to global sourcing is the search for low purchasing prices on the one hand, and ability to respond rapidly to changed customer demands, which normally requires short lead times and flexibility, on the other. Lowson (2001, 2002) highlight the many illusive traps with global, off-shore sourcing in comparison to domestic sourcing. Lowson points out that from a cost perspective, there are apart from many hidden costs (e.g. delays, use of air freight, administrative and quality costs, etc.) also inflexibility costs that must be taken into consideration when sourcing globally. These inflexibility costs “involves issues such as longer lead-times and a general lack of flexibility as part of a response to demand changes” (Lowson, 2002, p. 80). In the article, Lowson (2002) constructs a model that assesses these inflexibility costs and takes them into account in the discussion whether off-shore sourcing or domestic sourcing is to be preferred.

As an alternative to global sourcing, proximity sourcing is in general associated with improved market orientation and ability to respond to changes in customer demands. For instance, Abernathy et al. (2006) conclude that in a rapidly changing industry such as fashion and textile industry (which is in focus in their article) proximity sourcing can be advantageous based on a design, marketing and production point of view. These potential advantages need, however, to be properly exploited by the retailer. In a similar vein, Barnes and Lea-Greenwood (2006) conclude that for many retailers proximity sourcing, improving geographical as well as cultural proximity towards the supplier may leverage higher ability for end customer responsiveness.

The retailers’ relationships towards suppliers and manufacturers are another theme discussed in the reviewed articles. Supply chain responsiveness concerns the entire supply chain and appropriate relationships in the supply chain could therefore be of crucial importance for responsiveness performance. Doyle et al. (2006, p. 279) mean that close relationships with suppliers are essential for responsiveness, and that retailers may “benefit from an informal series of strategic alliances that serve to create efficiency and agility simultaneously”. Coordination between the manufacturer and retailer that are based on quantity-flexibility contracts, in which retailer commits to order quantities given certain prices, could here serve as an important platform (Li et al., 2016; Bicer and Hagspiel, 2016).

Wong and Johansen (2006) study the relationship between a manufacturer and a so-called JIT retailer, i.e., a retailer with only inventory in store, which requires quick response processes from the manufacturers. The authors identify six specific coordination strategies for how to align the two companies and achieve responsiveness towards the retailer’s customers. In a similar vein, Richey et al. (2012) examine the importance of collaboration at an operational level between supplier and retailer in order to achieve time-based logistics advantages. The authors conclude that in order to achieve a successful, reciprocal collaboration, the retailer must be able to adapt to the supplier conditions. Another crucial factor is the retailer’s ability to invest in technology that is compatible with the supplier’s.
4. Future research areas

Addressing the second review question, this research has identified future research areas within the scope of retail supply chain responsiveness. Based on the directions provided by the reviewed articles, we categorise the potential areas for future research into methodological issues, theoretical underpinnings, inclusion of context variables and outcomes of retail supply chain responsiveness.

4.1 Methodology issues

The most common area for future research appears to relate to methodology concerns, in particular study design, research in different sub-sectors of retailing, and perspective.

In terms of study design, several scholars call for in-depth case studies or survey studies with larger sample sizes to shed light on the various aspects of responsiveness in the retailing context (e.g. Seth and Panigrahi, 2015; Doyle et al., 2006). Many contributions suffer from being limited to reporting perceptual information from a few managers within the retail firms. In this regard, Combe et al. (2012) and Adjei et al. (2009) stress the importance of including multiple informants per firm in such studies. Comprehensiveness as well as triangulation would strengthen the rigidity of the research results.

Although all retail sectors are different from each other, and “one size doesn’t fit all”, some articles reviewed suggest that there are potential learnings to be made across different sub-sectors of retailing. For instance, Wong and Hvolby (2007) see room for replicating their study, which was done in the toy retail supply chain context, to other volatile industries such as fashion clothing and ski-wear. Similarly, several retail sectors have been considered to be in the “infancy” phase, and hence deserve further attention, including fast fashion retailing (Barnes and Lea-Greenwood, 2006, 2010). Common context factors across retail sectors are also discussed, e.g., product characteristics. One can expect that the results of the existing research could be dependent on whether or not the products are innovative (Masson et al., 2007) or durable (Bowersox et al., 1999).

Existing research on responsiveness could also be criticised regarding unclear perspectives taken in the studies. In particular, the market perspective is indicated as a promising future research topic among the reviewed articles. Since retailers are the final links in the supply chains facing consumers, and the fact that responsiveness is in direct relation to the customers or markets, some scholars draw the attention to including a consumer perspective in future research (Atapattu and Sedera, 2014). Of particular importance from retail perspective is how retailers can leverage on their position close to the end customers, being “match-makers” between supply and customer demand. The called for market and customer perspectives are, however, unclear and the existing literature is vague when it comes to definitions and details about the suggested perspectives to be taken.

4.2 Theoretical underpinning

The findings from this study reveal that relatively few articles are anchored in more rigid theory. Here, opportunities for improved sharpness and further development are still open. Promising combinations that are briefly suggested in literature are dynamic capabilities (Griffith et al., 2006) and coordination theory (Wong and Johansen, 2006). Another interesting theoretical ground would be power. The question of power emerges both in articles on general supply chain strategy (e.g. Bourlakis et al., 2014) and in articles discussing supplier–retailer relationships.

In addition to the lack of grounding in more rigid theory, several supply chain initiatives and practices have been repeatedly underlined by the existing literature for further exploration in relation to supply chain responsiveness, including integration and IT, and postponement. A future research area that deserves particularly more attention from a retail research perspective is the integration of information and capacities among supply chain
members so that improved overall responsiveness of the entire supply chain can be achieved. The supply chain-wide orchestration of capacities and information is in the very heart of the traditional SCM philosophy and the retailers are in pole position to take the initiative for increased supply chain-wide responsiveness (Randall et al., 2011; Helo and Luomala, 2011; Swoboda et al., 2014). Particularly investments on RFID and its application seem to be a promising area for future research. While some scholars stress the room for exploring the functionality of RFID in retailing (Sellitto et al., 2007; Azevedo and Carvalho, 2012; Azevedo and Ferreira, 2009), others call for taking a profitability perspective to consider the fixed costs associated with RFID (Goebel and Günther, 2009).

4.3 Inclusion of contextual variables

Inclusion of contextual variables is another common topic discussed for future research among the reviewed articles. In line with the general lack of rigidity and superficial case study methodologies, context variables are often not in-depth discussed. To address this, Yu and Ramanathan (2012) suggest including various dimensions of environmental complexity in future research. One example of this is inclusion of the international nature of retail supply chains, where findings could be scrutinised in relation to different countries and cultures (Okongwu and Santosa, 2008). Similarly, from a market point of view, Lowson (2001) maintains that various demand behaviours should be considered. In a later work, the same author stresses the paucity of research in investigating the volatility of product categories (Lowson, 2002).

For appropriate inclusion of context variables, changes in the environment over time must be captured, and since dynamic capabilities evolve over time, Richey et al. (2012) and Ganeshan et al. (2007) underline the relevance of longitudinal studies. Related to the existing research on dynamic capabilities, a general criticism is also that dynamic capabilities – such as responsiveness – are studied in isolation. Therefore, Griffith et al. (2006) highlight the need for developing an integrative framework for holistically capturing the contextual factors and responsiveness capabilities.

4.4 Outcomes of retail supply chain responsiveness

Although the research on retail supply chain responsiveness is on the rise, few studies actually address the direct or indirect outcomes resulting from responsiveness in retailing. Specifically, supply chain performance resulting from responsiveness requires more exploration (Sohel Rana et al., 2015; D’Avolio et al., 2015). Shi and Au-Yeung (2015) emphasise the room for further research on the relationship between such competencies, market performance and competitive advantages. Another research topic that has already been covered in the existing research, but continuously calls for updating and improvement, is different means for quantification and comprehensive calculations on the total costs involved in global sourcing practices (Lowson, 2001, 2002), as well as the trade-off between these costs and service performance (Bygballe et al., 2012; Randall et al., 2011).

5. Discussion

This paper explored the relatively sparse academic research done on retail supply chain responsiveness. Methodologically, our study has applied a systematic literature review approach, which not only systematically and validly categorised and organised the existing literature, but also identified the main potential areas for future research. A particular attention has been given to the creation of supply chain responsiveness.

Overall, existing research on retail supply chain responsiveness can be described as fragmented and few cross-references occur among the 46 articles reviewed. The articles are in general provided with sparse theoretical groundings and have superficial results.
We contend that from a theoretical point of view, retail supply chain responsiveness is a relatively immature area, with similar deficiencies to that of the general supply chain responsiveness literature. For instance, a majority of the reviewed articles are vague when it comes to definitions of the key terms such as responsiveness, flexibility and agility (the keywords that were used in this research), and quite a few of the articles typically call for further in-depth studies.

As such, our findings signal that a way forward to overcome these deficiencies, and to enable new insights towards a more comprehensive understanding of retailers’ role would be to develop a retail-specific framework of supply chain responsiveness. The four categories and their contents outlined in this paper could serve as a first attempt towards such a framework. Together, the categories comprise the opportunities for retailers to create responsiveness and could therefore serve as a basis for a future more comprehensive development.

It should be noted that responsiveness frameworks are already available in the “general” supply chain responsiveness literature, although to a limited extent (e.g. Reichhart and Holweg, 2007; Hoyt et al., 2007). However, as argued in the introduction of this research, applying such general frameworks may not be appropriate in a retail setting, mainly due to the retailers’ lack of control of manufacturing capacity. The absence of manufacturing capacity gives that focus of retail supply chain responsiveness instead to be geared towards the end customer- and market-oriented efforts and capabilities. Indeed, retailers have a unique position close to end customers, which means a potential for rich information and in-depth understanding of end customer needs and requirements (Swoboda et al., 2014).

The ability to sense and capture the end customer demand, i.e., “being close to the customers” (Christopher et al., 2004) is often discussed as a core competence for the retailers (Masson et al., 2007; Christopher et al., 2004; Atapattu and Sedera, 2014). This could, in contrast to manufacturing capabilities, be considered as the major tool for retailers for how to create supply chain responsiveness.

Although the presented four categories regarding retail supply chain responsiveness constitute a promising initial step towards a retail-specific framework, several possibilities for further developments are still open. For instance, an effort towards such a retail-specific framework could be made by addressing the four identified future research areas suggested in our research, e.g., in empirical studies. We judge all four areas to be important ingredients in the work towards a framework, either applied in isolation or in combination. For instance, linking responsiveness creation in the four categories with expected outcomes and clarifying the role of different contextual variables would make a contribution towards a more comprehensive framework.

In addition, such a work could also increase the practical relevance of the initial research that is presented in this paper. In its current state the practical relevance of the study at hand is limited to an overview and categorisation of four important themes that together spans retail supply chain responsiveness. In order to improve practical relevance, future studies should for instance provide guidance for how individual retailers in different contexts could develop and exploit different sources to retail supply chain responsiveness.

Beyond developing a specific framework, we also argue that the identified future research areas could serve as an important injection and encouragement towards more research in the area in general. The identified topics would all help to strengthen retail supply chain responsiveness as a research area and enable a ground for further structuring of existing fragmented findings. In addition, a major limitation with the research presented in this paper is its concentration on scholarly academic articles, further restricted by the language and publication format. Therefore, a future research opportunity would be to cover other type of publications (including books), and perhaps benefit from cross-comparing research results from different areas (e.g. extending the context to retail banking and services retailing).
References
Retail supply chain responsiveness


About the authors
Erik Sandberg works as Associate Professor at the Department of Management and Engineering, Linköping University. His research interests cover retail, supply chain management, strategy and the strategic role of logistics in the strategy of the firm. Erik Sandberg is the corresponding author and can be contacted at: erik.sandberg@liu.se

Hamid Jafari works as Assistant Professor at the Department of Industrial Engineering and Management, Jönköping University. His research focusses on the overlap of marketing and logistics, mainly in the retailing industry.
Human resource management, commitment and performance links in Iran and Turkey

N. Gökhan Torlak
Independent Scholar, Lansing, Istanbul, Turkey

Cemil Kuzey
Department of Computer Science and Information Systems, Arthur J. Bauernfeind College of Business, Murray State University, Murray, Kentucky, USA, and

Maral Ragom
Department of Sales Operations, ATU Duty Free, Istanbul, Turkey

Abstract

Purpose – The purpose of this paper is to investigate the influences of human resource management (HRM) practices on organizational commitment (OC), individual job performance (IJP) and organizational performance (OP) and the effects of OC on IJP and OP at travel agencies in Iran and Turkey. Given research site observation, improving HRM practices fostering OC, IJP and OP was essential to high service quality in hospitality industry.

Design/methodology/approach – Data were collected through a survey based on interview and e-mail from 440 employees working at travel agencies in both countries. Analysis methodology included demographic statistics, descriptive statistics, factor analysis, reliability analysis, correlation analysis and multiple regression analysis.

Findings – Only reward (RE), teamwork (TW), job description (JD), delegation (DL) and career management (CM) influenced affective commitment (AC); JD and CM affected continuance commitment (CC); and recruitment and selection (RS), RE, salary and wage (SW), TW, DL, job security (JS) and CM influenced normative commitment (NC) significantly and positively. Likewise, only RS, TR, SW, TW, DL and CM affected IJP; and TR, JS and CM affected OP significantly and positively. Finally, AC, CC and NC influenced IJP and only AC and CC impacted OP significantly and positively.

Research limitations/implications – Sample and number of respondents were limited. Generalization of findings was not possible. E-mail respondents raised the problem of who filled the form. Scales did not include personality traits, socio-cultural features and economic conditions which should be explored. However, the research provided recommendations to travel agencies in both countries.

Originality/value – This study filled the gap in hospitality industry in Iran and Turkey where research works into wide-ranging HRM practices–OC–IJP–OP links were few and far between.

Keywords – Organizational commitment, Organizational performance, Human resource management practices, Individual job performance, Travel agency

Paper type – Research paper

1. Introduction

Human resource management (HRM) practices assist managers in achieving organizational goals through effective management of work and people in organizations (Boxall and Purcell, 2003; Guest, 2011). HRM practices and their implementation can be considered as a system for management of organizational members (Guest, 2011). HRM practices also make possible human development and bind people to the organization. Employees can perform better and have growth within the organization through greater loyalty and increased efforts to aid the organization’s welfare state (Rhoades et al., 2001). Guest (2011) found that HRM controlled organizational performance (OP) through the way it affected employee attitudes and behavior. If HRM processes were available and well-organized and organizational members experienced them, employees would respond to HRM practices positively. This resulted in getting...
work-related well-being” including “health,” “happiness” and “relationships” for employees. Health related to physical well-being and safety. Happiness referred to job satisfaction, contentment and engagement. Finally, relationships included fairness, trust, openness, friendship and freedom from bullying and harassment (Grant et al., 2007). Furthermore, Bakker and Demerouti (2007) noted that work-related well-being required balancing demands of work – workload, emotional demands and work and non-work conflict – and resources of work – job autonomy, social support, development opportunities and feedback. If right HRM practices were provided leading to necessary resources, the feelings of physical discomfort would be alleviated irrespective of the level of job demands. These, in turn, enhanced employee well-being and consequently affected work attitudes, such as organizational commitment, individual job performance (IJP), OP and intention to quit the organization. Obviously, all organizations need HRM practices. However, they are not likely to take this issue seriously unless they observe HRM’s certain helpful impacts, such as employee job satisfaction, employee professional development and organizational competitiveness (Guest, 2002).

In this study, Zahra Ragom as an Intern and Researcher observed some specific difficulties in Iranian and Turkish travel agencies. Screening and hiring were conducted unfairly; training programs were improperly designed and administered, bonuses and opportunities for advancement and personal growth were not provided; pay was inequitable and inadequate; teamwork was rare and high-handed; role ambiguity was high; responsibility and authority were not granted; job security was unavailable; and employee promotions and development opportunities were unjust. These concerns led to downgraded employee expectations, lost commitment and high employee turnover. So, the authors assumed that these particular HRM practices in Iran and Turkey could be considered important for tourism organizations and, thus, were selected for this research. Furthermore, in 2016 there was a disastrous drop in the number of tourists visiting Turkey compared to previous 10 years (Statista, 2017) and the number of tourists visiting Iran has not increased significantly (World Bank, 2017). The authors argued that although Iran and Turkey had liberal policies and bilateral agreements, historical and cultural properties and controlled geographical areas acting like a point of passing among Europe, Middle East and Asia, these countries experienced political and economic uncertainty. The authors presumed these causes could be a part of the whole problem and effective use and integration of HRM practices leading to commitment and high performance at travel agencies should also be explored. However, most research works in the literature selected HRM practices individually and few treated them extensively. Thus, this study uniquely investigated the relationships among wide-ranging HRM practices – OC – performance at travel agencies in both countries. Specifically, the paper handled nine HRM practices – recruitment and selection (RS), training (TR), reward (RE), salary and wage (SW), teamwork (TW), job description (JD), delegation (DL), job security (JS) and career management (CM); three types of OC – affective commitment (AC), continuance commitment (CC) and normative commitment (NC); and two types of performance – IJP and OP. Thus, the study aimed to understand whether the owners/managers of travel agencies would be aware of developing a sincere and thoughtful approach that treated specific HRM processes cautiously resulting in happy, productive and committed employees who could increase organizational effectiveness and preserve service quality in the hospitality industry.

The paper used demographic and descriptive statistics to evaluate data collected from the employees working at travel agencies in Iran and Turkey. Exploratory factor analysis and reliability analysis examined data. Pearson correlation analysis tested relationships between dependent and independent variables. Finally, multiple regression analysis tested hypotheses.

The paper described HRM practices, OC, OP and IJP and listed significant contributions among HRM, OC, IJP and OP relationships in various fields in the literature; offered
14 hypotheses; explained research method; examined data through aforementioned statistical methods; and concluded with a discussion of findings and limitations of this study shedding light for further future research.

2. Theoretical framework and hypotheses

2.1 Human resource management (HRM) practices

HRM systematically integrates characteristics of employees and financial resources in the organization that aid all functions optimized, which lead to the achievement of organizational objectives resulting in a competitive advantage. HRM should utilize people most productively for the goals of the organization and meeting employees' needs and professional development (Palmer and Winters, 1993).

Recruitment and selection (RS). Recruitment generates a pool including qualified people to employ. However, selection is the process whereby specialists choose applicants from this pool depending on HRM strategies and the organization itself. Recently, organizations used their own RS processes that are compatible with organizational values (Bratton and Gold, 2007).

Training (TR). TR is a process through which skills, talent and knowledge of an employee are enhanced and increased. It should take place only when the need and objectives for such TR leading to effectiveness and productivity have been identified. Employees who are at early career stage need TR to grow on the job and become a manager. While TR is for a specific job purpose, development covers activities both improving job performance and personality (Saad et al., 2013).

Reward (RE). RE refers to tangible benefit or fringe compensation that employee receives as part of employment relationship (Trevor, 2008). RE is a powerful incentive for improving employee and unit performance. RE signifies how employees are valued and can also generate high levels of employee satisfaction. Both intrinsic RE, such as job enrichment, autonomy and delegation and extrinsic RE, such as stock option, bonus, gain sharing, promotion and benefits enhance performance and satisfaction. RE should be available, timely, performance contingent, durable and visible in order to motivate desired behavior (Scott et al., 1988).

Salary and wage (SW). SW is the sum of remuneration paid to employees as a result of their physical and mental effort that contribute to the functions of management. Organization is responsible for instituting and administering pay system meeting the needs of employees. Employees feel as being a meaningful part of the organization through their special contributions, significant achievements and incentives. Satisfactory SW encourages employees to obtain the desired level of performance (Agarwal and Ferratt, 1999).

Teamwork (TW). TW generates synergy – “the whole is greater than sum of its parts” – within the organization easing goal achievement (Senge, 1990). The need for cooperation among organizational units necessitates TW, which leads to improvements in productivity, communication and processes; creativity in problem solving; better use of organizational resources; making good decisions; and producing high-quality products and services (Robbins and Finley, 1995). TW eases “team learning” emphasizing “enriched debate,” “collective thinking and act” and dialogue that put aside members’ ambitions and passion. TW also pools skills and abilities to solve difficulties and provides opportunities to catch best performance in organizations (Senge, 1990).

Job description (JD). JD includes essential functions of job and job specifications or competencies enabling an individual to be successful in the position. JD should manifest assigned duties and responsibilities. However, JD does not involve rules, procedures, work objectives or desired abilities and personal attributes. A JD might be an inefficient tool for managing human resources if it has some missing parts for employees regarding planning, communication, control and decision-making responsibilities, self-management obligations and unplanned assignments (Grant, 1997).
Delegation (DL). DL is the assignment of operational and managerial responsibilities along with essential authority by a superior to a subordinate who is obliged to perform responsibilities orderly. DL occurs when the right to perform the job is given to delegatee who accepts this duty to execute the task. However, DL is not a process of resignation from the job. The delegator remains accountable for the overall performance of the subordinate who has been given authority (Koçel, 2005).

Job security (JS). The effect of JS on job satisfaction is significant for both males and females (Nikolaou et al., 2005). Besides, the impacts of job insecurity on job satisfaction, work behaviors (i.e. organizational citizenship behavior and deviant behavior) and negative emotions (i.e. anxiety, anger and burnout) indicate that job insecurity is negatively related to satisfaction which leads employees to behave counterproductive to organizational purposes (Reisel et al., 2010). Furthermore, research on cadets’ motivation to become police officers in China uncovered that JS among others was an important motivation factor for cadets to join the police force (Wu et al., 2009).

Career management (CM). Most advanced companies develop and apply an integrated CM system which is beneficial both for the organization and employees (Antoniu, 2010). Companies should encourage employees to gradually develop own occupational concept as a result of skills, abilities, needs, motivations and aspirations and to link individual needs and ambitions with organizational needs and opportunities. Moreover, CM specialists should evaluate, advise and inform staff on career planning and ensure individual development efforts with TR and development programs (Popescu Neveanu, 2003).

2.2 Organizational commitment (OC) and its antecedents
OC signifies an employee’s loyalty to, identification with and involvement in the organization, which is vital for competing in volatile environments. Committed employees ensure that outputs are of high quality and innovative, responsive to changes in customer demands, not harmful to community (Robbins et al., 2004).

Cook and Wall (1980) noted three elements of OC, namely, identification, involvement and loyalty. Identification comprised connection and pride of employees toward organization. Involvement included an employee’s effort for the company beyond personal gain and desire to help beyond regular time. Loyalty related to employee’s intentions to quit due to a better compensation offered elsewhere. Mowday et al. (1982) also evaluated two forms of OC, namely, attitudinal and behavioral. The former treated the processes by which people began to think about their relationship with the organization. The latter related to the processes by which people were locked to the organization and tried to handle this issue.

Meyer and Allen (1991) proposed that AC, CC and NC stressed a psychological state that characterized the employee’s relationship with the organization and had implication for preserving membership in the organization. AC referred to the employee’s emotional attachment to, identification with and involvement in the organization. Strong AC made employees continue with the organization. CC referred to an awareness of the costs associated with leaving the organization. Employees with strong CC stayed in the organization because they saw it as beneficial regarding costs and rewards. NC reflected a feeling of obligation to continue employment. Employees with strong OC felt that they ought to remain with the organization deriving from a “psychological contract” between the employee and the organization that highlighted reciprocal obligation in an exchange relationship.

Furthermore, although HRM practices had direct impact on OC (Subsection 2.4) some researchers suggested that HRM practices also had indirect relations with OC. For example, Kinicki et al. (1991) noted that perceptions of organization’s commitment...
to employee training and promotion affected HRM practices and work attitudes. Likewise, Koys (1991) found that employees' belief in organization's desire to screen and hire good employees and treat them fairly affected employee commitment to the organization. The assumption is that a specific execution of HRM practice will not ensure OC. However, the organization's enthusiasm to HRM will be more influential than practices (Guzzo and Noonan, 1994; Meyer and Smith, 2000).

2.3 Individual job performance (IJP) and organizational performance (OP)
Performance addresses the use of simple outcome-based financial indicators reflecting the fulfillment of the economic goals of the firm (Venkatraman and Ramanujam, 1986). However, performance meant both behaviors and results. Behaviors emanated from the performer and transformed performance from abstraction to action. Behaviors were the product of mental and physical effort applied to tasks and could be judged apart from results (Brumbrach, 1988). In managing performance both behavior and results are considered (Armstrong, 2006).

IJP and OP quantify effectiveness and efficiency of action with a set of metrics (Neely et al., 2005). The measures and indicators act as surrogates for organizational phenomena. Conceptually, OP is defined as the comparison of the value produced by a company with the shareholders expected to receive from the company (Henri, 2004). The shareholders approach to performance emphasizes that organizations can create and maintain sustained competitive advantage through rare, scarce, inimitable and valuable internal resources (Barney, 1991). Human resources can increase the human capital pool through training, rotation and coaching and influence employee behavior in the desired direction (Boxall and Purcell, 2003). Best practices resulting in sustained competitive advantage for the organization are possible if employees have willingness to stay and put in extra effort in the organization. However, stakeholders' approach to performance highlights that survival of an organization does not require only financial competitiveness. Organization's ability to legitimize its existence toward both inside and outside stakeholders through fairly treating individual employee values, beliefs and interests is also significant (Paauwe, 2004).

In a holistic view, performance measurement as a learning device helps strategic plans and organizational objectives (Henri, 2004). Campbell (1999) defined performance as behavior relevant to the attainment of an organization's verifiable goals. It aimed to create a high performing culture in which individuals and teams were responsible for constantly improving skills, processes and achieving targets. In particular, management performance can be expressed as the approximation of individual objectives of employees with organizational objectives only if employees support the culture of the organization. It provides for defined and agreed expectations concerning role responsibilities and accountabilities, skills and behaviors (Armstrong, 2006). Performance is measured through financial outcomes – profits, sales and market share; organizational outcomes – productivity, quality and efficiency; and human resource-related outcomes – attitudinal and behavioral effects among employees, such as satisfaction, commitment and quit (Dyer and Reeves, 1995). However, the majority of studies ignored the last one (Guest, 1997).

Employee performance appraisal aims to evaluate, compare and provide feedback on employee performance and manage human resources through revealing raises, promotions, training needs, etc., in organizations (Spence and Keeping, 2011).

2.4 Literature review
Tables I–III illustrate studies about relationships among HRM practices, OC, IJP and OP in terms of researchers, themes and findings.
Research works into relationships between HRM practices and OC. Given Table I, most researchers uncovered that HRM practices had significant and positive impacts on three antecedents of OC. Thus, $H1\text{--}H1c$ have been developed (Subsection 2.5).

Research works into relationships between HRM practices and IJP and OP. Given Table II, HRM practices directed processes that enhanced employee skills, job-related experience and ability, and thus focused on developing valuable, rare, inimitable and organized

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Themes</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen and Meyer (1997), Bartlett (2001)</td>
<td>Management’s timely TR and development programs and opportunities</td>
<td>More loyal and morally obliged employees who stay in the organization</td>
</tr>
<tr>
<td>Trevor et al. (1997), Pfeffer (1986)</td>
<td>Associations between SW, RE and AC, CC</td>
<td>Strong</td>
</tr>
<tr>
<td>Guest (1997), Fey et al. (2009)</td>
<td>Provision of career development opportunities to employees generating justice and righteousness</td>
<td>More committed personnel</td>
</tr>
<tr>
<td>Appelbaum et al. (2000)</td>
<td>Consideration of skills, knowledge and experience of employees in RS</td>
<td>Productive and loyal workforce leading to competitive high performing organizations</td>
</tr>
<tr>
<td>Meyer and Smith (2000)</td>
<td>Associations between JS and AC and NC</td>
<td>Strong and positive</td>
</tr>
<tr>
<td>Conway (2003)</td>
<td>Associations between HRM practices and CC and NC</td>
<td>Positive</td>
</tr>
<tr>
<td>Gelade and Ivery (2003)</td>
<td>Associations between HRM practices and OC in retail banking sector</td>
<td>Significant</td>
</tr>
<tr>
<td>Agarwala (2003)</td>
<td>Associations between innovative HRM practices and OC</td>
<td>Strong and positive</td>
</tr>
<tr>
<td>Paul and Anantharaman (2004)</td>
<td>Associations between career development opportunities and extensive TR and OC</td>
<td>Strong</td>
</tr>
<tr>
<td>Shahnuwaz and Juyal (2006)</td>
<td>Associations between HRM practices and OC in consultancy/research and fashion industries</td>
<td>Strong</td>
</tr>
<tr>
<td>Fiorito et al. (2007)</td>
<td>Association between DL and OC</td>
<td>Positive</td>
</tr>
<tr>
<td>Kun et al.’s (2008)</td>
<td>Associations between TR and development, RS, SW and OC in IT sector</td>
<td>Positive</td>
</tr>
<tr>
<td>Chew and Chan (2008)</td>
<td>Associations between remuneration, recognition and undertaking challenging employee assignments and OC</td>
<td>Significant</td>
</tr>
<tr>
<td>Gellatly et al. (2009)</td>
<td>Associations between development-oriented, stability-oriented and reward-oriented HRM practices and OC</td>
<td>However, training and career development were not significantly related to OC</td>
</tr>
<tr>
<td>Giauque et al. (2010)</td>
<td>Associations between HRM practices and commitment of knowledge workers in Swiss SMEs</td>
<td>Organizational support, procedural justice and reputation of the organization affected employees’ commitment</td>
</tr>
<tr>
<td>Nassar (2017)</td>
<td>Associations between RS, TR, performance appraisal, SW and communication and AC at four- and five-star hotels in Egypt</td>
<td>Strong and positive</td>
</tr>
</tbody>
</table>

Table I. Research works into relationships between HRM practices and OC
## Table II. Research works into relationships between HRM practices and IJP and OP

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Themes</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huselid (1995)</td>
<td>Associations between HRM practices and intermediate employee outcomes (i.e. turnover and productivity) and short- and long-term measures of corporate financial performance in top US organizations</td>
<td>Positive/Strong</td>
</tr>
<tr>
<td>Wright and Edwards (1998)</td>
<td>TW using semi-autonomous groups with job rotation at a British aluminum smelter</td>
<td>Increased job satisfaction and IJP and decreased overtime and industrial accidents rates. Conditions for this outcome included capital-intensive continuous process technology, management-union cooperation and strong work group solidarity</td>
</tr>
<tr>
<td>Jayaram et al. (1999)</td>
<td>Associations between HRM practices and manufacturing performance of suppliers to the Big Three in the North America</td>
<td>HRM practices significantly affected the performance of suppliers in dimensions of quality, flexibility, cost and time</td>
</tr>
<tr>
<td>Guest (2002)</td>
<td>Associations between HRM practices and performance</td>
<td>Worker attitudes and behavior mediated this relationship. Job design, direct participation and information provision were associated with higher performance</td>
</tr>
<tr>
<td>Procter and Burridge (2008)</td>
<td>Association between the extent of TW and productivity and financial performance Association between teams' decision-making autonomy, team intensity and productivity and quality levels</td>
<td>Positive/Significant</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Researchers</th>
<th>Themes</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delarue et al. (2008)</td>
<td>Association between TW and performance</td>
<td>Positive if teamwork was combined with structural change, performance could be further enhanced</td>
</tr>
<tr>
<td>Çalışkan (2010)</td>
<td>Associations between HRM practices and bottom line productivity and effectiveness</td>
<td>Positive</td>
</tr>
<tr>
<td>Bartel (2000), Collings et al. (2010), Cho et al. (2006)</td>
<td>Association between TR and IJP and OP in terms of turnover rates in hospitality industry</td>
<td>Positive/Significant Participation programs and incentive plans led to lower turnover rates for non-managerial employees</td>
</tr>
<tr>
<td>Park and Kruse (2014)</td>
<td>Association between group incentives and financial performance</td>
<td>OC partially mediated Relationships of group incentives with OC and financial performance were stronger in more innovative companies than in lesones</td>
</tr>
<tr>
<td>von Bonsdorff et al. (2015)</td>
<td>Associations between perceived team autonomy and company performance through highlighting organizational commitment as a mediating factor in this relationship in 25 small-sized companies in the retail trade industry</td>
<td>Team autonomy was both directly and indirectly associated with company performance</td>
</tr>
<tr>
<td>Guest (2017)</td>
<td>Association between HRM practices and performance</td>
<td>Workers’ interests and needs should be ethically considered. External threats for workers could be diminished by appropriate HRM practices. Promoting employee well-being at work and positive employment relationship required managing different interests, trust among the parties to the relationship and fair treatment, concerning high performance and reduced costs</td>
</tr>
</tbody>
</table>
resources and capabilities, which increased organizational competitiveness (Barney, 2011). This fueled IJP and OP. Thus, $H_2$ and $H_3$ have been developed (Subsection 2.5).

Research works into relationships between OC and IJP and OP. Given Table III, some researchers found that employee commitment was a key to competitive performance in organizations. Thus, $H_4$–$H_5c$ have been developed (Subsection 2.5).

2.5 Hypotheses

The following hypotheses were developed and then tested:

$H_1$. HRM practices have significant positive relationship with OC.

$H_{1a}$. HRM practices have significant positive relationship with AC.

$H_{1b}$. HRM practices have significant positive relationship with CC.

$H_{1c}$. HRM practices have significant positive relationship with NC.

$H_2$. HRM practices have significant positive relationship with IJP.

$H_3$. HRM practices have significant positive relationship with OP.

$H_4$. OC has significant positive relationship with IJP.
3. Research method
In the survey (see Table AI), questions for HRM practices were concerned with nine items. They were RS, TR, RE, SW, TW, JD, DL, JS and CM scales that included 14 questions. RS questions were adopted from the work of Dessler (2002); TR questions were modified from the work of Dessler (2002); RE questions were modified form the works of Devanna et al. (1984) and Herzberg (1968); SW questions were adopted from the works of Dutra (2001) and Bohlander and Snell (2009); TW, DL and JS questions were modified from the work of Pfeffer (1998); JD question was adopted from the work of Hackman and Oldham (1980); and CM question was adopted from the work of Hall (1990). Questions for OC were concerned with three subvariables – AC, CC and NC. OC scale that was related to 3 items included 18 questions and was based on the work of Allen and Meyer (1990, 1996). Questions that were related to IJP and OP included 16 questions and were based on the work of Özkan (2011). Above scales used in their respective and other prior studies were reliable and valid. In answering questionnaires five-level Likert-type scale was used. Answers to questions ranged from 1 (strongly disagree) to 5 (strongly agree).

The sample used for this research included 240 Iranian employees and 200 Turkish employees working in tourism agencies. The sampling strategy was convenient and the data collection method of the study was cross-sectional, which was conducted through personal interviews and e-mails. Out of 800 questionnaires, 503 employees responded (i.e. a response rate of 62.88 percent) to interviews and e-mails. However, 63 responses were incomplete, and thus deleted from the study. The survey comprised HRM practices, OC, IJP and OP sections.

4. Data analysis and research findings
The data were analyzed by using SPSS for Windows 21.0 program. While evaluating the data, definitive statistical methods were used. The bivariate relationship between the dependent and independent variables was tested by Pearson correlation analysis. Finally, multiple regression analysis was used to test proposed hypotheses. Demographic statistics results about Iranian and Turkish travel agencies are reported in Table IV. The sample included data from 240 individuals in tourism agencies in Iran while 200 participants were from Turkey. The results showed that the proportions of the working position were almost close to each other in each case. Majority of the participants were employees (Iranian: 61 percent; Turkish: 59 percent), while there were only 8 percent high-level managers in both travel agencies. Furthermore, approximately 54 percent of the participants had less than five years of experience in both cases, almost 72 percent had less than TL2,000 monthly income, and majority of the participants were male (Iranian 56 percent; Turkish: 58 percent). Finally, 53 percent of Iranian and 51 percent of Turkish participants were single, 40 percent of Iranian and 44 percent of Turkish
employees had a high school diploma, while only 13 percent of Iranian and 9 percent of Turkish participants had postgraduate education level.

The descriptive statistics results for the latent variables are provided in Table V. The average HRM practices ranged between 2.8 and 3.4, the average commitment level changed between 2.8 and 3.3, while the average performance level was stable at 3.3. There was not a high variation around the average of the constructs which were ranged between 0.596 and 1.248.

The exploratory factor analysis results are provided in Table VI. Principal component analysis with varimax rotation was employed. KMO-statistics, to determine the sample adequacy, results were above the threshold value of 0.7 which indicated that the sample size was sufficient for EFA. Factor loadings ranged between 0.503 and 0.905 which was well above the suggested value of 0.35 based on the sample size (Hair et al., 2010).
The discriminant validity was met since the factors were distinct and the variables were loaded significantly only on one factor. Moreover, the results showed that Cronbach’s $\alpha$ ranged between 0.739 and 0.929 which also satisfied the recommended threshold value of 0.70, thus the reliability of the constructs was satisfied.

The evaluations of bivariate linear associations among the constructs are illustrated in Table VII. The Pearson correlation analysis results revealed that RS had positive, statistically significant and very weak correlations with AC and CC, while there was a positive, significant and weak correlation with IJP and OP, and no significant correlation with NC.

TR had a positive, significant, very weak association with CC and weak association with AC, IJP and OP, while it had no significant association with NC.

RE had positive, significant and weak relationship with NC and CC, while it had positive, significant and moderately strong relationship with AC, IJP and OP.

SW had a positive, statistically significant and weak correlation with AC, CC, NC, IJP and OP.

TW had a positive, significant and weak correlation with AC, CC and NC; moderately strong correlation with OP; and strong correlation with IJP.

JD had positive, significant and very weak correlation with NC; weak correlation with CC and OP; and moderately strong correlation with AC and IJP.

DL was positively, significantly and very weakly associated with CC and NC, while it was weakly associated with AC and OP; and moderately strongly associated with IJP.

JS had a weak relationship with AC and CC; had a moderately strong relationship with IJP and OP; and had no significant relationship with NC.

CM had a positive, significant and weak association with CC; had a moderately strong association with AC; and strong association with IJP and OP.

Furthermore, AC had a positive, significant and weak correlation with OP, while it had a moderately strong correlation with IJP. CC had a positive, significant and strong association with IJP and OP. Finally, NC had a positive, significant and weak correlation with OP, while it had a moderately strong correlation with IJP.

To test the proposed hypotheses, researchers employed multiple regression analysis. The hypotheitical relationships are illustrated in Figure 1.

The impacts of HRM practices on AC, CC and NC are illustrated in Table VIII. Accordingly, RE ($\beta = 0.329$, $p < 0.01$), TW ($\beta = 0.145$, $p < 0.05$), JD ($\beta = 0.138$, $p < 0.01$),

<table>
<thead>
<tr>
<th>Variables</th>
<th>Average</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>3.299</td>
<td>0.827</td>
<td>1.670</td>
<td>5.000</td>
</tr>
<tr>
<td>CC</td>
<td>2.816</td>
<td>0.814</td>
<td>1.170</td>
<td>5.000</td>
</tr>
<tr>
<td>NC</td>
<td>2.933</td>
<td>0.680</td>
<td>1.330</td>
<td>4.500</td>
</tr>
<tr>
<td>OC</td>
<td>3.016</td>
<td>0.596</td>
<td>1.830</td>
<td>4.280</td>
</tr>
<tr>
<td>RS</td>
<td>3.207</td>
<td>1.248</td>
<td>1.000</td>
<td>5.000</td>
</tr>
<tr>
<td>TR</td>
<td>3.136</td>
<td>1.105</td>
<td>1.000</td>
<td>5.000</td>
</tr>
<tr>
<td>RE</td>
<td>3.058</td>
<td>0.812</td>
<td>1.500</td>
<td>4.500</td>
</tr>
<tr>
<td>SW</td>
<td>2.827</td>
<td>1.191</td>
<td>1.000</td>
<td>5.000</td>
</tr>
<tr>
<td>TW</td>
<td>3.420</td>
<td>0.937</td>
<td>1.000</td>
<td>5.000</td>
</tr>
<tr>
<td>JD</td>
<td>3.402</td>
<td>1.099</td>
<td>1.000</td>
<td>5.000</td>
</tr>
<tr>
<td>DL</td>
<td>3.194</td>
<td>0.830</td>
<td>1.000</td>
<td>4.500</td>
</tr>
<tr>
<td>JS</td>
<td>3.833</td>
<td>0.950</td>
<td>1.000</td>
<td>5.000</td>
</tr>
<tr>
<td>CM</td>
<td>2.959</td>
<td>1.224</td>
<td>1.000</td>
<td>5.000</td>
</tr>
<tr>
<td>IJP</td>
<td>3.344</td>
<td>0.796</td>
<td>1.000</td>
<td>5.000</td>
</tr>
<tr>
<td>OP</td>
<td>3.317</td>
<td>0.946</td>
<td>1.000</td>
<td>5.000</td>
</tr>
</tbody>
</table>

Note: $n = 440$

2005

HRM, commitment and performance links

Table V. Descriptive statistics
<table>
<thead>
<tr>
<th>Latent variables</th>
<th>Items</th>
<th>Factor loading</th>
<th>Cronbach’s α</th>
<th>Eigenvalue$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RS</strong></td>
<td>rs01</td>
<td>0.836</td>
<td>0.826</td>
<td>8.048</td>
</tr>
<tr>
<td></td>
<td>rs02</td>
<td>0.786</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TR</strong></td>
<td>tr03</td>
<td>0.751</td>
<td>0.852</td>
<td>6.485</td>
</tr>
<tr>
<td></td>
<td>tr04</td>
<td>0.735</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RE</strong></td>
<td>re05</td>
<td>0.885</td>
<td>0.798</td>
<td>5.096</td>
</tr>
<tr>
<td></td>
<td>re06</td>
<td>0.636</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SW</strong></td>
<td>sw07</td>
<td>0.755</td>
<td>0.835</td>
<td>4.325</td>
</tr>
<tr>
<td></td>
<td>sw08</td>
<td>0.883</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TW</strong></td>
<td>tw09</td>
<td>0.753</td>
<td>0.849</td>
<td>3.004</td>
</tr>
<tr>
<td></td>
<td>tw10</td>
<td>0.733</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>JD</strong></td>
<td>jd11</td>
<td>0.713</td>
<td>0.812</td>
<td>2.012</td>
</tr>
<tr>
<td><strong>DL</strong></td>
<td>dl12</td>
<td>0.707</td>
<td>0.804</td>
<td>1.925</td>
</tr>
<tr>
<td><strong>JS</strong></td>
<td>js13</td>
<td>0.742</td>
<td>0.872</td>
<td>1.589</td>
</tr>
<tr>
<td><strong>CM</strong></td>
<td>cm14</td>
<td>0.754</td>
<td>0.891</td>
<td>1.008</td>
</tr>
</tbody>
</table>

Total variance: 54.09 percent; KMO-stat.: 0.812; Bartlett test of sphericity ($p$-value): 0.000

<table>
<thead>
<tr>
<th>Latent variables</th>
<th>Items</th>
<th>Factor loading</th>
<th>Cronbach’s α</th>
<th>Eigenvalue$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AC</strong></td>
<td>ac17</td>
<td>0.848</td>
<td>0.776</td>
<td>6.982</td>
</tr>
<tr>
<td></td>
<td>ac20</td>
<td>0.814</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ac19</td>
<td>0.812</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ac16</td>
<td>0.793</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ac15</td>
<td>0.596</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ac18</td>
<td>0.567</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CC</strong></td>
<td>cc24</td>
<td>0.714</td>
<td>0.782</td>
<td>4.923</td>
</tr>
<tr>
<td></td>
<td>cc23</td>
<td>0.652</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cc21</td>
<td>0.633</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cc22</td>
<td>0.614</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cc25</td>
<td>0.543</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cc26</td>
<td>0.523</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NC</strong></td>
<td>nc27</td>
<td>0.817</td>
<td>0.739</td>
<td>1.852</td>
</tr>
<tr>
<td></td>
<td>nc30</td>
<td>0.610</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nc31</td>
<td>0.594</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nc28</td>
<td>0.567</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nc32</td>
<td>0.521</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nc29</td>
<td>0.503</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total variance: 52.338 percent; KMO-stat.: 0.717; Bartlett test of sphericity ($p$-value): 0.000

<table>
<thead>
<tr>
<th>Latent variables</th>
<th>Items</th>
<th>Factor loading</th>
<th>Cronbach’s α</th>
<th>Eigenvalue$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IJP</strong></td>
<td>ijp36</td>
<td>0.905</td>
<td>0.905</td>
<td>8.104</td>
</tr>
<tr>
<td></td>
<td>ijp37</td>
<td>0.806</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ijp40</td>
<td>0.772</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ijp34</td>
<td>0.752</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ijp38</td>
<td>0.706</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ijp39</td>
<td>0.689</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ijp35</td>
<td>0.644</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ijp33</td>
<td>0.612</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OP</strong></td>
<td>op41</td>
<td>0.845</td>
<td>0.929</td>
<td>1.781</td>
</tr>
<tr>
<td></td>
<td>op17</td>
<td>0.817</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>op45</td>
<td>0.795</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>op44</td>
<td>0.780</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>op48</td>
<td>0.779</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>op46</td>
<td>0.727</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>op42</td>
<td>0.711</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>op43</td>
<td>0.689</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total variance: 58.892 percent; KMO-stat.: 0.859; Bartlett test of sphericity ($p$-value) = 0.000

Note: $^a$Values obtained after rotation

Table VI. Exploratory factor analysis (EFA) results
Employees' perceptions of HRM practices

Organizational performance

Individual job performance

Organizational commitment
- Affective commitment
- Continuous commitment
- Normative commitment

H1, H1a, H1b, H1c

H2

H3

H4, H4a, H4b, H4c

H5, H5a, H5b, H5c

Figure 1. The proposed model

<table>
<thead>
<tr>
<th>Variables</th>
<th>AC</th>
<th>CC</th>
<th>NC</th>
<th>IJP</th>
<th>OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRM: RS</td>
<td>0.183**</td>
<td>0.120*</td>
<td>-0.056</td>
<td>0.269**</td>
<td>0.310**</td>
</tr>
<tr>
<td>HRM: TR</td>
<td>0.260**</td>
<td>0.159**</td>
<td>0.093</td>
<td>0.310**</td>
<td>0.237**</td>
</tr>
<tr>
<td>HRM: RE</td>
<td>0.515**</td>
<td>0.318**</td>
<td>0.200**</td>
<td>0.528**</td>
<td>0.487**</td>
</tr>
<tr>
<td>HRM: SW</td>
<td>0.298**</td>
<td>0.249**</td>
<td>0.268**</td>
<td>0.294**</td>
<td>0.294**</td>
</tr>
<tr>
<td>HRM: TW</td>
<td>0.397**</td>
<td>0.269**</td>
<td>0.203**</td>
<td>0.642**</td>
<td>0.558**</td>
</tr>
<tr>
<td>HRM: JD</td>
<td>0.400**</td>
<td>0.389**</td>
<td>0.124**</td>
<td>0.474**</td>
<td>0.358**</td>
</tr>
<tr>
<td>HRM: DL</td>
<td>0.249**</td>
<td>0.160**</td>
<td>0.190**</td>
<td>0.504**</td>
<td>0.376**</td>
</tr>
<tr>
<td>HRM: JS</td>
<td>0.285**</td>
<td>0.295**</td>
<td>0.02</td>
<td>0.568**</td>
<td>0.520**</td>
</tr>
<tr>
<td>HRM: CM</td>
<td>0.471**</td>
<td>0.392**</td>
<td>0.081</td>
<td>0.610**</td>
<td>0.604**</td>
</tr>
<tr>
<td>AC</td>
<td></td>
<td></td>
<td></td>
<td>0.532**</td>
<td>0.383**</td>
</tr>
<tr>
<td>CC</td>
<td></td>
<td></td>
<td></td>
<td>0.600**</td>
<td>0.632**</td>
</tr>
<tr>
<td>NC</td>
<td></td>
<td></td>
<td></td>
<td>0.473**</td>
<td>0.296**</td>
</tr>
</tbody>
</table>

Notes: *p < 0.05; **p < 0.01

Table VIII. Impact of HRM practices on OC

<table>
<thead>
<tr>
<th>Variables</th>
<th>AC</th>
<th>CC</th>
<th>NC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRM: RS</td>
<td>0.183**</td>
<td>0.120*</td>
<td>-0.056</td>
</tr>
<tr>
<td>HRM: TR</td>
<td>0.260**</td>
<td>0.159**</td>
<td>0.093</td>
</tr>
<tr>
<td>HRM: RE</td>
<td>0.515**</td>
<td>0.318**</td>
<td>0.200**</td>
</tr>
<tr>
<td>HRM: SW</td>
<td>0.298**</td>
<td>0.249**</td>
<td>0.268**</td>
</tr>
<tr>
<td>HRM: TW</td>
<td>0.397**</td>
<td>0.269**</td>
<td>0.203**</td>
</tr>
<tr>
<td>HRM: JD</td>
<td>0.400**</td>
<td>0.389**</td>
<td>0.124**</td>
</tr>
<tr>
<td>HRM: DL</td>
<td>0.249**</td>
<td>0.160**</td>
<td>0.190**</td>
</tr>
<tr>
<td>HRM: JS</td>
<td>0.285**</td>
<td>0.295**</td>
<td>0.02</td>
</tr>
<tr>
<td>HRM: CM</td>
<td>0.471**</td>
<td>0.392**</td>
<td>0.081</td>
</tr>
</tbody>
</table>

Table VII. Pearson correlation analysis results

Dependent variable | AC | CC | NC | IJP | OP |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RS</td>
<td>0.012</td>
<td>1.18</td>
<td>-0.031</td>
<td>0.43</td>
<td>0.112**</td>
</tr>
<tr>
<td>TR</td>
<td>0.023</td>
<td>0.29</td>
<td>0.007</td>
<td>0.56</td>
<td>0.028</td>
</tr>
<tr>
<td>RE</td>
<td>0.328***</td>
<td>3.31</td>
<td>0.004</td>
<td>0.95</td>
<td>0.251***</td>
</tr>
<tr>
<td>SW</td>
<td>0.083</td>
<td>0.98</td>
<td>0.006</td>
<td>1.41</td>
<td>0.162**</td>
</tr>
<tr>
<td>TW</td>
<td>0.145**</td>
<td>2.18</td>
<td>0.008</td>
<td>0.19</td>
<td>0.167**</td>
</tr>
<tr>
<td>JD</td>
<td>0.138***</td>
<td>3.31</td>
<td>0.204**</td>
<td>2.25</td>
<td>0.003</td>
</tr>
<tr>
<td>DL</td>
<td>0.158*</td>
<td>1.75</td>
<td>-0.048</td>
<td>0.27</td>
<td>0.131***</td>
</tr>
<tr>
<td>JS</td>
<td>0.134**</td>
<td>2.15</td>
<td>0.024</td>
<td>0.31</td>
<td>0.171**</td>
</tr>
<tr>
<td>CM</td>
<td>0.248*</td>
<td>1.70</td>
<td>0.212**</td>
<td>2.52</td>
<td>0.146**</td>
</tr>
</tbody>
</table>

F-statistics: 24.656***

| Notes: *p < 0.10; **p < 0.05; ***p < 0.01 |

| Notes: *p < 0.10; **p < 0.05; ***p < 0.01 |

Table VIII. Impact of HRM practices on OC
DL ($\beta = 0.158, p < 0.10$), JS ($\beta = 0.134, p < 0.05$) and CM ($\beta = 0.248, p < 0.10$) had a positive and statistically significant relationship with AC. Furthermore, JD ($\beta = 0.204, p < 0.05$) and CM ($\beta = 0.212, p < 0.05$) were positively and significantly associated with CC. Finally, RS ($\beta = 0.112, p < 0.05$), RE ($\beta = 0.251, p < 0.01$), SW ($\beta = 0.162, p < 0.05$), TW ($\beta = 0.167, p < 0.05$), DL ($\beta = 0.131, p < 0.01$), JS ($\beta = 0.171, p < 0.05$) and CM ($\beta = 0.146, p < 0.05$) had a statistically significant and positive impact on NC, while TR and JD had no impact on NC. The results revealed that $H_{1a}$–$H_{1c}$ were partially supported. The coefficient of determinations of the models indicated that 32.7 percent variation in AC, 20.9 percent variation in CC and 20.0 percent variation in NC were explained by the HRM practices as the independent variables.

Table IX illustrates the impact of HRM practices on IJP and OP. The results revealed that RS ($\beta = 0.063, p < 0.05$), TR ($\beta = 0.083, p < 0.05$), SW ($\beta = 0.049, p < 0.05$), TW ($\beta = 0.282, p < 0.01$), DL ($\beta = 0.214, p < 0.01$) and CM ($\beta = 0.264, p < 0.01$) had a positive and significant impact on IJP. However, the rest of the HRM practices showed no significant association with IJP. Moreover, TR ($\beta = 0.153, p < 0.01$), JS ($\beta = 0.146, p < 0.01$) and CM ($\beta = 0.375, p < 0.05$) were positively and statistically significantly associated with the OP, while the rest of the HRM practices were not significantly associated with OP. Thus, $H_2$ and $H_3$ were partially supported. Finally, 58.8 percent variation in IJP and 47.6 percent variation in OP were explained by the HRM practices as the independent variables.

The impacts of AC, CC and NC on IJP and OP are illustrated in Table X. The results suggested that AC ($\beta = 0.239, p < 0.01$), CC ($\beta = 0.424, p < 0.01$) and NC ($\beta = 0.263, p < 0.05$) had a positive and significant impact on IJP. AC ($\beta = 0.141, p < 0.01$) and

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>IJP</th>
<th>OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS</td>
<td>0.063**</td>
<td>2.115</td>
</tr>
<tr>
<td>TR</td>
<td>0.083**</td>
<td>2.151</td>
</tr>
<tr>
<td>RE</td>
<td>0.001</td>
<td>1.343</td>
</tr>
<tr>
<td>SW</td>
<td>0.049**</td>
<td>2.281</td>
</tr>
<tr>
<td>TW</td>
<td>0.282***</td>
<td>3.313</td>
</tr>
<tr>
<td>JD</td>
<td>0.083</td>
<td>1.365</td>
</tr>
<tr>
<td>DL</td>
<td>0.214***</td>
<td>3.109</td>
</tr>
<tr>
<td>JS</td>
<td>0.032</td>
<td>1.575</td>
</tr>
<tr>
<td>CM</td>
<td>0.264***</td>
<td>2.984</td>
</tr>
</tbody>
</table>

F-statistics: 70.517*** 45.395***

Notes: **p < 0.05; ***p < 0.01

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>IJP</th>
<th>OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>0.239***</td>
<td>3.13</td>
</tr>
<tr>
<td>CC</td>
<td>0.424***</td>
<td>3.93</td>
</tr>
<tr>
<td>NC</td>
<td>0.263***</td>
<td>2.33</td>
</tr>
</tbody>
</table>

F-statistics: 142.291*** 105.962***

Notes: **p < 0.05; ***p < 0.01

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>IJP</th>
<th>OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>0.239***</td>
<td>3.13</td>
</tr>
<tr>
<td>CC</td>
<td>0.424***</td>
<td>3.93</td>
</tr>
<tr>
<td>NC</td>
<td>0.263***</td>
<td>2.33</td>
</tr>
</tbody>
</table>

F-statistics: 142.291*** 105.962***

Notes: **p < 0.05; ***p < 0.01
CC ($\beta = 0.654$, $p < 0.01$) also had a statistically significant and positive association with OP, while NC had no significant impact on OP. Therefore, $H_{4a}–H_{5b}$ were fully supported, but $H_{5c}$ was not supported. According to the correlation of determination results, 49.1 percent variation on IJP and 41.8 percent variation on OP were explained by AC, CC and NC.

5. Conclusion
Given correlation analysis, the travel agencies in Iran and Turkey should recognize that fair screening and hiring, promotion and development processes and well-designed and administered training programs were related to employees’ emotional attachment to the organization, awareness of costs of leaving the organization, and IJP and OP. Likewise, intrinsic and extrinsic rewards, performance-based fair pay system, fully participative teamwork, clear job description, and assigned duties and authority were related to employees’ attachment to the organization, awareness of costs of leaving the organization, feelings of obligation to continue with the organization, and IJP and OP. Furthermore, employment security was related to employees’ attachment to the organization, awareness of costs of leaving the organization, and IJP and OP. Finally, employees’ attachment to the organization, awareness of costs of leaving the organization and feelings of obligation to continue with the organization were related to IJP and OP.

Regarding the effects of HRM practices on OC the research suggested that RE, TW, JD, DL and CM had significant and positive association with AC; JD and CM had significant and positive association with CC; and RS, RE, SW, TW, DL, JS and CM had significant and positive association with NC. However, TR and JD had significant and negative impact on NC. Thus, $H_{1a}–H_{1c}$ were partially supported. The travel agencies in Iran and Turkey should consider the following, which were espoused by some prior research works (Table I), when they formulate their future strategies concerning HRM and OC relationships:

- Timely, performance contingent, durable, visible and fair intrinsic and extrinsic rewards; well-designed and administered esprit de corps promoting enriched debate, collective act and dialogue; clearly defined duties and responsibilities; authority and responsibility granting processes; and fair individual development programs and promotions will support employees’ emotional attachment to the organization.
- Well-written job analysis documents providing autonomy, challenge, feedback and skill utilization and fair design of employees’ future development plans and programs will stimulate the awareness of employees concerning costs of leaving the organization.
- Equitable screening and hiring; fair and collective intrinsic rewards, such as enriched jobs, engaging work and opportunities for decision making and extrinsic rewards, such as employee stock ownership plans, bonuses and gain sharing; performance-based remuneration which is compatible with cost of living, supply of qualified personnel and success and conditions of business; harmony and good feelings in collective business processes among all individuals and units; employment security; assignment of specific duties, granting of appropriate authority and creation of commitment for employees to perform assigned duties; and fair promotion and development programs will generate a feeling of obligation to continue employment.

Regarding the effects of HRM practices on IJP and OP the research suggested that RS, TR, SW, TW, DL and CM had significant and positive association with IJP. Furthermore, TR, JS and CM had significant and positive association with OP. The rest of HRM practices had no significant relationship with OP. Thus, $H_{2}$ and $H_{3}$ were partially supported. The travel agencies in Iran and Turkey should consider the following, which were
supported by some prior research works (Table II), when they formulate their future strategies concerning HRM, IJP and OP relationships:

- Fair screening and hiring policies that ease to comprehend levels of skill, knowledge, capacity and experience of candidates for the job; properly designed and administered on the job and out of job training programs enhancing employees’ skills, talent and knowledge; satisfactory, fair and competitive conditions of employment; fully participative teamwork; authority granting, duty assignment, obligation creation; and fair promotion and development programs will raise IJP, such as work-related knowledge, motivation, productivity, effectiveness and satisfaction.

- Well-designed and administered on the job and out of job training programs, employment security, and fair promotion and development programs enhancing employee competences will increase OP, such as profitability, market share, innovativeness, growth rate, competitiveness and investment to human resources.

Regarding the effects of OC on IJP and OP the research suggested that AC, CC and NC had significant and positive association with IJP. Likewise, AC and CC had significant and positive association with OP. Thus, $H_{4a}–H_{5b}$ were fully supported. However, $H_{5c}$ was not supported. The travel agencies in Iran and Turkey should consider the following, which were championed by some prior research works (Table III), when they formulate their future strategies concerning OC, OP and IJP relationships:

- Timely and fair training, mentoring and development programs which are supported by top management; higher base pay level which is generally interpreted as information about one’s worth to the organization; and 360° feedback by which individuals receive feedback from others in performance appraisal that fosters continuous learning culture will enhance IJP and OP, such as employee motivation, productivity, work-related knowledge, satisfaction, effectiveness as well as profitability, market share, growth rate and innovativeness.

- Employees’ feelings of obligation to continue with the organization will not affect the organization’s outcome-based financial indicators, such as business success, market share, growth rate, profitability, innovativeness, competitiveness and investment to human resources.

Few studies treated HRM–OC–performance links in hospitality industry in Iranian and Turkish research literature. Thus, this paper made a contribution to the Iranian and Turkish tourism business environments.

In view of the limitations of this study the first issue was about the generalizability of research due to geographic boundaries of data collected. The inferences drawn from such samples in Iran and Turkey may not be fully generalizable to the tourism agencies from other countries in different cultures. Also, since the questionnaires were distributed by using convenient sampling method, the control pertaining to who fills the questionnaire was limited. Related with this, a major constraint was the disability of getting the respondents to fill in the questionnaire in a complete way although they were just on feet and continuing their works, so that a limited number of complete questionnaires were obtained. Furthermore, the sample size was limited to the respondents in Iran and Turkey. Hence, this situation constrained to generalize the findings to the overall population.

Finally, this study did not address personality traits of employees, socio-cultural features, economic conditions, etc. In order to make a better explanation of the relationships among the constructs of this study, a broader geographic region might also have been chosen with a higher number of respondents as sample groups. This would have enabled the practitioners to get more valid and realistic research findings.
References


Hall, D.T. (1990), Careers in Organizations, Goodyear, Santa Monica, CA.


Further reading

Appendix

HRM
Rs01 The organization disseminates information about recruitment
Rs02 The organization discloses information to applicants regarding the steps and criteria of hiring
Tr03 The organization designs my development and education promoting my personal and professional growth
Tr04 The organization administers my development and education promoting my personal and professional growth
Re05 I get bonuses and premiums
Re06 I have freedom to use my own judgment
Sw07 My salary/wage is influenced by my results
Sw08 My salary/wage is compatible with my skills, training and education
Tw09 Team working is encouraged
Tw10 Team included members from all levels of the organization
Jd11 My job requires the completion of a whole, relatively identifiable piece of work
Dl12 Specific and understandable duties and appropriate authority are assigned to employees
Js13 I feel certain that I will not be laid off in the next 12 months
Cm14 My organization develops career plans with individuals

OC
Ac15 I would be very happy to spend the rest of my career with this organization
Ac16 I really feel as if this organization’s problems are my own
Ac17 I do not feel a strong sense of “belonging” to my organization
Ac18 I do not feel “emotionally attached” to this organization
Ac19 I do not feel like “part of the family” at my organization
Ac20 This organization has a great deal of personal meaning for me
Cc21 Right now, staying with my organization is a matter of necessity as much as desire
Cc22 It would be very hard for me to leave my organization right now, even if I wanted to
Cc23 Too much in my life would be disrupted if I decided to leave my organization now
Cc24 I feel that I have very few options to consider leaving this organization
Cc25 It would not be too costly for me to leave my organization now
Cc26 One of the few serious consequences of leaving this organization would be the scarcity of available alternatives

Nc27 I think that people these days move from company to company too often
Nc28 If I got another offer for a better job elsewhere I would not feel it was right to leave my organization
Nc29 Jumping from organization to organization does not seem at all unethical to me
Nc30 I was taught to believe in the value of remaining loyal to one organization
Nc31 One of the major reasons I continue to work in this organization is that I believe loyalty is important and feel a sense of moral obligation to remain
Nc32 I do not think that to be a “company man” or “company woman” is sensible anymore

IJP
Ijp33 Employees perform higher than those who work for rivals
Ijp34 Employees have better work-related knowledge than those who work for rivals
Ijp35 Employees are more successful than those who work for rivals
Ijp36 Employees are more motivated than those who work for rivals
Ijp37 Employees are more committed than those who work for rivals
Ijp38 Employees generate more value than those who work for rivals
Ijp39 Employees are more satisfied than those who work for rivals
Ijp40 Employees execute their tasks better than those who work for rivals

OP
Op41 Organization is mostly more successful than its rivals
Op42 Organization has a higher market share than its rivals
Op43 Organization grows faster than its rivals
Op44 Organization is more profitable than its rivals
Op45 Organization is more innovative than its rivals
Op46 Organization provides better services than its rivals

Table AI.
Items: HRM, OC, OP and IJP

2016
About the authors
Dr N. Gökhan Torlak, PhD, studied Industrial Relations and Economics at Istanbul University from 1983 to 1987, where he received BA degree. Then, he received MA degree in Management Systems at Hull University in 1991 and PhD degree in Management Systems at the University of Lincolnshire and Humberside in 1999. Torlak spent his academic life teaching at Bilgi University from 2001 to 2002 and Fatih University since 2002 where he chaired Department of Management from 2004 to 2008. He became Associate Professor in 2010 and Full Professor in 2015 in the area of Management and Organization. He supervised 20 MA students and 1 PhD student to successful completion. He published numerous articles and books in peer-reviewed and prestigious national and international journals, delivered conference papers and undertaken consultancy engagements and EU projects. His research interests are related to applications of systems methodologies, methods and techniques in the solution of methodological, managerial and organizational problems. Dr N. Gökhan Torlak is the corresponding author and can be contacted at: gokhantorlak@ttmail.com

Dr Cemil Kuzey, PhD, is Assistant Professor in the Department of Computer Science and Information Systems, Arthur J. Bauernfeind College of Business at Murray State University. Dr Kuzey is primarily teaching Operation Research and Statistics topics for Social Sciences. He received PhD degree in Business Administration from the Department of Quantitative Analysis at Istanbul University, Turkey. Among his academic pursuits, he took several graduate courses at the Ontario Institute for Studies in Education, University of Toronto, Canada. His research interests are related to operation research, data mining and business intelligence.

Maral Ragom, MA, graduated from the Department of Management, Fatih University. She received MA degree from the same university. She is currently working as Manager in the Department of Sales Operations, ATU Duty Free Turkish Airlines.

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Effects of strategic information systems on competitive strategy and performance

Adilson Carlos Yoshikuni and Alberto Luiz Albertin

Department of Information Technology and Quantitative Methods Applied to Administration,
School of Business Administration of Sao Paulo, Sao Paulo, Brazil

Abstract

Purpose – This study argues that strategic information systems (SISs) are necessary for organizations’ survival and corporate performance in turbulent economic environments. Applying Miles and Snow’s strategy typology, the purpose of this paper is to explore how SIS supports business strategy and corporate performance.

Design/methodology/approach – This study uses quantitative survey data from 389 Brazilian companies during economic crises and analyzes them using structural equation modeling.

Findings – There is strong evidence that SIS promotes capacity and flexibility to create competitive strategies in response to environmental changes. SIS significantly and positively predicts firms’ use of prospector strategies, reducing the need to sacrifice efficiency for innovation. SIS can predict corporate performance more strongly than firms’ strategic orientations can.

Practical implications – The results provide organizations insights on how SIS enables strategic planning processes to create competitive strategy and improve performance during economic turbulence.

Originality/value – This research demonstrates SIS’s positive effects during economic turbulence on competitive strategy and performance, revealing that corporate performance is influenced more by SIS (strategic process) than strategic orientation (content). Hence, this study fills a research gap in the information systems strategy literature by contributing new insights about SIS.

Keywords Balanced scorecard, Strategic orientation, Corporate performance, IS strategy, IT/IS business value, Strategic information systems

1. Introduction

Several studies have demonstrated that creating business strategy value requires effective use of strategic information systems (SISs) (Chan and Huff, 1992; Chen et al., 2010; Marabelli and Galliers, 2017; Newkirk et al., 2003; Philip, 2007; Teo and King, 1997; Wang and Byrd, 2017). Recognizing that SIS is widely utilized, Chen et al. (2010) and Merali et al. (2012) suggest that conceptions emerging from the SIS literature both differ from and complement each other regarding contextual elements of the SIS process, content, desired impact and alignment.

Understanding the importance of SIS streams and the infeasibility of examining them all, this study focuses on SIS as use of information systems (IS) to support the process and content of business strategy in complex environments. According to Newkirk and Lederer (2006) and Singh et al. (2002), SIS supports all strategic planning processes, such as strategic awareness, situation analysis, strategy conception, strategy formulation and strategy implementation/control planning; moreover, SIS enables the content of business strategies and influences corporate performance (Chan and Reich, 2007; Chan et al., 2006; Marabelli and Galliers, 2017; Whittington, 2014).

Strategic planning is a systematic process for establishing strategies geared toward providing firms with competitive advantages and improving corporate performance in certain environmental conditions (Grant, 2003; Hill et al., 2014; Wolf and Floyd, 2017; Yoshikuni and Jeronimo, 2013). Several typologies describe strategic planning...
(Mintzberg et al., 2009), the most famous being Porter’s (1986) typology of generic strategies and Miles and Snow’s (Miles et al., 1978) typology of strategic behaviors. Porter’s (1986) typology presents types of cost leadership and differentiates broad and segmented targets (focus). In other words, Porter’s perspectives classify strategy based on the extremes of innovation and efficiency, but it cannot measure the strategy’s ambidexterity. Miles and Snow’s (Miles et al., 1978) typology can test the trade-off between innovation and efficiency using the analyzer strategic orientation archetype (strategy’s ambidexterity) and firms without a strategic orientation, such as the reactor typology. Thus, the use of their typology is prioritized in this study.

Given today’s global marketplace and increasingly complex economy, SIS is critical to many organizations’ survival, and business managers consistently rank it among the top IS management issues (Whittington, 2014). Brazilian managers assume that information technology/information systems (IT/IS) can strengthen corporate performance (Meirelles, 2016), and Brazilian firms spent 7.6 percent of their revenue on IT/IS solutions to address and adapt to economic turbulence.

Thus, the study sought to fill research gaps in the IS strategy literature by contributing new insights about SIS (Chen et al., 2010; Newkirk and Lederer, 2006; Peppard et al., 2014; Whittington, 2014). Specifically, SIS must incorporate strong planning capabilities to help organizations effectively adapt to changing factors – internal and external – to enable strategic orientation (Miles et al., 1978). Furthermore, SIS affects corporate performance (as measured by the balanced scorecard) (Kaplan and Norton, 1992, 2008) during extreme economic turbulence.

2. Theoretical framework

This section describes the factors investigated with regard to SIS, strategic orientation and corporate performance.

2.1 Strategic information systems

This study adopted the framework of IT/IS business value (Kohli and Grover, 2008; Melville et al., 2004) for how IT/IS resources are applied within strategic business processes to improve performance. According to Whittington (2014), SIS is used as an IT/IS application to promote strategy-as-practice in order to support deliberate managerial planning for strategic positioning and performance; that is, technology applications enable strategic practice to develop and execute content strategies. Arvidsson et al. (2014) investigate SIS embedded in strategy practices, enabling a firm to analyze diverse scenarios and increase the speed of strategy development, to explore emergent strategy grounded in the activities of multiple organizational sub-communities in which firms develop specific and original strategy content. Thus, many strategizing practices by SIS, from strategy formulation to strategy communication, enable orientation strategy to gain corporate performance, and show that business strategy process is inseparable from the influence of IS on content strategies (Whittington, 2014).

SIS is defined as a portfolio of IS applications supporting an organization’s business plans (Sabherwal and Chan, 2001) to enable the process and content of business strategy to achieve its objectives. These applications include operational support systems, business collaboration systems, management IS and decision support systems (Laudon and Laudon, 2006; O’Brien and Marakas, 2007; Sabherwal and Chan, 2001; Singh et al., 2002). According to Mentzas (1997), Newkirk and Lederer (2006) and Yoshikuni and Jeronimo (2013), SIS supports the strategic planning process by enabling strategic awareness through disseminating strategic objectives/goals for every organization; making it possible to map the external environment’s opportunities and threats when analyzing the general environment in which a company exists; designing strategy by aligning internal
resources – technological, people and organizational – to utilize opportunities and mitigate threats; selecting and formulating strategies to develop new business processes leveraged by IT/IS resources; and implementing and monitoring business strategies by supporting change processes, execution and control of action plans.

Recently, research has been intensifying on the creation of business value through IS and IT resources (Marabelli and Galliers, 2017; Melville et al., 2004; Merali et al., 2012). Effective use of IS in business strategy processes is considered a key factor for chief information officers and chief executive officers (Anwar and Hasnu, 2016; Philip, 2007). SIS studies seek to guide research in this theoretical framework (Chan and Huff, 1992; Earl, 1993; Ein-Dor and Segev, 1978; King, 1978; Lederer and Salmela, 1996; Teo and King, 1997) and clarify how effective use of IS contributes to strategic planning processes and sharing of organizational perspectives to maintain and achieve corporate objectives (Chen et al., 2010; Jääskeläinen and Luukkanen, 2017; Mentzas, 1997; Newkirk et al., 2003; Newkirk and Lederer, 2006; Philip, 2007).

Top managers define strategic objectives to provide the basis for developing strategic planning practices (Bernado et al., 2017; Hill et al., 2014). SIS enables organizational collaboration competency to promote effective utilization and management of relevant stakeholder groups’ inputs into the SIS process (Philip, 2007; Segars and Grover, 1999). Effective SIS use facilitates communication with and monitoring of employees, which helps determine whether objectives are being met (Jääskeläinen and Luukkanen, 2017; Karpovsky and Galliers, 2015; Segars et al., 1998). SIS promotes collaboration at all organizational levels, including top management, in internal and external SIS processes to improve the effectiveness of the strategic planning process (Newkirk and Lederer, 2006; Yeh et al., 2012). Managers use strategic IS to organize planning teams and obtain strategic commitment from top management to improve its effectiveness in performing tasks and achieving communication goals (Dameron et al., 2015; Jääskeläinen and Luukkanen, 2017; Muriithi et al., 2016; Philip, 2007; Segars et al., 1998).

Environmental analysis involves scanning all external factors that affect (but are not directly controlled by) the organization to identify opportunities for improving operations (Hill et al., 2014; Mintzberg et al., 2009; Porter, 1986). Through environmental analysis, firms can secure information that describes advancements, opportunities and threats in the external environment (Dameron et al., 2015; De Lorenzi Cancellier et al., 2014; Newkirk and Lederer, 2006; Singh et al., 2002; Xu et al., 2011). Hence, SIS must incorporate strong planning capabilities to be flexible enough to adapt IS effectively to changing external factors (Davenport et al., 2010; George et al., 2014; Leidner et al., 2011; Marabelli and Galliers, 2017) and to enable organizational competencies to conceive strategies.

Strategy conception relates to the development, evaluation and selection of organizational strategies (Bernado et al., 2017; Hill et al., 2014; Mintzberg et al., 2009; Porter, 1986). Specifically, it involves identification of potentially problematic issues, generation of alternative courses of action and analysis of proposed strategic approaches (Kaplan and Norton, 2008; Porter, 1986; Rouhani et al., 2016; Segars et al., 1998; Shollo and Galliers, 2016; Singh et al., 2002). Thus, IS can strengthen a firm’s capacity to coordinate and integrate, and can increase its ability to alter current strategies (Jääskeläinen and Luukkanen, 2017; Kim et al., 2011; Newkirk and Lederer, 2006; Pavlou and El Sawy, 2006, 2010).

Managers must develop and implement strategic actions that are consistent with the company’s business strategy, thereby, facilitating the achievement of long-term organizational goals (Hill et al., 2014; Jääskeläinen and Luukkanen, 2017; Mintzberg et al., 2009; Porter, 1986). During strategy implementation, SIS provides information about how project plans are realized, facilitates communication and coordination of work activities among work personnel, supports budgetary processes and matches job requirements with personnel qualifications (Jääskeläinen and Luukkanen, 2017; Kim et al., 2011; León-Soriano et al., 2010; Muriithi et al., 2016; Newkirk and Lederer, 2006; Teo and King, 1997).
Finally, strategic control relates to monitoring the implementation of a strategy and assessing its outcomes (Hill et al., 2014; Kaplan and Norton, 1992, 2008; Wolf and Floyd, 2017). Strategic control ensures effective and efficient use of resources to accomplish organizational objectives (Jääskeläinen and Luukkanen, 2017; León-Soriano et al., 2010; Singh et al., 2002). Specifically, strategic use of IS generates and integrates data for careful evaluation by organizational personnel (Kaplan and Norton, 2008). Moreover, it allows for comparison of corporate performance with budgets, goals, standards and targets (León-Soriano et al., 2010; Muriithi et al., 2016; Newkirk and Lederer, 2006).

Thus, SIS is a set of IT/IS applications that collect, process, analyze and provide data/information for decision making. These applications enable a holistic, interactive, decentralized and dynamic view of the organization; generate organizational knowledge; and facilitate learning in the strategic planning process. Hence, SIS-embedded strategic planning enables strategy content to gain competitive advantage and improve firm performance.

2.2 Strategic orientation

Strategic orientation relates to the way in which a firm adapts to environmental changes to achieve corporate performance (Chan, 1997; Moore, 2005). Among the strategic typologies in strategic management, Miles and Snow’s typology (Miles et al., 1978) is one of the most enduring, scrutinized and applied frameworks (Anwar and Hasnu, 2016; Avci et al., 2011; Chatzoglou et al., 2017; Conant et al., 1990; Frambach et al., 2016).

This typology considers that managers’ divergent strategic orientations can induce similar businesses to operate differently when facing environmental changes (Efrat and Shoham, 2013). They argued that these business strategies may exist simultaneously within industries, and viable strategies (prospectors, defenders and analyzers), if properly implemented, would yield similar results and outperform non-viable strategies (reactors) (Anwar and Hasnu, 2016; Parnell et al., 2015). Ultimately, Miles and Snow (Miles et al., 1978) defined four strategic orientation archetypes as part of their typology: prospector, defender, analyzer and reactor firms.

Prospector firms monitor market trends in order to be the first entrant in a new market or the first developer of a new product (Avci et al., 2011; Parnell et al., 2012, 2015). These firms are externally oriented and constantly redefine markets. In addition, they adopt new production systems and technologies with little hesitation (Efrat and Shoham, 2013; Frambach et al., 2016; Moore, 2005).

However, defender firms protect their status in the current markets and seek market stability (Conant et al., 1990; Efrat and Shoham, 2013; Parnell et al., 2015). They are risk averse and adhere to systems that improve the efficiency of existing operations (Moore, 2005; Parnell, 2010; Sabherwal and Chan, 2001). They seek only proven opportunities and, thus, tend to lag behind industry competitors in terms of innovation (Anwar and Hasnu, 2016; Avci et al., 2011; Chatzoglou et al., 2017).

Analyzer firms are hybrids of defenders and prospectors (Conant et al., 1990; Frambach et al., 2016; Parnell et al., 2015). These firms primarily seek to minimize risk and maximize opportunities for profit, developing a balance between the two (Avci et al., 2011; Parnell, 2010; Sabherwal and Chan, 2001). Analyzers tend to focus on efficiency and productivity when the market is stable, but they engage in cautious scanning and innovation during market turbulence (Anwar and Hasnu, 2016; Chan, 1997; Parnell et al., 2012).

Finally, reactor firms tend to be inconsistent in their adaptive patterns. They respond to changes in competitive circumstances only when forced (Avci et al., 2011; Chan, 1997). According to Miles and Snow (Miles et al., 1978), the reactors have a dysfunctional orientation because of their inconsistent strategic approach (Miles et al., 1978). They often suffer poor performance relative to firms with different strategic orientations (Anwar and Hasnu, 2016; Moore, 2005).
In summary, defenders play it safe by operating in a narrow, stable domain, whereas prospectors frequently take risks with untried products. Analyzers are highly risk averse; they look for opportunities to grow but only add new products/services that have already been shown to work successfully by another organization (a prospector). Reactors do not follow a conscious strategy, which is not considered viable in the long run (Anwar and Hasnu, 2016; Chan, 1997; Sabherwal and Chan, 2001).

2.3 Corporate performance
To analyze critical features of the competitive environment, it is necessary to extend extant measurements of corporate performance so they can assess multiple dimensions of organizational success (Chan, 1997; Jääskeläinen and Luukkanen, 2017; Mithas et al., 2011; Mostaghel et al., 2015; Norreklit, 2000; Reefeke and Trocchi, 2013; Sen et al., 2017; Sohn et al., 2003; Yoshikuni et al., 2014). Many researchers consider the balanced scorecard (Kaplan and Norton, 1992) an effective and comprehensive tool for measuring corporate performance based on financial success, customer performance (CUPE), internal process efficiency and organizational learning (Bento et al., 2013; Callado and Jack, 2015; Kaplan and Norton, 2008; Park et al., 2017; Perkins et al., 2014; Yoshikuni and Albertin, 2017).

Financial performance (FIPE) – a function of productivity and growth-based corporate performance – is related to the degree to which a firm converts tangible and intangible assets into shareholder value (Atkinson et al., 2011; Mithas et al., 2011; Norreklit, 2000; Perkins et al., 2014). Productivity strategy concerns the efficient management of costs, expenses and investment performance; growth strategy is primarily associated with revenue generation (Callado and Jack, 2015; Kaplan and Norton, 2008; Lipe and Salterio, 2000; Park et al., 2017; Sohn et al., 2003; Stewart, 2001; Yoshikuni et al., 2014).

CUPE specifies how a firm can create value for the market (León-Soriano et al., 2010; Mostaghel et al., 2015; Sohn et al., 2003; Stewart, 2001). Specifically, a firm can promote customer satisfaction by delivering desired product attributes to the market, thereby, demonstrating added value and improving customer retention (Bento et al., 2013; Kaplan and Norton, 2008; Lipe and Salterio, 2000; Mithas et al., 2011; Norreklit, 2000; Park et al., 2017; Reefeke and Trocchi, 2013; Stewart, 2001; Yoshikuni et al., 2014).

Internal process performance (IPPE) identifies activities in the value chain that transform assets into benefits for clients and shareholders (Mithas et al., 2011; Norreklit, 2000; Park et al., 2017; Perkins et al., 2014; Sohn et al., 2003). Generally, researchers consider that “internal business processes” encompass all activities in the internal value chain. This perspective dictates that three processes are common to all firms: innovation, operations and post-sales (Callado and Jack, 2015; Kaplan and Norton, 2008; Lipe and Salterio, 2000; Park et al., 2017; Reefeke and Trocchi, 2013).

Finally, the organizational growth and learning perspective (GLPE) indicates how intangible assets are aligned and integrated to create organizational value (Bento et al., 2013; Lipe and Salterio, 2000; Mithas et al., 2011; Park et al., 2017; Perkins et al., 2014; Reefeke and Trocchi, 2013). This component is measured based on human capital (i.e. employee training), information capital (i.e. IT/IS support and alignment with strategy) and organizational capital (i.e. corporate cultural attitudes) (Kaplan and Norton, 2008; Park et al., 2017; Sohn et al., 2003; Stewart, 2001; Yoshikuni and Albertin, 2017; Yoshikuni et al., 2014).

3. Research model, hypothesis development and control variables
Based on the literature review, it was postulated that SIS-embedded strategic planning (Whittington, 2014) enables the competitive strategy described by Miles et al. (1978) and influences corporate performance (Kaplan and Norton, 1992). The model is presented in Figure 1 with the hypotheses.
3.1 Strategic information systems and strategic orientation

Strategic planning helps companies develop capabilities to achieve organizational objectives (Mintzberg et al., 2009; Wolf and Floyd, 2017). A company’s strategic orientation defines the stance it adopts to achieve these objectives (Miles et al., 1978). Strategic orientations are characterized as analyzer, prospector, defender and reactor. Thus, the SIS-embedded strategic planning process enables the business strategy’s content (Chan et al., 2006; Whittington, 2014).

Porter (1990), a seminal study of competitive strategy, defines strategic positioning (content strategy of cost leadership and differentiation) as dependent on an effective strategic planning process. Mintzberg et al. (2009) define content strategies as deliberate planning (goal formation, alternative generation and choice), and (or) emergent forces (decisions and actions arising within an organization) in that they are developed by strategic activities disseminated across all organizational levels. Both of these studies demonstrate that strategic content comprises outcomes influenced by a strategic business process - either more or less formal. According to recent strategy theory research of Wolf and Floyd (2017) on the landscape of strategic planning, practice strategy is defined as a dimension that promotes strategic planning effectiveness, thereby, enabling content strategy and impacting organizational performance. Then, strategic practices refer to the routines and norms of strategic work, which are included in the strategic planning process by dimensions of practitioners, praxis and technologies (Whittington, 2014). Therefore, the proposal model of strategic planning defined by Wolf and Floyd (2017) describes strategic practices as antecedents to support content strategy and influence performance by IT/IS resources (SIS).

Hence, as mentioned earlier (Section 2.1), the SIS is a set of IS applications and IT resources that is embedded in strategic planning and enables organization to execute their business strategies in practice (Marabelli and Galliers, 2017; Whittington, 2014) to impact firm performance (Melville et al., 2004). SIS provides a wide range of information on strategic planning routines, enabling an organization’s participatory capacity to think, analyze, formulate, coordinate and monitor business strategy (Singh et al., 2002; Yoshikuni and Jeronimo, 2013).

According to Chen et al. (2010) and Merali et al. (2012), effective use of SIS helps organizations successfully execute the strategic planning process. These authors argue that successful strategic planning implies the company’s ability to effectively promote its employees’ cooperative work in thinking, analyzing and developing strategies supported by IS. SIS facilitates the strategic awareness phase by promoting communication, integration and cooperation from top to bottom, and from bottom to top (Chen et al., 2010; Jääskeläinen and Luukkanen, 2017; Segars and Grover, 1999), and without boundaries – local or

![Diagram of Proposed SIS model of competitive strategy and performance](image-url)
global – so that all employees understand the strategic priorities (Newkirk et al., 2003; Newkirk and Lederer, 2006; O’Brien and Marakas, 2007). SIS promotes organizational commitment through teamwork (Chan and Reich, 2007; Segars and Grover, 1999). In light of Wolf and Floyd (2017), any business strategy’s success depends on employees’ understanding of the strategy’s development and execution. Therefore, SIS has relevant dissemination awareness strategies for strategic orientations, such as defender, analyzer and prospector (Bernado et al., 2017; Jääskeläinen and Luukkanen, 2017; King, 1978).

SIS enables the strategic planning process to map external factors from the general environment (Newkirk and Lederer, 2006) and to develop strategies to capture opportunities and mitigate threats (Kaplan and Norton, 2008; Porter, 1986). According to Sabherwal and Chan (2001), SIS promotes flexibility for prospectors to monitor their product/market trends more closely and to spend more on marketing than defenders do; it also promotes flexibility for analyzers to accomplish imitation successfully through extensive marketing surveillance. According to Chen et al. (2010) and Leinher et al. (2011), SIS has a similar influence on innovation and conservative strategies as it does on prospector and defender strategies.

Pavlou and El Sawy (2010) find that effective use of SIS allows organizations to respond in real time to external environmental challenges by reconfiguring existing resources. SIS supports strategic planning conception by developing dynamic capabilities for reconfiguring existing operational skills to respond to environmental changes better (Yoshikuni and Albertin, 2017). Moreover, it enables capabilities to reconfigure existing resources spontaneously in order to build new operational capabilities and address urgent, unpredictable and new environmental situations (Pavlou and El Sawy, 2006, 2010). Therefore, among defenders, SIS contributes to long-term planning, or futurity, which is a key characteristic of that configuration. Among prospectors, SIS contributes to proactiveness by helping the organization to develop strategic decisions quickly and effectively. Finally, among analyzers, SIS contributes to levels of internal and external analyses conducted by organizations (Chan et al., 2006; Sabherwal and Chan, 2001).

SIS enables flexibility and agility during strategic planning formulation (Jääskeläinen and Luukkanen, 2017; Kearns and Sabherwal, 2006) and implementation of business strategies, with an emphasis on operational efficiency and flexibility for innovation (Chen et al., 2010; Marabelli and Galliers, 2017). According to Chan and Huff (1992), SIS can support business strategies, such as aggressiveness, analysis, proactiveness, risk-taking defensiveness and futurity/innovativeness. According to Gupta et al. (1997) and Sabherwal and Chan (2001), strategies described by Chan and Huff (1992) reflect three types of SIS that correspond to the defenders, analyzers and prospectors. SIS supports efficiency, flexibility and comprehensiveness, which align with the defender, prospector and analyzer business strategies, respectively. Thus, SIS supports defenders’ emphasis on cost containment, prospectors’ desire for flexibility and innovation and analyzers’ endeavors to achieve efficiency and innovation simultaneously.

Firms need to know how a strategy is working and why it might not be working, and thus, the monitoring phase is necessary for all strategies (Mintzberg et al., 2009). Hence, as mentioned before (Section 2.1), SIS is a set of IT/IS applications that collect, process, analyze and provide data/information for decision making. This enables a holistic, interactive, decentralized and dynamic view of the organization and generates organizational knowledge and learning in the strategic planning process (Yoshikuni and Jeronimo, 2013). This way, the prospector orientation supported by SIS enables firm creativity by generating new products and services as well as new business approaches (Chan and Reich, 2007; Gupta et al., 1997; Sabherwal and Chan, 2001). Moreover, the defender orientation is promoted by SIS for control, that is, for organizational efficiency and productivity (Chan et al., 2006; Martinez-Simarro et al., 2015; Philip, 2007), and for creation of a stable condition for maintaining current products and customer relationships (Chan and Reich, 2007;
Gupta et al., 1997). Finally, SIS contributes to high levels of internal (production) and external (marketing) analyses, enabling comprehensive decision making to develop analyzer strategies (Chan, 1997; Croteau and Bergeron, 2001; Sabherwal and Chan, 2001). Therefore, planning capabilities provided by SIS result in creation of value and benefits for the strategic planning process by disseminating strategic awareness; analyzing external factors; promoting cooperation for conception; and developing, implementing and monitoring competitive strategies (Mentzas, 1997; Newkirk et al., 2003; Newkirk and Lederer, 2006) for defenders, analyzers and prospectors (Miles et al., 1978). Thus, the following hypotheses were postulated:

**H1a.** SIS is positively associated with the adoption of the analyzer strategic orientation.

**H1b.** SIS is positively associated with the adoption of the defender strategic orientation.

**H1c.** SIS is positively associated with the adoption of the prospector strategic orientation.

It is understood that reactor firms do not have a clear, consistent strategy (Chan, 1997; Sabherwal and Chan, 2001), do not use IS strategically (Gupta et al., 1997) and are characterized by extreme organizational inertia and respond to environmental pressures only when forced (Anwar and Hasnu, 2016). Thus, the following was hypothesized:

**H1d.** SIS is not positively associated with the adoption of the reactor strategic orientation.

Hence, an SIS-embedded strategic planning process (Newkirk and Lederer, 2006; Singh et al., 2002) enables competitive strategy content for defenders, analyzers and prospectors (Miles et al., 1978; Sabherwal and Chan, 2001) in all its stages (strategic awareness, environmental analysis, development and business strategy execution and monitoring).

### 3.2 Strategic orientation and corporate performance

A business strategy is the outcome of decision making that guides an organization with respect to the environment, structures and processes to improve corporate performance (Bernado et al., 2017; Croteau and Bergeron, 2001; Hill et al., 2014; Mintzberg et al., 2009). A business strategy defines a company’s long-term plan to achieve its goals.

The relationship between strategy and performance has been examined in numerous works, both theoretically and empirically (Anwar and Hasnu, 2016; Chatzoglou et al., 2017). Many studies show that Miles and Snow’s assumption of its effect of improving firm performance in the long run is overwhelming (Anwar and Hasnu, 2016; Parnell, 2010). Drawing from extant research and empirical findings showing that a firm’s strategic orientation affects its corporate performance, this study tested the following hypotheses:

**H2a.** The analyzer orientation positively affects corporate performance.

**H2b.** The defender orientation positively affects corporate performance.

**H2c.** The prospector orientation positively affects corporate performance.

However, the effect on reactors is uncertain or inappropriate and is generally linked with poor performance (Avci et al., 2011; Conant et al., 1990; Moore, 2005). Moreover, several studies demonstrate that the defender, analyzer and prospector strategies outperform the reactor strategy (Anwar and Hasnu, 2016; Parnell, 2010; Parnell et al., 2015).

Nevertheless, recent studies on orientation strategy in developing countries, such as China and Turkey, show that reactor strategies have a negative effect on firm performance (e.g. Parnell et al., 2012). Drawing from extant research and empirical findings, the following hypothesis was tested:

**H2d.** The reactor orientation negatively affects corporate performance.
3.3 Strategic information systems and corporate performance

Over the past 30 years, studies have continued to build on empirical evidence that reveals positive effects of IS strategy on corporate performance (Gerow et al., 2015; Sabherwal and Chan, 2001; Yayla and Hu, 2012). These studies show that organizations perform well when key IT/IS resources are aligned to support effective management of business strategy (Colman et al., 2015).

According to Gerow et al. (2015), the SIS literature frequently emphasizes the positive aspects of alignment in theoretical frameworks and empirical research on firm performance (e.g. increased sales revenue, improved operational efficiency, cost reductions and enhanced customer value). Recent studies demonstrate that IT/IS resources enable firm capabilities with tacit, socially complex firm-specific resources that are shown to enhance the customer value proposition and explain variations in FIPE (Merali et al., 2012).

Wade and Hulland (2004) suggest in their study of strategy and IS based on the resource-based view theory that IT/IS resources directly and indirectly influence competitive position and performance. Hence, once IT/IS resources are embedded in the organizational structure, firms can use them to create dynamic capabilities to renew and re-invent their organizations’ resource base in order to adapt to the changing competitive context and, to re-position themselves to maintain or improve their competitive positioning (Merali et al., 2012; Pavlou and El Sawy, 2010). Yoshikuni and Albertin (2017), in a recent study, investigate the strong effects of IT/IS resource-enabled dynamic capabilities on business process improvement in order to understand customer needs and impacts on FIPE. Based on these arguments, the following hypothesis was developed:

\[ H3. \text{ SIS is positively associated with corporate performance.} \]

3.4 Control variables

According to Chan et al. (2006) and Melville et al. (2004), organizational size (i.e. number of employees (SIZE)) and sector are industry characteristics that influence the relationship between SIS and corporate performance. Anwar and Hasnu (2016) and Parnell et al. (2012) demonstrate that there are different effects on the relationship between orientation strategy and performance based on different sectors and firm sizes. Two hypotheses were proposed to examine the moderation effects on exogenous and endogenous variables:

\[ H4a. \text{ The effect of SIS on the orientation strategy is moderated by control variables.} \]
\[ H4b. \text{ The effect of the orientation strategy on firm performance is moderated by control variables.} \]
\[ H4c. \text{ The effect of SIS on firm performance is moderated by control variables.} \]

4. Methodology

This section describes the sample, data and analytic methods.

4.1 Sample

After a thorough literature review, a survey instrument was developed (see Appendix for details) and a pre-test survey was conducted to check the clarity of the items' contents, response time and related observations (Kim et al., 2011; Yayla and Hu, 2012). The respondents were three IT/IS researchers and two researchers from the business field. Table I summarizes the measures and sources of the variables used in the analyses. Then, content validity of the instrument was tested through a pilot test with 42 organizational informants, who were Executive Master of Business Administration (EMBA) professionals enrolled in a large university in Brazil (D’Arcy and Devaraj, 2012; Yayla and Hu, 2012).
The sample was selected from Brazilian companies using directories provided by the Center for Applied Information Technology (GVCia) of Fundação Getulio Vargas[1]. Key informant methodology was used to obtain the sample, and respondents were chosen based on their position, experience and professional knowledge (Kim et al., 2011). The target respondents included senior business administrators with adequate knowledge of IS and business strategizing processes.

The survey was administered via e-mail, and questionnaires were distributed to 1,577 organizations. Respondents had two weeks to respond, during which they could review the questionnaire with other company executives. Respondents could resolve possible doubts with the authors by e-mail or telephone. A total of 394 (23 percent) questionnaires were returned. Among them, 47 had missing data; these responses were removed from the analyses, yielding a final sample size of 389.

To identify potential bias in the subsamples, the organization groups of “EMBA” and “GVCia” were compared with those of the final sample ($n = 389$). All $t$-tests comparing the responses provided by these two samples showed no significant differences. A dummy variable was included in the model to represent the sample (i.e. 42 organizations vs the main study) of which a participant was a part. The path from the dummy variable to corporate performance was not significant ($\beta = 0.01$, $p > 0.05$). The Statistical Package for the Social Sciences (SPSS, Version 20.0) was used for all analyses, and any incomplete response sets from the final data set were eliminated.

4.2 Data treatment
Past studies have shown that research using partial-least-squares path modeling (PLS-PM) methods must use a sample of no less than ten times the number of structural paths arriving at a given reflexive construct (Henseler et al., 2009; Urbach and Ahlemann, 2010). Given this rule and the conceptual model, the minimum sample size in this study was 40 respondents. The sample of 389 far exceeds this threshold.

Corporate performance (CP) is a latent, second-order variable composed of multiple reflective constructs, including FIPE, CUPE, IPPE and GLPE, according to Yoshikuni et al. (2014). Moreover, corporate performance was modeled as a latent, second-order variable according to the guidelines of Bento et al. (2013), Wetzel et al. (2009) and Wold (1982). This allowed the execution of the PLS-PM algorithm.

Finally, statistical techniques were applied to detect and (where possible) control for common method bias. Consistent with Chin et al. (2013), the measured latent marker variable (MLMV) technique was used for the model at the corporate performance level.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measure</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategic information systems</td>
<td>IS support strategic planning</td>
<td>Singh et al. (2002), Newkirk and Lederer (2006)</td>
</tr>
<tr>
<td></td>
<td>Analyzer</td>
<td></td>
</tr>
<tr>
<td>Strategic orientation</td>
<td>Defender</td>
<td>Miles et al. (1978)</td>
</tr>
<tr>
<td></td>
<td>Prospector</td>
<td>Moore (2005)</td>
</tr>
<tr>
<td></td>
<td>Reactor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Financial</td>
<td></td>
</tr>
<tr>
<td>Corporate performance</td>
<td>Internal process</td>
<td>Kaplan and Norton (1992)</td>
</tr>
<tr>
<td></td>
<td>Organizational learning and growth</td>
<td>Yoshikuni et al. (2014)</td>
</tr>
<tr>
<td>Control variables</td>
<td>Organizational size (i.e. number of employees (SIZE)) and sector</td>
<td>Anwar and Hasnu (2016), Chan et al. (2006), Melville et al. (2004)</td>
</tr>
</tbody>
</table>

Table I. Measures and sources of variables used in the analyses

Effects of SISs on competitive strategy
4.3 Analysis

PLS-PM was used to analyze all variables and evaluate the relationships among them. PLS-PM is a well-established method for simultaneous analysis of multiple variables (e.g. asymmetric variable distributions and limited data) (see Ringle et al., 2012, 2014). The SmartPLS 2.0 M3 program was used to perform all PLS-PM analyses (Ringle et al., 2005).

In evaluating the normality of each measure, none was found to be sufficiently non-normal to warrant correction. All skewness values were less than 3, and all kurtosis levels were less than 10 (see Marôco, 2010). Table II summarizes these statistics.

Of the respondents who returned questionnaires, 30 percent were C-level executives (e.g. chief executive officers), 37 percent were management and coordination personnel and 33 percent were supervisors with decision-making powers. Of the firms represented in the sample, 13 firms (3 percent) were engaged in agribusiness (generic value chain of these businesses related to agriculture and livestock). Moreover, 100 firms (26 percent) were involved in manufacturing of durable goods (non-perishable goods, such as cars, household appliances and furniture) and non-durable goods (commodities for basic needs, such as food, drink, clothes, shoes and cosmetics). The remaining 276 firms (71 percent) were services providers (corresponding to trade activities in goods and provision of services, such as merchandise trade to public administration, transportation, financial and real estate activities, business or personal services, education and health and social promotion). Of the represented firms, 3 percent had fewer than 9 employees at the time of data collection; 8 percent had 10–49 employees; 9 percent had 50–99 employees; 11 percent had 100–249 employees; 11 percent had 250–499 employees; and 58 percent had more than 500 employees. The sample was heavily populated by firms in the services and manufacturing sectors (97 percent).

5. Results

This section describes the analysis results for the measurement and structural models as well as the reflective constructs. This section also presents the results of the hypothesis tests.

5.1 Measurement model

The reflective constructs in the measurement model were evaluated by checking their internal consistency, indicator reliability, convergent validity and discriminant validity (Hair et al., 2013; Henseler et al., 2016).

![Table II. Pearson correlations and descriptive statistics for latent variables](image-url)
The Fornell–Larcker (1981) criterion was used to evaluate the convergent validity and average variance extracted (AVE; values greater than 0.50 are preferable) of the reflective constructs (Henseler et al., 2009; Ringle et al., 2014). Composite reliability (CR) is the most reasonable measure of reliability for PLS-PM, because it prioritizes variables according to their respective reliabilities (Ringle et al., 2014). CR values greater than 0.70 are considered internally consistent (D’Arcy and Devaraj, 2012; Henseler et al., 2016). To analyze the validity of the model’s constructs, the Fornell–Larcker criterion was compared with the square root of the constructs’ AVE values with highest latent variable correlation with any other construct (Henseler et al., 2009) (see Table II). Furthermore, a bootstrapping method with 1,000 replications was used to determine the statistical significance of the tests.

The convergent validity and square root of the strategy indicator’s AVE (the value of which is on the diagonal) was also evaluated. Table II shows that all AVE and internal consistency values (which should exceed 0.70) were acceptable. Moreover, indicators with outer loadings between 0.50 and 0.70 were considered because of the increases in AVE and CR beyond the threshold values suggested by Hair et al. (2013). Table III shows that the indicators have higher factor loadings on their assigned constructs and lower factor loadings on other constructs, thereby, indicating discriminant validity (Chin, 1998; Ringle et al., 2014; Urbach and Ahlemann, 2010).

The second-order corporate performance variable yielded an AVE value of 0.63 and CR estimate of 0.89. A comparison of the Fornell–Larcker criterion with the square root of corporate performance (0.944) AVE values shows the criterion to be satisfied.

<table>
<thead>
<tr>
<th>First latent variable</th>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – SIS</td>
<td>SIS 1</td>
<td>0.706</td>
<td>0.355</td>
<td>0.314</td>
<td>0.461</td>
<td>-0.243</td>
<td>0.401</td>
<td>0.355</td>
<td>0.444</td>
<td>0.517</td>
</tr>
<tr>
<td></td>
<td>SIS 2</td>
<td>0.733</td>
<td>0.209</td>
<td>0.171</td>
<td>0.276</td>
<td>-0.157</td>
<td>0.273</td>
<td>0.255</td>
<td>0.417</td>
<td>0.355</td>
</tr>
<tr>
<td></td>
<td>SIS 3</td>
<td>0.864</td>
<td>0.335</td>
<td>0.251</td>
<td>0.386</td>
<td>-0.201</td>
<td>0.361</td>
<td>0.279</td>
<td>0.510</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>SIS 4</td>
<td>0.815</td>
<td>0.328</td>
<td>0.283</td>
<td>0.309</td>
<td>-0.199</td>
<td>0.356</td>
<td>0.275</td>
<td>0.466</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>SIS 5</td>
<td>0.764</td>
<td>0.319</td>
<td>0.248</td>
<td>0.389</td>
<td>-0.204</td>
<td>0.311</td>
<td>0.286</td>
<td>0.396</td>
<td>0.269</td>
</tr>
<tr>
<td>2 – Analyzer SO</td>
<td>ANAL 1</td>
<td>0.187</td>
<td>0.613</td>
<td>0.406</td>
<td>0.201</td>
<td>0.039</td>
<td>0.186</td>
<td>0.223</td>
<td>0.232</td>
<td>0.186</td>
</tr>
<tr>
<td></td>
<td>ANAL 2</td>
<td>0.268</td>
<td>0.659</td>
<td>0.363</td>
<td>0.468</td>
<td>-0.108</td>
<td>0.207</td>
<td>0.177</td>
<td>0.374</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>ANAL 3</td>
<td>0.375</td>
<td>0.843</td>
<td>0.537</td>
<td>0.387</td>
<td>-0.076</td>
<td>0.339</td>
<td>0.381</td>
<td>0.403</td>
<td>0.370</td>
</tr>
<tr>
<td>3 – Defender SO</td>
<td>DEFE 1</td>
<td>0.254</td>
<td>0.474</td>
<td>0.752</td>
<td>0.243</td>
<td>-0.035</td>
<td>0.270</td>
<td>0.312</td>
<td>0.257</td>
<td>0.261</td>
</tr>
<tr>
<td></td>
<td>DEFE 2</td>
<td>0.124</td>
<td>0.257</td>
<td>0.550</td>
<td>0.049</td>
<td>0.110</td>
<td>0.171</td>
<td>0.168</td>
<td>0.139</td>
<td>0.142</td>
</tr>
<tr>
<td></td>
<td>DEFE 3</td>
<td>0.291</td>
<td>0.524</td>
<td>0.800</td>
<td>0.539</td>
<td>-0.133</td>
<td>0.311</td>
<td>0.368</td>
<td>0.415</td>
<td>0.307</td>
</tr>
<tr>
<td>4 – Prospector SO</td>
<td>PROS 1</td>
<td>0.307</td>
<td>0.345</td>
<td>0.384</td>
<td>0.810</td>
<td>-0.244</td>
<td>0.285</td>
<td>0.212</td>
<td>0.429</td>
<td>0.225</td>
</tr>
<tr>
<td></td>
<td>PROS 2</td>
<td>0.450</td>
<td>0.450</td>
<td>0.423</td>
<td>0.848</td>
<td>-0.307</td>
<td>0.333</td>
<td>0.377</td>
<td>0.507</td>
<td>0.371</td>
</tr>
<tr>
<td></td>
<td>PROS 3</td>
<td>0.382</td>
<td>0.460</td>
<td>0.374</td>
<td>0.860</td>
<td>-0.271</td>
<td>0.292</td>
<td>0.318</td>
<td>0.478</td>
<td>0.305</td>
</tr>
<tr>
<td>5 – Reactor SO</td>
<td>REAC 1</td>
<td>-0.213</td>
<td>-0.045</td>
<td>-0.051</td>
<td>-0.209</td>
<td>0.759</td>
<td>-0.104</td>
<td>-0.103</td>
<td>-0.230</td>
<td>-0.181</td>
</tr>
<tr>
<td></td>
<td>REAC 2</td>
<td>-0.196</td>
<td>-0.090</td>
<td>-0.091</td>
<td>-0.228</td>
<td>0.764</td>
<td>-0.144</td>
<td>-0.130</td>
<td>-0.192</td>
<td>-0.163</td>
</tr>
<tr>
<td></td>
<td>REAC 3</td>
<td>-0.148</td>
<td>-0.028</td>
<td>0.024</td>
<td>-0.317</td>
<td>0.616</td>
<td>-0.036</td>
<td>0.007</td>
<td>-0.164</td>
<td>-0.033</td>
</tr>
<tr>
<td>6 – Financial performance</td>
<td>FIPE 1</td>
<td>0.367</td>
<td>0.295</td>
<td>0.320</td>
<td>0.311</td>
<td>-0.114</td>
<td>0.923</td>
<td>0.449</td>
<td>0.494</td>
<td>0.370</td>
</tr>
<tr>
<td></td>
<td>FIPE 2</td>
<td>0.416</td>
<td>0.329</td>
<td>0.320</td>
<td>0.314</td>
<td>-0.086</td>
<td>0.887</td>
<td>0.434</td>
<td>0.518</td>
<td>0.380</td>
</tr>
<tr>
<td></td>
<td>FIPE 3</td>
<td>0.433</td>
<td>0.351</td>
<td>0.359</td>
<td>0.366</td>
<td>-0.161</td>
<td>0.921</td>
<td>0.492</td>
<td>0.553</td>
<td>0.427</td>
</tr>
<tr>
<td>7 – Customer performance</td>
<td>CUPE 1</td>
<td>0.271</td>
<td>0.305</td>
<td>0.356</td>
<td>0.254</td>
<td>-0.111</td>
<td>0.350</td>
<td>0.773</td>
<td>0.469</td>
<td>0.318</td>
</tr>
<tr>
<td></td>
<td>CUPE 2</td>
<td>0.222</td>
<td>0.272</td>
<td>0.321</td>
<td>0.248</td>
<td>-0.059</td>
<td>0.473</td>
<td>0.806</td>
<td>0.547</td>
<td>0.340</td>
</tr>
<tr>
<td></td>
<td>CUPE 3</td>
<td>0.382</td>
<td>0.299</td>
<td>0.282</td>
<td>0.343</td>
<td>-0.114</td>
<td>0.312</td>
<td>0.692</td>
<td>0.461</td>
<td>0.426</td>
</tr>
<tr>
<td>8 – Internal process</td>
<td>IPPE 1</td>
<td>0.467</td>
<td>0.252</td>
<td>0.216</td>
<td>0.241</td>
<td>-0.207</td>
<td>0.440</td>
<td>0.450</td>
<td>0.720</td>
<td>0.496</td>
</tr>
<tr>
<td></td>
<td>IPPE 2</td>
<td>0.410</td>
<td>0.389</td>
<td>0.353</td>
<td>0.336</td>
<td>-0.137</td>
<td>0.468</td>
<td>0.609</td>
<td>0.816</td>
<td>0.446</td>
</tr>
<tr>
<td></td>
<td>IPPE 3</td>
<td>0.408</td>
<td>0.441</td>
<td>0.355</td>
<td>0.735</td>
<td>-0.289</td>
<td>0.352</td>
<td>0.358</td>
<td>0.670</td>
<td>0.390</td>
</tr>
<tr>
<td>9 – Growth and learning</td>
<td>GLPE 1</td>
<td>0.223</td>
<td>0.219</td>
<td>0.223</td>
<td>0.125</td>
<td>-0.094</td>
<td>0.265</td>
<td>0.290</td>
<td>0.318</td>
<td>0.682</td>
</tr>
<tr>
<td></td>
<td>GLPE 2</td>
<td>0.411</td>
<td>0.291</td>
<td>0.252</td>
<td>0.228</td>
<td>-0.107</td>
<td>0.344</td>
<td>0.341</td>
<td>0.496</td>
<td>0.757</td>
</tr>
<tr>
<td></td>
<td>GLPE 3</td>
<td>0.444</td>
<td>0.337</td>
<td>0.307</td>
<td>0.428</td>
<td>-0.220</td>
<td>0.350</td>
<td>0.420</td>
<td>0.511</td>
<td>0.798</td>
</tr>
</tbody>
</table>

Table III. Cross-loadings to determine discriminant validity of the first model.
5.2 Structural model

To test for multicollinearity among the model’s constructs, their variance inflation factor (VIF) values were evaluated. All VIF values were well below Marôco’s (2010) recommended limit of 5, suggesting that there was no multicollinearity among the variables in the data.

Tables IV–VI (Cases 1–3) show the moderating effects of all latent and control variables on the relationships between the exogenous and endogenous variables. First, the relationships between all latent variables are statistically significant \((p < 0.05\); see Table IV, Case 1). Although SIS and adoption of the reactor strategic orientation are negatively related \((\beta = -0.264; p < 0.001)\), the relationships between SIS and adoption of other strategic orientations are positive \((\beta_{\text{analyzer}} = 0.408, \beta_{\text{defender}} = 0.335, \beta_{\text{prospector}} = 0.482; \text{all } p < 0.001)\). This pattern of effects is also evident for corporate performance. The results show the adoption of the reactor strategic orientation to be negatively associated with corporate performance \((\beta = -0.103; p < 0.05)\), but that of the analyzer \((\beta = 0.221)\), defender \((\beta = 0.207)\) and prospector \((0.275)\) strategic orientations to be positively related to corporate performance \((p < 0.001)\).

By including a direct path between SIS and CP, the strong positive relationship between the two variables is verified \((\beta = 0.376, p < 0.001)\). Moreover, the variance explained by this new model \((R^2)\) increased from 37.4 to 47.3 percent (see Table V, Case 2).

The control variables (sector and size) had no moderating effects on the significant relationships described above (see Table VI, Case 3). However, firm size seems to have moderated the (originally non-significant) relationship between the reactor strategic orientation and corporate performance \((\beta = -0.125; p < 0.05)\).

5.3 Controlling for common method bias

Chin et al.’s (2013) MLMV technique is applied to control for common method bias (Table VII, Case 4), because this study used one instrument to obtain data from single

<table>
<thead>
<tr>
<th>Relationship</th>
<th>( \beta )</th>
<th>SE</th>
<th>( t )</th>
<th>( p )</th>
<th>( R^2 ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIS ( \rightarrow ) Analyzer SO</td>
<td>0.408</td>
<td>0.042</td>
<td>9.714</td>
<td>0.000</td>
<td>16.7</td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Defender SO</td>
<td>0.335</td>
<td>0.045</td>
<td>7.480</td>
<td>0.000</td>
<td>11.2</td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Prospector SO</td>
<td>0.482</td>
<td>0.039</td>
<td>12.471</td>
<td>0.000</td>
<td>23.2</td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Reactor SO</td>
<td>-0.264</td>
<td>0.052</td>
<td>5.065</td>
<td>0.000</td>
<td>7.0</td>
</tr>
<tr>
<td>Analyzer SO ( \rightarrow ) CP</td>
<td>0.221</td>
<td>0.058</td>
<td>3.835</td>
<td>0.000</td>
<td>37.4</td>
</tr>
<tr>
<td>Defender SO ( \rightarrow ) CP</td>
<td>0.207</td>
<td>0.055</td>
<td>3.759</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Prospector SO ( \rightarrow ) CP</td>
<td>0.275</td>
<td>0.060</td>
<td>4.555</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Reactor SO ( \rightarrow ) CP</td>
<td>-0.103</td>
<td>0.045</td>
<td>2.268</td>
<td>0.023</td>
<td></td>
</tr>
</tbody>
</table>

Table IV. Case 1: relationships between all latent variables

<table>
<thead>
<tr>
<th>Relationship</th>
<th>( \beta )</th>
<th>SE</th>
<th>( t )</th>
<th>( p )</th>
<th>( R^2 ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIS ( \rightarrow ) Analyzer SO</td>
<td>0.406</td>
<td>0.044</td>
<td>9.177</td>
<td>0.000</td>
<td>16.5</td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Defender SO</td>
<td>0.334</td>
<td>0.047</td>
<td>7.152</td>
<td>0.000</td>
<td>11.2</td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Prospector SO</td>
<td>0.479</td>
<td>0.041</td>
<td>11.830</td>
<td>0.000</td>
<td>23.0</td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Reactor SO</td>
<td>-0.263</td>
<td>0.052</td>
<td>5.063</td>
<td>0.000</td>
<td>6.9</td>
</tr>
<tr>
<td>Analyzer SO ( \rightarrow ) CP</td>
<td>0.142</td>
<td>0.054</td>
<td>2.629</td>
<td>0.009</td>
<td>47.3</td>
</tr>
<tr>
<td>Defender SO ( \rightarrow ) CP</td>
<td>0.188</td>
<td>0.050</td>
<td>3.733</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Prospector SO ( \rightarrow ) CP</td>
<td>0.159</td>
<td>0.058</td>
<td>2.772</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>Reactor SO ( \rightarrow ) CP</td>
<td>-0.094</td>
<td>0.040</td>
<td>1.226</td>
<td>0.220</td>
<td></td>
</tr>
<tr>
<td>SIS ( \rightarrow ) CP</td>
<td>0.376</td>
<td>0.045</td>
<td>8.377</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

Table V. Case 2: relationships between all latent variables and the direct effect of SIS on CP
respondents at a single point in time. Specifically, four items designed to have the lowest possible correlation with the other constructs under investigation were used (see the list "Formative indicators used for MLMV analysis" below). These items were intended to capture common method variance, if any existed within the data.

Formative indicators used for MLMV analysis:

(1) MLMV_1: It is easy for me to reach my goals.

(2) MLMV_2: I would never abandon the desire to have my own business.

(3) MLMV_3: I have a positive attitude toward others.

(4) MLMV_4: I always imagine my house in the future.

To verify the MLMV analysis results, the differences between path coefficients across two groups (Cases 3 and 4) were explored. A parametric approach to the PLS-multi-group analysis (PLS-MGA, Hair et al., 2013) was used with the specification of group-specific path coefficients, standard errors (obtained from a 5,000-case bootstrapping procedure) and

<table>
<thead>
<tr>
<th>Relationship</th>
<th>$\beta$</th>
<th>SE</th>
<th>$t$</th>
<th>$p$</th>
<th>$R^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIS $\rightarrow$ Analyzer SO</td>
<td>0.372</td>
<td>0.049</td>
<td>7.597</td>
<td>0.000</td>
<td>19.20</td>
</tr>
<tr>
<td>Sector $\rightarrow$ Analyzer SO</td>
<td>0.019</td>
<td>0.049</td>
<td>0.378</td>
<td>0.706</td>
<td></td>
</tr>
<tr>
<td>Size $\rightarrow$ Analyzer SO</td>
<td>0.066</td>
<td>0.046</td>
<td>1.435</td>
<td>0.151</td>
<td></td>
</tr>
<tr>
<td>SIS $\times$ Sector $\rightarrow$ Analyzer SO</td>
<td>-0.106</td>
<td>0.124</td>
<td>0.855</td>
<td>0.392</td>
<td></td>
</tr>
<tr>
<td>SIS $\rightarrow$ Defender SO</td>
<td>0.313</td>
<td>0.051</td>
<td>6.087</td>
<td>0.000</td>
<td>12.30</td>
</tr>
<tr>
<td>Sector $\rightarrow$ Defender SO</td>
<td>-0.014</td>
<td>0.049</td>
<td>0.289</td>
<td>0.773</td>
<td></td>
</tr>
<tr>
<td>Size $\rightarrow$ Defender SO</td>
<td>0.032</td>
<td>0.052</td>
<td>0.616</td>
<td>0.538</td>
<td></td>
</tr>
<tr>
<td>SIS $\times$ Sector $\rightarrow$ Defender SO</td>
<td>0.089</td>
<td>0.049</td>
<td>1.814</td>
<td>0.070</td>
<td></td>
</tr>
<tr>
<td>SIS $\times$ Size $\rightarrow$ Defender SO</td>
<td>-0.045</td>
<td>0.068</td>
<td>0.665</td>
<td>0.506</td>
<td></td>
</tr>
<tr>
<td>SIS $\rightarrow$ Prospector SO</td>
<td>0.438</td>
<td>0.045</td>
<td>9.727</td>
<td>0.000</td>
<td>25.20</td>
</tr>
<tr>
<td>Sector $\rightarrow$ Prospector SO</td>
<td>-0.054</td>
<td>0.067</td>
<td>0.816</td>
<td>0.415</td>
<td></td>
</tr>
<tr>
<td>Size $\rightarrow$ Prospector SO</td>
<td>0.123</td>
<td>0.049</td>
<td>2.455</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>Sector $\times$ Prospector SO</td>
<td>0.044</td>
<td>0.081</td>
<td>0.544</td>
<td>0.587</td>
<td></td>
</tr>
<tr>
<td>Size $\times$ Prospector SO</td>
<td>0.018</td>
<td>0.097</td>
<td>0.185</td>
<td>0.854</td>
<td></td>
</tr>
<tr>
<td>SIS $\rightarrow$ Defender SO 0.267</td>
<td>0.054</td>
<td>4.949</td>
<td>0.000</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td>Sector $\rightarrow$ Defender SO</td>
<td>0.017</td>
<td>0.049</td>
<td>0.337</td>
<td>0.736</td>
<td></td>
</tr>
<tr>
<td>Size $\rightarrow$ Defender SO</td>
<td>0.001</td>
<td>0.054</td>
<td>0.011</td>
<td>0.991</td>
<td></td>
</tr>
<tr>
<td>SIS $\times$ Sector $\rightarrow$ Defender SO</td>
<td>0.047</td>
<td>0.098</td>
<td>0.479</td>
<td>0.632</td>
<td></td>
</tr>
<tr>
<td>SIS $\times$ Size $\rightarrow$ Defender SO</td>
<td>-0.129</td>
<td>0.113</td>
<td>1.143</td>
<td>0.253</td>
<td></td>
</tr>
<tr>
<td>Sector $\rightarrow$ Defender SO</td>
<td>0.137</td>
<td>0.053</td>
<td>2.574</td>
<td>0.010</td>
<td>49.90</td>
</tr>
<tr>
<td>Size $\rightarrow$ Defender SO</td>
<td>0.168</td>
<td>0.048</td>
<td>3.503</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Sector $\times$ Defender SO</td>
<td>0.189</td>
<td>0.055</td>
<td>3.434</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Size $\times$ Defender SO</td>
<td>0.048</td>
<td>0.040</td>
<td>1.194</td>
<td>0.232</td>
<td></td>
</tr>
<tr>
<td>Analyzer SO $\rightarrow$ CP</td>
<td>0.004</td>
<td>0.037</td>
<td>0.105</td>
<td>0.916</td>
<td></td>
</tr>
<tr>
<td>Sector $\rightarrow$ Analyzer SO</td>
<td>0.053</td>
<td>0.041</td>
<td>1.276</td>
<td>0.202</td>
<td></td>
</tr>
<tr>
<td>Size $\rightarrow$ Analyzer SO</td>
<td>0.024</td>
<td>0.063</td>
<td>0.379</td>
<td>0.705</td>
<td></td>
</tr>
<tr>
<td>SIS $\times$ Sector $\rightarrow$ Analyzer SO</td>
<td>-0.001</td>
<td>0.069</td>
<td>0.008</td>
<td>0.994</td>
<td></td>
</tr>
<tr>
<td>SIS $\times$ Size $\rightarrow$ Analyzer SO</td>
<td>0.338</td>
<td>0.044</td>
<td>7.650</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

Table VI. Case 3: interaction effects of the control variables on the relationships among SIS, SO and CP
The normality of the data was verified, and, consistent with Hair et al. (2013), the differences between \( \beta_3 \) and \( \beta_4 \) were observed to be non-significant. Taken together, these results indicate that common method bias is not a concern in the data used (see Table VIII).

The coefficient of determination \( (R^2) \), which measures variance in strategic orientation and corporate performance, provides an indication of the structural model’s predictive power. Cohen (1988) suggests that in social and behavioral sciences, \( R^2 \) values of 0.2, 0.3 and 0.5 represent small, medium and large effects, respectively. As evidenced by the \( R^2 \) values in Case 4, the coefficients of determination indicate that the relationships of the analyzer, defender and prospector strategic orientations are characterized by a large effect. Only the reactor strategic orientation induced a small effect (\( R^2 = 0.99 \) percent).

<table>
<thead>
<tr>
<th>Relationship</th>
<th>( \beta )</th>
<th>SE</th>
<th>( t )</th>
<th>( p )</th>
<th>( R^2 ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIS ( \rightarrow ) Analyzer SO</td>
<td>0.300</td>
<td>0.051</td>
<td>5.919</td>
<td>0.000</td>
<td>24.10</td>
</tr>
<tr>
<td>Sector ( \rightarrow ) Analyzer SO</td>
<td>0.013</td>
<td>0.045</td>
<td>0.289</td>
<td>0.772</td>
<td></td>
</tr>
<tr>
<td>Size ( \rightarrow ) Analyzer SO</td>
<td>0.068</td>
<td>0.045</td>
<td>1.487</td>
<td>0.157</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Sector ( \rightarrow ) Analyzer SO</td>
<td>0.108</td>
<td>0.063</td>
<td>1.716</td>
<td>0.086</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Size ( \rightarrow ) Analyzer SO</td>
<td>-0.060</td>
<td>0.093</td>
<td>0.640</td>
<td>0.522</td>
<td></td>
</tr>
<tr>
<td>MLMV ( \rightarrow ) Analyzer SO</td>
<td>0.242</td>
<td>0.050</td>
<td>4.848</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Defender SO</td>
<td>0.233</td>
<td>0.053</td>
<td>4.324</td>
<td>0.000</td>
<td>18.90</td>
</tr>
<tr>
<td>Sector ( \rightarrow ) Defender SO</td>
<td>-0.022</td>
<td>0.049</td>
<td>0.451</td>
<td>0.652</td>
<td></td>
</tr>
<tr>
<td>Size ( \rightarrow ) Defender SO</td>
<td>0.030</td>
<td>0.053</td>
<td>0.578</td>
<td>0.563</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Sector ( \rightarrow ) Defender SO</td>
<td>0.092</td>
<td>0.052</td>
<td>1.755</td>
<td>0.079</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Size ( \rightarrow ) Defender SO</td>
<td>-0.052</td>
<td>0.064</td>
<td>0.812</td>
<td>0.417</td>
<td></td>
</tr>
<tr>
<td>MLMV ( \rightarrow ) Defender SO</td>
<td>0.067</td>
<td>0.052</td>
<td>1.211</td>
<td>0.226</td>
<td></td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Prospector SO</td>
<td>0.368</td>
<td>0.045</td>
<td>8.114</td>
<td>0.000</td>
<td>30.10</td>
</tr>
<tr>
<td>Sector ( \rightarrow ) Prospector SO</td>
<td>-0.061</td>
<td>0.071</td>
<td>0.896</td>
<td>0.378</td>
<td></td>
</tr>
<tr>
<td>Size ( \rightarrow ) Prospector SO</td>
<td>0.126</td>
<td>0.051</td>
<td>2.485</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Sector ( \rightarrow ) Prospector SO</td>
<td>0.041</td>
<td>0.076</td>
<td>0.540</td>
<td>0.589</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Size ( \rightarrow ) Prospector SO</td>
<td>-0.010</td>
<td>0.081</td>
<td>1.063</td>
<td>0.288</td>
<td></td>
</tr>
<tr>
<td>MLMV ( \rightarrow ) Prospector SO</td>
<td>0.233</td>
<td>0.048</td>
<td>4.908</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>SIS ( \rightarrow ) Reactor SO</td>
<td>-0.233</td>
<td>0.056</td>
<td>4.599</td>
<td>0.000</td>
<td>9.00</td>
</tr>
<tr>
<td>Sector ( \rightarrow ) Reactor SO</td>
<td>0.023</td>
<td>0.053</td>
<td>0.436</td>
<td>0.663</td>
<td></td>
</tr>
<tr>
<td>Size ( \rightarrow ) Reactor SO</td>
<td>0.067</td>
<td>0.056</td>
<td>1.133</td>
<td>0.294</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Sector ( \rightarrow ) Reactor SO</td>
<td>0.043</td>
<td>0.084</td>
<td>0.458</td>
<td>0.647</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Size ( \rightarrow ) Reactor SO</td>
<td>-0.115</td>
<td>0.108</td>
<td>1.063</td>
<td>0.288</td>
<td></td>
</tr>
<tr>
<td>MLMV ( \rightarrow ) Reactor SO</td>
<td>-0.113</td>
<td>0.058</td>
<td>1.942</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>Analyzer SO ( \rightarrow ) CP</td>
<td>0.117</td>
<td>0.054</td>
<td>2.186</td>
<td>0.029</td>
<td>50.80</td>
</tr>
<tr>
<td>Defender SO ( \rightarrow ) CP</td>
<td>0.155</td>
<td>0.049</td>
<td>3.153</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Prospector SO ( \rightarrow ) CP</td>
<td>0.175</td>
<td>0.059</td>
<td>2.958</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Reactor SO ( \rightarrow ) CP</td>
<td>-0.036</td>
<td>0.040</td>
<td>0.890</td>
<td>0.374</td>
<td></td>
</tr>
<tr>
<td>Analyzer SO ( \times ) Sector ( \rightarrow ) CP</td>
<td>-0.037</td>
<td>0.063</td>
<td>0.591</td>
<td>0.554</td>
<td></td>
</tr>
<tr>
<td>Analyzer SO ( \times ) Size ( \rightarrow ) CP</td>
<td>-0.028</td>
<td>0.038</td>
<td>0.649</td>
<td>0.516</td>
<td></td>
</tr>
<tr>
<td>Defender SO ( \times ) Sector ( \rightarrow ) CP</td>
<td>0.065</td>
<td>0.046</td>
<td>1.361</td>
<td>0.173</td>
<td></td>
</tr>
<tr>
<td>Defender SO ( \times ) Size ( \rightarrow ) CP</td>
<td>0.012</td>
<td>0.050</td>
<td>0.244</td>
<td>0.807</td>
<td></td>
</tr>
<tr>
<td>Prospector SO ( \times ) Sector ( \rightarrow ) CP</td>
<td>-0.027</td>
<td>0.052</td>
<td>0.520</td>
<td>0.603</td>
<td></td>
</tr>
<tr>
<td>Prospector SO ( \times ) Size ( \rightarrow ) CP</td>
<td>-0.066</td>
<td>0.083</td>
<td>0.795</td>
<td>0.427</td>
<td></td>
</tr>
<tr>
<td>Reactor SO ( \times ) Sector ( \rightarrow ) CP</td>
<td>0.063</td>
<td>0.073</td>
<td>0.864</td>
<td>0.387</td>
<td></td>
</tr>
<tr>
<td>Reactor SO ( \times ) Size ( \rightarrow ) CP</td>
<td>-0.122</td>
<td>0.048</td>
<td>2.558</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>Sector ( \rightarrow ) CP</td>
<td>0.002</td>
<td>0.039</td>
<td>0.038</td>
<td>0.969</td>
<td></td>
</tr>
<tr>
<td>Size ( \rightarrow ) CP</td>
<td>0.058</td>
<td>0.043</td>
<td>1.343</td>
<td>0.179</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Sector ( \rightarrow ) CP</td>
<td>0.025</td>
<td>0.066</td>
<td>0.377</td>
<td>0.706</td>
<td></td>
</tr>
<tr>
<td>SIS ( \times ) Size ( \rightarrow ) CP</td>
<td>0.010</td>
<td>0.065</td>
<td>0.132</td>
<td>0.879</td>
<td></td>
</tr>
<tr>
<td>SIS ( \rightarrow ) CP</td>
<td>0.327</td>
<td>0.048</td>
<td>6.816</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>MLMV ( \rightarrow ) CP</td>
<td>0.108</td>
<td>0.044</td>
<td>2.431</td>
<td>0.015</td>
<td></td>
</tr>
</tbody>
</table>
5.4 Direct and indirect effects of exogenous and endogenous variables

All possibilities for mediation were evaluated to identify the direct and indirect effects of SIS on CP. First, the direct effect of SIS on CP was estimated. This analysis revealed a strong positive relationship between SIS and CP ($\beta = 0.585; p < 0.001$) and featured a large coefficient of determination ($R^2 = 34.2 \text{ percent}$). Following Zhao et al. (2010), it was concluded that there are likely indirect effects intrinsic to the model as well. Accordingly, the mediator variables from the PLS-PM analysis were included, and the variance accounted for (VAF) associated with each calculation was evaluated (see Table IX).

Owing to significant indirect effects ($p$ value $< 0.001$), the VAF value was analyzed, as it determines the size of the ratio of the indirect effect to the total effect (Preacher and Hayes, 2008). According to Hair et al. (2013), when the VAF is less than 20 percent, there is almost no mediation present. A VAF exceeding 80 percent indicates full mediation. However, a VAF of 20–80 percent suggests partial mediation. The results indicate that none of the strategic orientation types mediate the relationship between SIS and CP.

5.5 Comparing differences between path coefficients in the structural model

According to Hair et al. (2013), the parametric approach is useful for exploring the differences between path coefficients in the structural model. In this vein, the differences between path coefficients associated with the relationship between SIS and strategic orientation variables were evaluated. Table X summarizes the results of these analyses.

Tables XI and XII summarize the results of the analyses of differences between path coefficients associated with the relationship between SIS and strategic orientation variables on corporate performance.

6. Discussion and conclusion

This study investigated the relationship between SIS and strategic orientation, between SIS and corporate performance and between strategic orientation and corporate performance. Furthermore, this study explored how the effective use of SIS to support business strategy affects these outcomes. Specifically, this study explored the effects of SIS on four distinct
strategic orientations – analyzer, defender, prospector and reactor – and the direct effects of strategic orientation types on corporate performance.

The coefficient of determination associated with the inclusion of corporate performance and all strategic orientation types (except reactor) indicated that the model had good explanatory power. Given these findings, this study expands the extant theory and can assist practitioners to use SIS effectively in developing countries during periods of economic turbulence to gain superior corporate performance. The research method used the statistical technique of PLS-PM with SmartPLS software, which was proven an appropriate tool for the analysis.

Table XIII demonstrates the hypotheses and original evidence.

The tests for H1 reveal significant path coefficients, indicating that incorporation of IS into the strategic planning process positively influences the likelihood of a firm adopting an analyzer, defender or prospector strategic orientation. The study demonstrates that SIS provides firm capabilities for disseminating strategic awareness; analyzing external factors; and promoting cooperation for designing, developing, implementing and monitoring competitive strategies (defender, analyzer and prospector). Moreover, the results show that SIS is negatively related to the adoption of the reactor strategic orientation. This finding is consistent with expectations, as reactor firms tend to respond to the competitive environment inconsistently and without the steady use of SIS.

Table X.

<table>
<thead>
<tr>
<th>Endogenous variable (1)</th>
<th>β1</th>
<th>SE</th>
<th>Endogenous variable (2)</th>
<th>β2</th>
<th>SE</th>
<th>(β1−β2)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyzer SO</td>
<td>0.300</td>
<td>0.051</td>
<td>Defender SO</td>
<td>0.233</td>
<td>0.053</td>
<td>0.066</td>
<td>0.898</td>
<td>0.370</td>
</tr>
<tr>
<td>Analyzer SO</td>
<td>0.300</td>
<td>0.051</td>
<td>Prospector SO</td>
<td>0.368</td>
<td>0.045</td>
<td>0.088</td>
<td>1.006</td>
<td>0.316</td>
</tr>
<tr>
<td>Defender SO</td>
<td>0.233</td>
<td>0.053</td>
<td>Reactor SO</td>
<td>−0.233</td>
<td>0.056</td>
<td>0.533</td>
<td>6.671</td>
<td>0.000</td>
</tr>
<tr>
<td>Prospector SO</td>
<td>0.368</td>
<td>0.045</td>
<td>Reactor SO</td>
<td>−0.233</td>
<td>0.056</td>
<td>0.601</td>
<td>7.482</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table XI.

<table>
<thead>
<tr>
<th>SIS vs SO</th>
<th>Analyzer SO</th>
<th>Defender SO</th>
<th>Prospector SO</th>
<th>Reactor SO</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIS</td>
<td>0.376</td>
<td>0.045</td>
<td>0.137</td>
<td>−0.048</td>
</tr>
<tr>
<td>SIS</td>
<td>0.376</td>
<td>0.045</td>
<td>0.168</td>
<td>0.048</td>
</tr>
<tr>
<td>SIS</td>
<td>0.376</td>
<td>0.045</td>
<td>−0.036</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Table XII.

<table>
<thead>
<tr>
<th>Exogenous variable (1, 2)</th>
<th>β1</th>
<th>SE</th>
<th>Exogenous variable (1, 2)</th>
<th>β2</th>
<th>SE</th>
<th>(β1−β2)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyzer SO</td>
<td>0.117</td>
<td>0.054</td>
<td>Defender SO</td>
<td>0.155</td>
<td>0.049</td>
<td>0.038</td>
<td>0.517</td>
<td>0.606</td>
</tr>
<tr>
<td>Analyzer SO</td>
<td>0.117</td>
<td>0.054</td>
<td>Prospector SO</td>
<td>0.175</td>
<td>0.059</td>
<td>0.057</td>
<td>0.720</td>
<td>0.472</td>
</tr>
<tr>
<td>Defender SO</td>
<td>0.155</td>
<td>0.049</td>
<td>Reactor SO</td>
<td>−0.036</td>
<td>0.040</td>
<td>0.153</td>
<td>2.285</td>
<td>0.023</td>
</tr>
<tr>
<td>Defender SO</td>
<td>0.155</td>
<td>0.049</td>
<td>Reactor SO</td>
<td>−0.036</td>
<td>0.040</td>
<td>0.191</td>
<td>3.003</td>
<td>0.003</td>
</tr>
<tr>
<td>Prospector SO</td>
<td>0.175</td>
<td>0.059</td>
<td>Reactor SO</td>
<td>−0.036</td>
<td>0.040</td>
<td>0.211</td>
<td>2.948</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Furthermore, SIS was more strongly related to the prospector strategic orientation than the defender one. These results suggest that SIS produces business value through the promotion of environmental adaptation by pioneering new products or responding to emergent opportunities. Specifically, SIS allows firms to communicate objectives more effectively, scan the environment, promote organizational flexibility and innovate in a volatile environment. The study also found that an IS-incorporated business strategy promotes organizational flexibility, allowing firms to be creative and remain competitive in unpredictable business environments. Hence, effective use of SIS is believed to contribute to the building of capacity to reconfigure existing operational skills in order to respond to environmental changes better. Moreover, SIS can enable improvisation capabilities to reconfigure existing resources spontaneously in building new operational capabilities and to face urgent, unpredictable and new environmental situations during economic turbulence.

In addition, the results show no significant difference in how SIS affects adoption of the analyzer or defender strategic orientation. These results provide evidence for the notion that SIS pushes firms to protect their market share by improving the efficiency and productivity of current operations, while simultaneously monitoring the turbulent environment for new growth opportunities.

The tests for $H2$ indicate that a firm’s adoption of the analyzer, defender or prospector strategic orientation positively influences its corporate performance. By contrast, the results show that the adoption of the reactor strategic orientation is negatively associated with corporate performance. The results confirm those of other studies conducted in the USA (Moore, 2005; Parnell et al., 2015), which indicate no differences between the effects of the environment of a stable economy vs that of an economy facing a crisis.

The tests for $H3$ reveal significant path coefficients, indicating that SIS positively influences corporate performance. The research result is consistent with those of other studies carried out in countries with stable economies (Leidner et al., 2011) and in developing economies (Yayla and Hu, 2012). This finding demonstrates that IT/IS resources should be used to support and enable capabilities of business strategy (i.e. alignment) in order to drive firm performance. The study shows that SIS incorporated in the strategic planning process makes a greater contribution to corporate performance than does the strategic orientation adopted by the prospector, defender, analyzer or reactor company. Thus, it can be concluded that the enabling of strategy-as-practice is supported by SIS, which becomes more effective than the strategic firm posture, that is, practices involving the strategic planning process are more effective than strategy content is. Hence, this study shows that SIS enables firm capabilities to blend rational and top-down decision making, and promotes an integrative, communicative device and a key coordinating mechanism for strategic decisions.
The tests for H4a were supported by the moderation of firm size in the relationship between SIS and the prospector strategic orientation as well as that between the reactor strategic orientation and corporate performance. The research confirms the results of past studies (Anwar and Hasnu, 2016; Chan and Reich, 2007; Chan et al., 2006; Parnell et al., 2015) that size influences competitive strategy and corporate performance. Organization size – small, medium or large – is believed to influence dependent variables based on the availability of financial resources and maturity in the use of SIS. However, the moderation of sector did not support (H4b and H4c) the relationship between SIS and strategic orientation and between strategic orientation and corporate performance. These results suggested that a high concentration of services and manufacturing sectors (97 percent) did not allow verification of the moderation of control variables.

The post hoc analysis verified positive cause-and-effect relationships among the perspectives of an organization’s strategy. The analysis confirms Park et al.’s (2017) finding that companies should try to improve their performance in their learning and growth perspective in order to influence internal processes and impact customer satisfaction to improve FIPE. The results show that path coefficients of the reflective models were strong and significant: GLPE→IPPE (β = 0.607; p < 0.001; R² = 0.369), IPPE→CUPE (β = 0.655; p < 0.001; R² = 0.429) and CUPE→FIPE (β = 0.507; p < 0.001; R² = 0.257). Thus, a balanced scorecard is an appropriate model to measure corporate performance.

Taken together, the results show that effective use of SIS enables strategic orientation and creation of several benefits to influence firm performance. Furthermore, IS can contribute to the creation of business value to support all strategic planning processes. As such, the results indicate that the adoption of SIS solutions during economic crises can help firms perform well, even in volatile operating environments.

7. Implications and future research
This study contributes to Whittington’s (2014) theory and research questions, which focus on the necessity of SIS and strategy researchers to have a joint agenda. This is because the IS field has longstanding interest in the strategic role of SIS, while strategy researchers are increasingly recognizing the significance of material technology in strategy work.

This study’s results provide additional empirical support for the argument that IS use supports business strategy. The results suggest that both the effects of effective use of SIS on different types of competitive strategies and the strategic orientation influence corporate performance.

The effects of SIS on corporate performance vary depending on competitive strategy. Specifically, effective use of IS strategy enhances the prospector, defender and analyzer strategies, suggesting that these types of organizations should pay close attention to how they use IS in order to support their business strategies. The results also imply that SIS empowers the strategic planning process and enables strategy content.

Hence, the study confirms that, in practice, it is not sufficient simply to monitor an organization’s IT investment level. It is also necessary to understand and monitor how firms use technology to create business strategy value.

An important consideration for planning by practitioners is that not all firms use SIS in the same way to improve business strategy. It would appear that prospectors and analyzers have more to gain from the use of SIS, suggesting that they utilize SIS with greater effort to scan the environment and create new products and services through IT applications, such as customer relationship management, big data and customer analytics, rather than through operational support systems.

However, among defenders, the effects of SIS use are different to those among prospectors. These organizations use SIS with available data and information to make better decisions and improve operational efficiency (productivity and scale economies).
Meanwhile, this kind of firm focuses on operational efficiency by reneging on innovation initiatives; in the long run, it may have difficulty staying in the market.

Hence, these results suggest that it is important for planners to be aware that components of the strategic planning process are supported by SIS, because mechanisms used to attain effective SIS depend on the organization’s business strategy orientation. Assuming that each strategic orientation requires specific organizational capabilities to achieve superior performance, each company must align IT/IS resources that enable key processes to be effective in achieving strategic initiatives.

Another important observation is that effective use of SIS is more strongly related to exploration activities (prospector and analyzer) than to exploitation strategy (defender) in turbulent economies. In other words, the effective use of SIS in this scenario contributes to mitigating the risk of launching new products and services in a recessionary market by innovation strategy. Hence, this research confirms past studies (Chen et al., 2010; Leidner et al., 2011) that when SIS is effectively employed, it may create distinguish between firms’ performance and strategic business improvements attributable by use of SIS.

In agreement with past studies (Conant et al., 1990; Moore, 2005), the direct effects of Miles and Snow’s strategic types are equal to corporate performance by firms with defender, analyzer and prospector strategic orientations. Reactors showed inconsistent behavior, but other studies in developing countries (Parnell et al., 2012) demonstrate a significant negative effect on corporate performance. It is believed that reactor firms find it extremely difficult to survive in turbulent economies without a strategic planning process. In the context of this turbulent environment, SIS was demonstrated as a better alternative to support the strategic planning process and competitive strategy content.

Finally, this study’s results suggest that effective use of SIS ensures a stronger impact of the competitive prospector strategy and corporate performance in turbulent economic environments. The results show that it is important for firms to understand how SIS supports the strategic planning process and enables competitive strategy in periods of economic crises.

Future studies, for example, could investigate how SIS contributes to strategy-as-practice in the areas of praxis, practices and practitioners. According to Whittington (2014), firms do not have substantial empirical experience in applying various intimate methodologies, particularly ethnography, to business strategy from within.

Another potential focus of SIS researchers is to understand how digital technology impacts business strategy transformation. This would aid understanding of the importance of these technologies for the strategic planning process and content strategy.

Another topic to be investigated is the mediation of digital technology in the participation of strategy practitioners, that is, the level of collaboration of stakeholders in the strategic planning process.

Further study on how SIS could contribute to the strategic planning process should be conducted to understand the messy unfolding of practices involving strategic initiatives. Based on the research questions by Marabelli and Galliers (2017), strategists should be aware that strategizing is an emergent and emerging process and that it needs to be treated as such.

A final question to be resolved is how SIS can create dynamic and improvisation capabilities for firms to engage in exploration and exploitation innovation. Greater insight into these variables would provide further information on how SIS creates strategy business value that can affect firms’ capabilities for achieving superior performance in turbulent business environments.

Thus, this study contributes to the field of IS and strategy theory and presents many implications for practitioners and researchers in the field.

8. Study limitations

Although this study provides substantial insights into how SIS promotes business value through competitive strategy and corporate performance, it has limitations. First, as
mentioned earlier (Section 4), cross-sectional design does not allow a researcher to fully establish the causality between independent variables and dependent variables, and a carefully designed longitudinal study could address this question more successfully.

Second, the sample for the study was not perfectly random, because the difficulty of collecting data from Brazilian organizations precluded full randomization. Furthermore, data were collected for only two major sectors which hindered the ability to generalize across other sectors. However, this limitation is also acceptable, as it enabled the observation of variations within the two sectors, thereby, ruling out the effects of the agribusiness industry on SIS.

Note

1. FGV is active in the areas of information and research, both applied and academic, with more than 90 study centers. FGV was recognized as the top think tank in Latin America for seven consecutive years (FGV, 2015). The GVCia is a leading IT/IS applied research center in Brazil and has been publishing studies on IT/IS theory for more than 25 years (Meirelles, 2016).

References


Effects of SISs on competitive strategy


Further reading


Appendix. Abbreviated questionnaire

All items were presented in the form of seven-point Likert-type scales ranging from 1 (strongly agree) to 7 (strongly disagree).

- Strategic IS enables a firm to...
  - (SIS_1) disseminate its objectives to all levels.
  - (SIS_2) scan all external factors that affect it.
  - (SIS_3) formulate business strategies.
  - (SIS_4) implement strategies consistently with the firms’ business strategy in order to achieve goals.
  - (SIS_5) monitor the strategy and compare outcomes with other firms.

Effects of SISs on competitive strategy
Strategic orientation
Analyzer orientation
The firm develops a strategy to...

- (ANAL_1) adopt industry innovations only after lengthy consideration.
- (ANAL_2) focus first on serving current customers and second on capturing new customers.
- (ANAL_3) realize that present developments are indeed opportunities allowing for the assumption of necessary risks.

Defender orientation
The firm develops a strategy to...

- (DEFE_1) maintain a safe niche using a traditional store format.
- (DEFE_2) stick with use of the current store format.
- (DEFE_3) concentrate on improving current retailing methods rather than developing new methods.

Prospector orientation
The firm develops a strategy to...

- (PROS_1) be an innovation leader in the market.
- (PROS_2) move into new markets frequently.
- (PROS_3) be the first in the industry to develop new ways to market goods.

Reactor orientation
The firm develops a strategy to...

- (REAC_1) make unavoidable changes due to excessive pressure from the environment.
- (REAC_2) respond to environmental pressure by cutting costs.
- (REAC_3) enact fundamental changes when it faces negative events, such as a crisis.

Corporate performance by BSC
Financial performance

- (FIPE_1) The company reaches its goals of profitability to satisfy shareholders.
- (FIPE_2) The business is efficient in terms of spending (i.e. cost management, expenses, and investments) to meet productivity goals.
- (FIPE_3) The company reaches its goals with respect to revenues.

Customer performance

- (CUPE_1) Customers remain loyal to the company.
- (CUPE_2) The market associates the company’s image (brand) with the quality of the services and/or products it represents.
- (CUPE_3) Customers are satisfied with value provided by the company.

Internal process performance
The firm is efficient and effective in promoting...

- (IPPE_1) business process innovation.
• (IPPE_2) Business process operations.
• (IPPE_3) Business process post-sale activities.

Growth and learning performance
• (GLPE_1) Employees are satisfied with the firm’s human capital policies (attraction, retention and development).
• (GLPE_2) The firm is recognized by the market as a good place to work.
• (GLPE_3) Employees have the essential skills to manage their routines and strategic activities.

About the authors
Adilson Carlos Yoshikuni is Professor in the Educational Development Institute of Management at Fundação Getulio Vargas – FGV and Post-doctoral Researcher at the Center for Applied Information Technology (GVCIA) at FGV-EAESP. He earned PhD Degree in Business Administration from the School of Business Administration of São Paulo at FGV-EAESP. His research interests include the business value of information technology (IT), information systems (IS) for competing positioning, strategic information systems (SIS), digital business strategy and dynamic capacity enabled by IS. Adilson Carlos Yoshikuni is the corresponding author and can be contacted at: ayoshikuni@terra.com.br

Alberto Luiz Albertin is Full-time Professor at the School of Business Administration of São Paulo da Fundação Getulio Vargas (FGV-EAESP), Chief of the Department of Information Technology and Quantitative Methods Applied to Administration of FGV-EAESP and Coordinator of the Center for Applied Information Technology (GVCIA). He earned MSc and PhD Degrees in Business Administration from São Paulo University (FEA-USP). His research interests include issues in e-business and e-commerce, business value of IT, IS strategy and SIS.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Exploring the relationships of strategic entrepreneurship and social capital to sustainable supply chain management and organizational performance

Syed Awais Ahmad Tipu
Department of Management, College of Business Administration, University of Sharjah, Sharjah, United Arab Emirates, and
Kamel Fantasy
Department of Business and Administration, Faculty of Business and Economics, University of Winnipeg, Winnipeg, Canada

Abstract

Purpose – The purpose of this paper is to draw upon the resource-based view (RBV) of the firm in an attempt to explore how a firm’s resources (i.e. assets and capabilities) such as social capital (SC) and strategic entrepreneurship (SE) relate to sustainable supply chain management (SSCM) and organizational performance (OP).

Design/methodology/approach – Data were collected by questionnaire survey from the supply chain and logistics managers of 242 manufacturing firms in Pakistan. The structural equation modeling approach was used to test the hypotheses.

Findings – The results provide support for the proposed hypotheses. The results indicate that SC and SE are positively related to OP. However, the findings show a positive but weak association of SC and SE with SSCM. In a developing country context of Pakistan, organizations are more likely to employ SC and SE for achieving OP. However, relatively less emphasis is placed on linking SC and SE to SSCM. Pakistani organizations need to integrate SSCM into their business strategies. It is concluded that organizations in Pakistan though have some degree of involvement in SSCM but still face some challenges.

Originality/value – The current study attempts to narrow the gap in the available literature in three important aspects. First, it makes the contribution to the literature on SSCM by employing RBV and exploring the relationships of a firm’s resources (i.e. SC) and capabilities (i.e. SE) to SSCM and OP. Second, it employs a relatively more comprehensive measure of SE compared to the limited measures in existing empirical research. Third, the examination of the links of SE and SC to SSCM and OP is of particular importance in the context of a developing country such as Pakistan.

Keywords Performance, Sustainable supply chain management, Social capital, Resource-based view, Strategic entrepreneurship

Paper type Research paper

1. Introduction

Organizations need to employ strategic and entrepreneurial perspectives in order to survive in today’s competitive business environment (Hitt et al., 2001). The integration of strategy and entrepreneurship is called strategic entrepreneurship (SE) which underlines both opportunity seeking and advantage seeking behaviors (Ireland and Webb, 2007). This requires looking for future opportunities that could be exploited to deliver more value (Hitt et al., 2011). Today’s firms also need to develop mutual trust and relationships – social capital (SC) – with business partners, given the higher level of interconnectedness of their business operations.

It is pertinent to examine the role of SE and SC in the context of supply chain as supply chain partners increasingly strive to achieve sustainable competitive advantage at the inter-supply chain level. This is because the inter-company competition is elevated to inter-supply chain competition. Supply chain partners integrate and act as a single unified
entity in order to compete with competitors’ supply chains (Tan, 2001). In this context, sustainable supply chain management (SSCM) provides a possible source of competitive advantage to organizations (Gold et al., 2010). SSCM refers to the “management of material, information and capital flows as well as cooperation among companies along the supply chain while taking goals from all three dimensions of sustainable development, i.e., economic, environmental and social, into account which are derived from customer and stakeholder requirements” (Seuring and Muller, 2008, p. 1700).

The current paper draws upon the resource-based view (RBV) of the firm (Barney, 1991; Grant, 1991; Peteraf, 1993; Wernerfelt, 1984) in an attempt to explore the relationships of SE and SC to SSCM and organizational performance (OP). The RBV underlines the role of resources and capabilities in explaining persistent performance differences among firms (Alvarez and Barney, 2002). Resources are tangible and intangible assets of a firm (Barney and Arikan, 2001), whereas capabilities are the capacity of a group of resources to perform some task or activity (Grant, 1991). SC is an intangible asset which facilitates action through establishing relationships between individuals and organizations. It is one of the most important resources which need to be strategically managed (Sirmon and Hitt, 2003). Organizations need to strategically manage their resources in order to engage in SE (Ireland et al., 2003). Resources are bundled to create capabilities such as proactiveness, innovativeness and risk taking which are important dimensions of SE.

The potential influence of resources and capabilities on performance provides the necessary impetus for the current study to employ the lens of RBV in examining the relationships of SE and SC to SSCM and OP. This approach is in line with Barney’s (2012) assertion to consider RBV in the context of supply chain management. The RBV of the firm is pertinent as the study variables such as SE and SC have theoretical underpinnings in the RBV. Researchers have employed the RBV in examining SE (Ireland et al., 2003) and SSCM (Gold et al., 2010). In particular, Gavronski et al. (2011) draw upon RBV for proposing a theoretical model of development of green supply chain management capabilities. The current study extends the current debate by exploring the relationships of SE and SC to SSCM and OP.

Existing literature pays scant attention to linking SE to SSCM. The current study addresses this gap by exploring the relationship between SE and SSCM. The review of the literature did not reveal studies explicitly exploring the links between SE and SSCM. The notion of strategy has been brought forth in the SSCM literature but the particular linkage between SE and SSCM remains underexplored. Paulraj (2011) examines the relationship between enviropreneurship and sustainability performance. Enviropreneurship relates to entrepreneurial orientation (EO) that addresses the needs of environment and society and also satisfies the economic objectives of organizations (Menon and Menon, 1997). EO depicts the degree of innovation, risk propensity and proactiveness in organizations (Monsen and Boss, 2009). The notion of EO could be considered to represent SE (Monsen and Boss, 2009). However, the enviropreneurship scale employed by Paulraj (2011) contains only a few items which do not provide a full account of the three dimensions of EO including innovation, risk propensity and proactiveness. For example, only one statement measures risk propensity. Moreover, the enviropreneurship scale assesses the three facets of EO only in terms of environmentally friendly product development approaches in organizations.

The current study attempts to narrow the gap in the current literature in three important aspects. First, the current paper makes a contribution to the literature on SSCM by employing RBV and exploring the relationships of a firm’s resources (i.e. SC) and capabilities (i.e. SE) to SSCM and OP. Second, the current research employs relatively more comprehensive measures of SE compared to the very limited existing empirical research. This paper adopts the scale developed by Covin and Slevin (1990), which provides separate measure for proactiveness, risk taking and innovativeness using four items for each variable. Third, the examination of
the links of SE and SC to SSCM and OP is of particular importance in Pakistan, given that Pakistani industry faces serious domestic issues and constraints including power shortages, macroeconomic and political instability and bureaucratic corruption (Hussain et al., 2012). These factors make it challenging for the firm to gain competitive advantage and sustain higher level of performance.

The remainder of the paper is structured as follows. First, the existing literature on the possible relationships of SE and SC to SSCM and OP is discussed in an attempt to develop research hypotheses. A detailed methodology is presented including the measurement of variables. The results are discussed followed by managerial implications, conclusions and future research directions.

2. Literature review

2.1 Strategic entrepreneurship

SE integrates the concepts of strategy and entrepreneurship. According to Hitt et al. (2001, p. 480), SE is an “entrepreneurial action with a strategic perspective.” Entrepreneurship focuses more on opportunistic behavior in which organizations identify market opportunities for potential exploitation. However, strategic management emphasizes achieving sustainable competitive advantage (Hitt et al., 2011).

Most of the literature on SE is centered on the earlier work of Ireland et al. (2003). They have proposed a dynamic model of SE which represents various aspects of SE. They draw upon the RBV of the firm and integrate perspectives from organizational learning, creative cognition, human capital and SC. Given the fact that the SE field is in its infancy (Ireland, 2007), Monsen and Boss (2009) suggest for consideration other measures and constructs in order to empirically investigate different dimensions of the model of SE proposed by Ireland et al. (2003). Monsen and Boss (2009) refer to this model and argue that “applying creativity and developing innovation” – one of the boxes in the model – could be considered to present an important aspect of SE. They contend that innovativeness, risk taking and proactiveness are the key elements of this box that could be used to examine an important aspect of SE. They suggest for consideration the EO construct to underpin the integration of strategy and entrepreneurship; hence, the notion of SE, EO relates to strategic management and underlines the entrepreneurial actions of a firm with regard to innovativeness, risk taking and proactiveness (Ireland et al., 2001). Lumpkin and Dess (2001, p. 431) provide the following explanation for these dimensions:

Innovativeness refers to a willingness to support creativity and experimentation in introducing new products/services, and novelty, technological leadership and R&D in developing new processes. Risk taking means a tendency to take bold actions such as venturing into unknown new markets, committing a large portion of resources to ventures with uncertain outcomes, and/or borrowing heavily. Pro-activeness is an opportunity-seeking, forward-looking perspective involving introducing new products or services ahead of the competition and acting in anticipation of future demand to create change and shape the environment.

Covin and Slevin (1989, 1990) draw on the degree of innovativeness, risk-taking and proactiveness in order to define the notion of strategic posture. The refined form of strategic posture is called EO. Monsen and Boss (2009) use the EO measure developed by Covin and Slevin (1990) to assess SE. This measure of SE assesses the strategic posture of any organization in terms of risk taking, innovativeness and proactiveness. In line with the RBV, the current paper draws on the same conceptualization of SE and considers risk taking, innovativeness and proactiveness as capabilities.

2.2 Sustainable supply chain management (SSCM)

SSCM refers to the three aspects of sustainable development including social, environmental and economic (Seuring and Muller, 2008; Touboulic and Walker, 2015). Organizations are
increasingly under pressure to contribute toward societal welfare, lower the environmental impact and being profitable (McWilliams et al., 2016). Stakeholders play critical roles in influencing supply chain partners to engage in SSCM (Meixell and Luoma, 2015). In addition to enhancing the economic bottom line of an organization, SSCM also considers matters such as fair labor conditions and environmentally friendly production (Seuring, 2013). Some of the examples of SSCM initiatives include cost savings due to reduced packaging and design for reuse and recycling, improved working conditions and lower disposal costs (Carter and Easton, 2011).

SSCM has become increasingly important for all sizes of organizations in a variety of industries (Seuring, 2013). Supply chain partners need to increase coordination of their key inter-organizational processes to achieve SSCM (Seuring and Muller, 2008). In this context, Sarkis and Dhavale (2015) also suggest for consideration the environmental impacts and social responsibilities of the suppliers. Mutual consideration of economic, environmental and social aspects encourages transparency in the processes and collaboration with others in the value chain for achieving collaborative advantage (Glavas and Mish, 2015).

2.3 Strategic entrepreneurship and sustainable supply chain management

The available literature reports the relationships of different dimensions of SE (i.e. risk taking, proactiveness and innovativeness) to SSCM. The need to effectively manage risk makes the organizations more actively involved in identifying, mitigating and eliminating risk. This is very crucial in the context of a supply chain as supply chain partners undertake joint monitoring activities to manage risk (Carter and Rogers, 2008; Seuring and Muller, 2008). For effectively managing risk, organizations need to be more proactive (Pagell and Wu, 2009) which relates to the SE orientation. Bowen et al. (2001) find that the corporate environmental proactivity predicts the implementation of product-based green supply. Environmental proactivity describes the situation in which management goes beyond basic compliance with regulations and gives high priority to environmental issues. In the same vein, other researchers also report that the corporate environmental approach and management commitment to environmental issues contribute toward establishing green supply (Min and Galle, 1997; Noci, 1997).

The organizational capabilities such as stakeholder integration and manufacturing innovations are important antecedents of environmental management (Hart, 1995; Florida, 1996). The pursuit of innovation provides the necessary impetus for SSCM (Edgeman et al., 2015). Christmann (2000) contends that innovative organizations are more likely to lead the drive of environmental sustainability. Pagell and Wu (2009) also find that an organizational capability to innovate is an important antecedent to SSCM.

Lee (2008) argues that firms need to be more proactive in order to effectively undertake a green supply chain initiative. The literature provides support with regard to the positive relationship between firm’s resources and capabilities and the likelihood of being proactively involved in environmental management (Klassen and Whybark, 1999; Ghobadian et al., 1998; Hart, 1995).

Considering the above literature, we hypothesize as follows:

\[ H1. \] SE positively relates to SSCM.

2.4 Organizational performance

OP provides the basis for assessment of progress toward achieving predetermined objectives. This helps an organization to identify weak and strong areas and streamline its future initiatives (Fantazy et al., 2010). Both financial and non-financial measures are employed in the literature to assess OP (e.g. see Hoque and James, 2000). A balanced approach, as suggested by Maskell (1991), is more appropriate in which organizations
employ both financial and non-financial measures. This approach potentially enables the organizations to effectively compete in today’s competitive business environment since employing only traditional measures of OP is not adequate (Agami et al., 2012). Some of the dimensions of financial measures include accounting returns, growth and market returns (Combs et al., 2006). Non-financial measures refer to customer satisfaction in terms of level of customer perceived value of product (Fantazy et al., 2010).

2.5 Strategic entrepreneurship and organizational performance
Given the limited literature on SE, the current paper draws upon the notion of EO in order to highlight the relation between SE and OP. Monsen and Boss (2009) contend that EO could be considered to represent SE. The literature provides empirical support with regard to the positive influence of EO on performance (Fellnhofer et al., 2016; Zahra and Covin, 1995). Covin and Slevin (1991) report that considerable amount of evidence exists in the literature that confirms a positive relationship between strategic posture and firm performance. Different dimensions of EO – innovativeness, proactiveness, risk-taking propensity – could positively influence OP. Soininen et al. (2013) find that growth-oriented attitude positively relates to innovativeness and proactiveness. Zahra (1986) finds that degree of innovation, risk taking and futurity significantly predict the financial performance in terms of net income-to-sales ratio.

The meta-analysis conducted by Rauch et al. (2009) largely settles the debate on the EO-performance link. They find moderately large effects of EO on performance and report strongly positive correlation between EO and both financial and nonfinancial measures.

Thus, we propose the following hypothesis:

H2. SE positively relates to OP.

2.6 Social capital
The RBV theory assumes resource heterogeneity and resource immobility (Barney, 1991; Barney and Arikan, 2001; Priem and Butler, 2001). Resource heterogeneity means that firms control different bundles of resources and resource immobility relates to the possibility of persistent differences in the resource bundles owned by separate firms. This provides motivation to a firm to search external factor markets to acquire complementary resources. This relates to the SC of a firm which is defined as the “ability of actors to secure benefits by virtue of membership in social networks” (Portes, 1998, p. 6). Reciprocal relationships are crucial in developing SC (Lee and Jones, 2015). SC of an organization represents a total set of value-creating resources which an organization enjoys due to its intra- and inter-firm relationships (Koka and Prescott, 2002). The RBV of the firm considers SC as an intangible asset.

SC can be divided into three categories which include: cognitive capital, structural capital and relational capital (Nahapiet and Ghoshal, 1998). Relational capital represents relational embeddedness which underlines the relations of trust and reciprocity based on previous interaction (Lawson et al., 2008). Cognitive capital shows that the parties involved interact with each other having a shared meaning and common interpretations. This encourages them to develop shared values and goals (Krause et al., 2007). Structural capital involves structural configuration and boundary spanning roles of participants. It encourages managerial communication and technical exchanges between the participants (Lawson et al., 2008).

2.7 Social capital and sustainable supply chain management
Gold et al. (2010) take the relational perspective of the RBV and assert that the supply chain partners collaborate on environmental and social issues in order to achieve sustained inter-firm competitive advantage. It is crucial in the dynamic environment given the shift of competition from an inter-firm to inter-supply chain level. They further
contend that the supply chain-wide collaboration is critical in achieving sustainability in relation to the economic, environmental, and social dimensions of SSCM. Frosch (1994) argues that the proximity among supply chain partners facilitates inter-firm linkages which positively influence the environmental dimension of SSCM. In the same vein, Geffen and Rothenberg (2000) also highlight the role of suppliers’ relations in enhancing the quest for SSCM. Vachon and Klassen (2008) explore the influence of environmental collaborative activities on manufacturing performance. They find that a plant’s environmental performance increases as a result of more environmental collaboration. These findings highlight the possible relation between SC and SSCM. Hence, the current study proposes the following hypothesis:

H3. SC positively relates to SSCM.

2.8 Social capital and organizational performance

Previous research has suggested that the high level of SC enables an organization to perform better than its competitors (Nahapiet and Ghoshal, 1998). Bernardes (2010) finds relational embeddedness aspect of SC as a critical antecedent to performance. Tipu and Arain (2011) report that entrepreneurs may develop business relationships with suppliers for gaining favorable credit policies.

Researchers have employed both financial and nonfinancial measures to assess the SC–performance link. Stam and Elfring find a significant positive relationship between bridging ties and a venture’s financial and nonfinancial performance. Acquaah (2007) reports that the SC derived from networking and social relationships with managers in other firms, government officials, and community leadership enhances OP. In the same vein, Peng and Luo (2000) also confirm that the personal ties with managers at other firms positively affect a focal firm’s market share.

Krause et al. (2007) find a positive relationship between buying firms’ perceptions of shared values and goals with key suppliers and buyers’ performance improvement in terms of cost. Cooke and Wills (1999) reveal that SC improves business performance in terms of turnover, employment, market share or profit. Bosma et al. (2004) conclude that SC positively influences entrepreneurial performance in terms of employment generation and increase in survival time. The above discussion leads us to propose the following hypothesis:

H4. SC positively relates to OP.

2.9 Sustainable supply chain management and organizational performance

The existing literature provides support with regard to the relationship between SSCM and OP. Improved environmental performance of a firm provides various advantages such as improved revenues, new market opportunities and increased market share (Rao and Holt, 2005). Alvarez et al. (2001) argue that green supply chain management (GSCM) positively influences an organization’s economic performance. Rao and Holt’s (2005) study of companies in South East Asia also finds that the greening of different phases of supply chain leads toward competitiveness and economic performance. Markley and Davis (2007) contend that the sustainable supply chain positively relates to the profitability of a firm. Drawing upon the existing literature, the current paper proposes the following hypothesis:

H5. SSCM positively relates to OP.

3. Basic research model and hypotheses

The research model as shown in Figure 1 has its theoretical underpinnings in the RBV of the firm. As per the RBV, a firm’s resources include assets and capabilities. The model presents
SC as an intangible asset or resource and SE as a capability in terms of risk taking, innovativeness and proactiveness. Other studies also refer to SC as a resource (Ireland et al., 2003) and risk taking (Luo and Park, 2004), innovativeness (Benitez-Amado et al., 2010), and proactiveness (Lerner and Almor, 2002) as capabilities. The model also acknowledges the vital link of SC and SE to SSCM and OP. The relevant hypotheses are also shown in the model.

4. Methodology

4.1 Sample description

The current study employs a snowball sampling technique. Researchers face sampling difficulties due to infrastructural inadequacies and unavailability of reliable organizational data in developing countries such as Pakistan (Elahi, 2008; Nasif et al., 1991; Sekaran, 1983). Moreover, reference materials for the construction of sampling frames are often non-existent or incomplete (Ryan and Tipu, 2013). Consequently, the selection of subjects is often based on opportunistic availability (Tipu et al., 2012). Therefore, the current study employs the contextually appropriate convenience technique of snowball sampling and follows Harzing and 32 country collaborators (2005) in using confederates to support data collection. The research confederates provided the respondents with the survey instrument in English. The sample comprises 242 manufacturing firms in the cities of Lahore, Sialkot and Islamabad in Pakistan with supply chain and logistics managers as the key informants. The cities included in the current research had a wide range of industries with manufacturing activity. Industry segmentation includes several industry types with the largest proportion of respondents working in the computer and electronics sector (computer and electronics 28.51 percent, textile 15.70 percent, furniture 13.64 percent, medical and pharmaceutical 12.80 percent, sport and leisure 11.15 percent, metal and automotive 10.33 percent, others 7.85 percent). The sample demographics summary can be seen in Table I.

<table>
<thead>
<tr>
<th>Industry category</th>
<th>No</th>
<th>%</th>
<th>Company size</th>
<th>No</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer and electronics</td>
<td>69</td>
<td>28.51</td>
<td>Micro (0–9 employees)</td>
<td>5</td>
<td>2.07</td>
</tr>
<tr>
<td>Textile</td>
<td>38</td>
<td>15.70</td>
<td>Small (10–99 employees)</td>
<td>90</td>
<td>37.19</td>
</tr>
<tr>
<td>Furniture</td>
<td>33</td>
<td>13.64</td>
<td>Medium (100–250 employees)</td>
<td>105</td>
<td>43.39</td>
</tr>
<tr>
<td>Medical and pharmaceutical</td>
<td>31</td>
<td>12.81</td>
<td>Large (30) (251–999 employees)</td>
<td>30</td>
<td>12.39</td>
</tr>
<tr>
<td>Sport and leisure</td>
<td>27</td>
<td>11.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metal and automotive</td>
<td>25</td>
<td>10.33</td>
<td>Very large (1,000+ employees)</td>
<td>12</td>
<td>4.96</td>
</tr>
<tr>
<td>Others</td>
<td>19</td>
<td>7.85</td>
<td>Total</td>
<td>242</td>
<td>100</td>
</tr>
</tbody>
</table>

Table I.
Sample demographics summary
4.2 Questionnaire design
Drawing upon the literature review, the questionnaire is divided into five sections (Sections 2–5 are shown in questionnaire provided in Table AI):

- **Section 1** contains the information on the profiles of the firms, such as firm name, address, the respondent’s position within the company and the type of manufacturing industry. Also collected is the number of employees in the organization, the type of products produced, the approximate turnover and the number of years the firm has implemented a supply chain program.

- **Section 2** includes questions on dimensions of SE in terms of four items of risk taking, 4 items of innovativeness and 4 items of pro-activeness.

- **Section 3** includes questions on dimensions of SC in terms of four items of cognitive capital, eight items of structural capital and four items of relational capital.

- **Section 4** includes questions on dimensions of SSCM in terms of five economic performance, six items of environmental performance and three items of social performance.

- **Section 5** includes questions on dimensions of OP in terms of two items of financial performance, and four items of none financial performance.

4.3 Measurement of variables
To test the above hypotheses, we used multi-item scales adopted from prior studies for the measurement of constructs. A description of the constructs is included in Table AI. An important question for social researchers is to apply a reflective or a formative perspective for measures. It cannot be taken for granted that researchers will always make the correct choice when operationalizing constructs in organizational research efforts. Depending on this choice, researchers should use multi-item measures generated by scale development (reflective) and index construction (formative) approaches (Diamantopoulos and Siguaw, 2006). All the four latent constructs of this study are shown in the research model in Figure 1 and are assumed to be reflective in this study.

4.3.1 **Strategic entrepreneurship.** In order to measure SE, the 12 items scale – adapted from Covin and Slevin (1990) – is used in the current study. Using a semantic differential type of scale, the respondents are asked to select the number from 1 to 7 in order to rate the importance of the potential of SE in their organizations in terms of (innovativeness, pro-activeness and risk taking). A confirmatory factory analysis (CFA) was conducted to validate the scales that fits the data which resulted in the deletion of two items: (SE4) and (SE8). This procedure allowed us to select ten items with high validity loading and consistency reliability Cronbach (α = 0.913). Statistical details are presented in Table III.

4.3.2 **Social capital.** In order to measure SC, 12 items are drawn from Lawson et al. (2008) and 4 items from Krause et al. (2007). The respondents are asked to recall their business relationship with one of the supply chain partners and rate the importance of the potential of SC in terms of (cognitive, relational, and structural capital) using a Likert scale of seven points from 1 = strongly disagree to 7 = strongly agree. A CFA was conducted to validate our scales which required deletion of four items SC6, SC14, SC12 and SC16. This technique allowed us to select twelve items with high loading and decent consistency reliability (Cronbach’s α = 0.903). Statistical details are presented in Table III.

4.3.3 **Sustainable supply chain management.** Drawing upon the earlier work of Paulraj (2011), the current study employs a 14-items scale to measure SSCM. The respondents are asked to rate if their firm and supply chain partners work closely to achieve SSCM using a Likert scale of seven points from 1 = strongly disagree to 7 = strongly agree. A CFA was
conducted to validate our scales which required deletion of two items SSCM4 and SSCM13. This process allowed us to select twelve items with high loading and decent consistency reliability (Cronbach’s $\alpha = 0.791$). Statistical details are presented in Table III.

4.3.4 Organizational performance. In order to measure OP, five items are drawn from Fantazy et al. (2010) and one item from Paulraj et al. (2012). The respondents are asked to evaluate their performance and compare their firm’s performance with their major competitors using a Likert scale of seven points from 1 = very weak to 7 = very strong. Sample performance measures include “average net profit” and “order lead-time.” A CFA was conducted to validate our scales that fits our data which resulted in deletion of no items. This procedure allowed us to select all the six items with high validity loading and consistency reliability (Cronbach’s $\alpha = 0.808$). Statistical details are presented in Table IV.

5. Data analysis and results
In total, 48 items in this research were designed to measure four main constructs: OP (dependent variable), SC, SE and SSCM (predictors). The value of KMO (Kaiser–Meyer–Olkin) measure of sampling adequacy (0.897) demonstrates that the sample is appropriate to run factor analysis. Table II displays the Kaiser–Meyer–Olkin and Bartlett’s test results.

The data were analyzed in two phases. First, the exploratory factor analysis (EFA) (principal component analysis with varimax rotation) was performed on all the multiple scale items to determine the item retention and uni-dimensionality of the constructs. The EFA revealed that two of the SE items, four of the SC items and two of SSCM items did not meet our criteria. In the EFA approach, a one-factor solution was selected for SC, SE, SSCM and OP. SPSS software version 19 was used to perform all the necessary analysis. See Table III.

To test the research model, LISREL 8.75 was used as the statistical package. A two-stage approach was used: first, CFA was used to confirm the factor structure extracted in the EFA. Second, the full model was tested to generate the path coefficients.

To examine the psychometric properties of the measurement model, Cronbach’s $\alpha$, convergent validity and discriminant validity were assessed for the remaining 40 indicators. The reliability measures for this model are provided in Tables II and III along with uni-dimensionality measures. An $\alpha$ value above 0.70 is deemed sufficient for confirmatory research (Gefen et al., 2000). Cronbach’s coefficient $\alpha$ was computed to assess the reliability of the items used to measure the research model variables. Table III shows that the Cronbach’s $\alpha$ values obtained in this study ranged between 0.791 and 0.913, suggesting that the measurement used in this study is reliable (Hair et al., 1992). All constructs had values above 0.70 meaning that the items measured the construct reliably (Gliem and Gliem, 2003). According to Germain et al.’s (1994) suggestion, principal component analyses were used to test unidimensionality. Therefore, each construct should be independently subject to principal components analysis to determine the eigenvalues. As a rule, eigenvalues that are greater than 1 while the second and subsequent factors should have eigenvalues less than 1 that provide support for the unidimensionality of these scales (Hoe, 2008). However, two of our constructs (SC and SSCM) do not comply with this rule. Since the subsequent factors are only slightly above 1 and therefore can be ignored (Hoe, 2008). The eigenvalues for each factor are presented in Table III.

KMO and Bartlett’s Test

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaiser–Meyer–Olkin measure of sampling adequacy</td>
<td>0.897</td>
</tr>
<tr>
<td>Bartlett’s test of sphericity Approx. $\chi^2$</td>
<td>5,901.969</td>
</tr>
<tr>
<td>df</td>
<td>1,128</td>
</tr>
<tr>
<td>Sig.</td>
<td>0.000</td>
</tr>
</tbody>
</table>

| Table II. KMO and Bartlett’s test results |
### Construct 1: Strategic Entrepreneurship (SE) (α = 0.913)

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor loading</th>
<th>t-value</th>
<th>Component</th>
<th>Total</th>
<th>% of variance</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE1</td>
<td>0.816</td>
<td>9.47</td>
<td>1</td>
<td>6.326</td>
<td>52.714</td>
<td>52.714</td>
</tr>
<tr>
<td>SE1</td>
<td>0.790</td>
<td>8.45</td>
<td>2</td>
<td>0.945</td>
<td>7.873</td>
<td>60.587</td>
</tr>
<tr>
<td>SE3</td>
<td>0.782</td>
<td>4.36</td>
<td>3</td>
<td>0.750</td>
<td>6.249</td>
<td>66.836</td>
</tr>
<tr>
<td>SE12</td>
<td>0.762</td>
<td>9.60</td>
<td>4</td>
<td>0.636</td>
<td>5.304</td>
<td>72.140</td>
</tr>
<tr>
<td>SE6</td>
<td>0.742</td>
<td>9.44</td>
<td>5</td>
<td>0.570</td>
<td>4.748</td>
<td>76.887</td>
</tr>
<tr>
<td>SE9</td>
<td>0.733</td>
<td>10.45</td>
<td>6</td>
<td>0.544</td>
<td>4.533</td>
<td>81.420</td>
</tr>
<tr>
<td>SE11</td>
<td>0.707</td>
<td>7.08</td>
<td>7</td>
<td>0.476</td>
<td>3.971</td>
<td>85.391</td>
</tr>
<tr>
<td>SE2</td>
<td>0.706</td>
<td>5.76</td>
<td>8</td>
<td>0.429</td>
<td>3.575</td>
<td>88.966</td>
</tr>
<tr>
<td>SE10</td>
<td>0.669</td>
<td>9.32</td>
<td>9</td>
<td>0.376</td>
<td>3.134</td>
<td>92.100</td>
</tr>
<tr>
<td>SE7</td>
<td>0.580</td>
<td>12.35</td>
<td>10</td>
<td>0.349</td>
<td>2.909</td>
<td>95.000</td>
</tr>
<tr>
<td>SE8</td>
<td>0.410</td>
<td>–</td>
<td>11</td>
<td>0.331</td>
<td>2.756</td>
<td>97.764</td>
</tr>
<tr>
<td>SE4</td>
<td>0.436</td>
<td>–</td>
<td>12</td>
<td>0.286</td>
<td>2.236</td>
<td>100.000</td>
</tr>
</tbody>
</table>

### Construct 2: Social Capital (SC) (α = 0.903)

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor loading</th>
<th>t-value</th>
<th>Component</th>
<th>Total</th>
<th>% of variance</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC1</td>
<td>0.811</td>
<td>6.09</td>
<td>1</td>
<td>6.594</td>
<td>41.215</td>
<td>41.215</td>
</tr>
<tr>
<td>SC13</td>
<td>0.811</td>
<td>8.23</td>
<td>2</td>
<td>1.141</td>
<td>7.131</td>
<td>48.346</td>
</tr>
<tr>
<td>SC15</td>
<td>0.751</td>
<td>10.44</td>
<td>3</td>
<td>1.053</td>
<td>6.578</td>
<td>54.924</td>
</tr>
<tr>
<td>SC11</td>
<td>0.690</td>
<td>12.45</td>
<td>4</td>
<td>0.862</td>
<td>5.391</td>
<td>60.315</td>
</tr>
<tr>
<td>SC2</td>
<td>0.687</td>
<td>9.04</td>
<td>5</td>
<td>0.768</td>
<td>4.799</td>
<td>65.114</td>
</tr>
<tr>
<td>SC10</td>
<td>0.623</td>
<td>5.55</td>
<td>6</td>
<td>0.751</td>
<td>4.684</td>
<td>69.808</td>
</tr>
<tr>
<td>SC4</td>
<td>0.577</td>
<td>6.55</td>
<td>7</td>
<td>0.688</td>
<td>4.174</td>
<td>73.982</td>
</tr>
<tr>
<td>SC9</td>
<td>0.607</td>
<td>6.05</td>
<td>8</td>
<td>0.623</td>
<td>3.866</td>
<td>77.867</td>
</tr>
<tr>
<td>SC3</td>
<td>0.550</td>
<td>3.67</td>
<td>9</td>
<td>0.599</td>
<td>3.742</td>
<td>81.620</td>
</tr>
<tr>
<td>SC8</td>
<td>0.521</td>
<td>15.23</td>
<td>10</td>
<td>0.513</td>
<td>3.204</td>
<td>84.824</td>
</tr>
<tr>
<td>SC5</td>
<td>0.514</td>
<td>8.08</td>
<td>11</td>
<td>0.492</td>
<td>3.078</td>
<td>87.902</td>
</tr>
<tr>
<td>SC7</td>
<td>0.506</td>
<td>4.67</td>
<td>12</td>
<td>0.468</td>
<td>2.926</td>
<td>90.826</td>
</tr>
<tr>
<td>SC14</td>
<td>0.446</td>
<td>–</td>
<td>13</td>
<td>0.423</td>
<td>2.645</td>
<td>93.474</td>
</tr>
<tr>
<td>SC6</td>
<td>0.410</td>
<td>–</td>
<td>14</td>
<td>0.378</td>
<td>2.359</td>
<td>95.833</td>
</tr>
<tr>
<td>SC16</td>
<td>0.410</td>
<td>–</td>
<td>15</td>
<td>0.344</td>
<td>2.150</td>
<td>97.983</td>
</tr>
<tr>
<td>SC12</td>
<td>0.407</td>
<td>–</td>
<td>16</td>
<td>0.323</td>
<td>2.017</td>
<td>100.000</td>
</tr>
</tbody>
</table>

### Construct 3: Sustainable Supply Chain Management (SSCM) (α = 0.791)

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor loading</th>
<th>t-value</th>
<th>Component</th>
<th>Total</th>
<th>% of variance</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSCM12</td>
<td>0.843</td>
<td>9.02</td>
<td>1</td>
<td>5.502</td>
<td>39.300</td>
<td>39.300</td>
</tr>
<tr>
<td>SSCM1</td>
<td>0.779</td>
<td>7.53</td>
<td>2</td>
<td>1.170</td>
<td>8.380</td>
<td>47.660</td>
</tr>
<tr>
<td>SSCM10</td>
<td>0.778</td>
<td>12.45</td>
<td>3</td>
<td>0.960</td>
<td>6.855</td>
<td>54.514</td>
</tr>
<tr>
<td>SSCM2</td>
<td>0.774</td>
<td>6.89</td>
<td>4</td>
<td>0.874</td>
<td>6.242</td>
<td>60.756</td>
</tr>
<tr>
<td>SSCM6</td>
<td>0.681</td>
<td>3.88</td>
<td>5</td>
<td>0.779</td>
<td>5.566</td>
<td>66.322</td>
</tr>
<tr>
<td>SSCM9</td>
<td>0.670</td>
<td>7.90</td>
<td>6</td>
<td>0.693</td>
<td>4.953</td>
<td>71.276</td>
</tr>
<tr>
<td>SSCM10</td>
<td>0.649</td>
<td>5.67</td>
<td>7</td>
<td>0.678</td>
<td>4.843</td>
<td>76.119</td>
</tr>
<tr>
<td>SSCM7</td>
<td>0.640</td>
<td>9.00</td>
<td>8</td>
<td>0.694</td>
<td>4.745</td>
<td>80.863</td>
</tr>
<tr>
<td>SSCM5</td>
<td>0.562</td>
<td>3.56</td>
<td>9</td>
<td>0.581</td>
<td>4.150</td>
<td>85.013</td>
</tr>
<tr>
<td>SSCM14</td>
<td>0.527</td>
<td>6.66</td>
<td>10</td>
<td>0.553</td>
<td>3.819</td>
<td>88.852</td>
</tr>
<tr>
<td>SSCM11</td>
<td>0.517</td>
<td>8.65</td>
<td>11</td>
<td>0.488</td>
<td>3.344</td>
<td>92.176</td>
</tr>
<tr>
<td>SSCM10</td>
<td>0.510</td>
<td>8.48</td>
<td>12</td>
<td>0.414</td>
<td>2.958</td>
<td>95.134</td>
</tr>
<tr>
<td>SSCM4</td>
<td>0.468</td>
<td>–</td>
<td>13</td>
<td>0.386</td>
<td>2.759</td>
<td>97.893</td>
</tr>
<tr>
<td>SSCM13</td>
<td>0.452</td>
<td>–</td>
<td>14</td>
<td>0.295</td>
<td>2.167</td>
<td>100.000</td>
</tr>
</tbody>
</table>

### Construct 4: Organizational Performance (OP) (α = 0.808)

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor loading</th>
<th>t-value</th>
<th>Component</th>
<th>Total</th>
<th>% of variance</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP1</td>
<td>0.886</td>
<td>9.89</td>
<td>1</td>
<td>3.368</td>
<td>56.131</td>
<td>56.131</td>
</tr>
<tr>
<td>OP6</td>
<td>0.880</td>
<td>4.44</td>
<td>2</td>
<td>0.849</td>
<td>14.146</td>
<td>70.277</td>
</tr>
<tr>
<td>OP2</td>
<td>0.803</td>
<td>5.54</td>
<td>3</td>
<td>0.534</td>
<td>8.903</td>
<td>79.180</td>
</tr>
<tr>
<td>OP5</td>
<td>0.731</td>
<td>6.32</td>
<td>4</td>
<td>0.514</td>
<td>8.571</td>
<td>87.751</td>
</tr>
<tr>
<td>SC1</td>
<td>0.700</td>
<td>6.78</td>
<td>5</td>
<td>0.415</td>
<td>6.919</td>
<td>94.670</td>
</tr>
<tr>
<td>OP4</td>
<td>0.559</td>
<td>9.76</td>
<td>6</td>
<td>0.320</td>
<td>5.331</td>
<td>100.000</td>
</tr>
</tbody>
</table>

### Table III.

Factor loading, standard error, t-value and internal consistency (α)

**Sources:** SPSS 19, extraction method: principal component analysis; rotation method: Varimax with Kaiser normalization
Convergent validity is suggested if the factor loadings are 0.50 or higher (Bagozzi and Yi, 1988). We found that the lowest loading was 0.51 which was above the required upper limit. This finding provides strong support for the convergent validity of the scale. Next, to check the reliabilities of latent variables, composite reliability (CR) and the average variance extracted (AVE) were calculated using the procedure outlined by Fornell and Larcker (1981).

As shown in Table IV, CRs and AVEs of the scales are above 0.72 and 0.5, respectively, except for the SSCM variable at 0.49, which is nevertheless very close to the ceiling. These results suggest satisfactory convergent validity of the measurement model.

5.1 Structural research model
The fitted structure equation model (SEM) obtained from LISREL software accepted for the study is illustrated in Figure 2, with a structural model determining the significance of the

---

**Table IV.**

<table>
<thead>
<tr>
<th></th>
<th>SE</th>
<th>SC</th>
<th>SSCM</th>
<th>OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>0.73</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSCM</td>
<td>0.69</td>
<td>0.75</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>OP</td>
<td>0.57</td>
<td>0.69</td>
<td>0.79</td>
<td>1.00</td>
</tr>
<tr>
<td>AVE</td>
<td>0.84</td>
<td>0.75</td>
<td>0.49</td>
<td>0.69</td>
</tr>
<tr>
<td>√AVE</td>
<td>0.91</td>
<td>0.86</td>
<td>0.70</td>
<td>0.83</td>
</tr>
<tr>
<td>Composite reliability</td>
<td>0.86</td>
<td>0.90</td>
<td>0.88</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Notes: $\chi^2 = 1,524.33$; df = 734; p-value = 0.2133; RMSEA = 0.035
relationships between the independent and dependent variables. Many researchers have chosen SEM models since the models provide a statistical approach for explicitly considering measurement error in the observed variables including both dependent and independent variables in a given model (Kline, 2004). SEM is a multivariate analysis methodology for empirically examining sets of relationships represented in the form of linear causal models (Joreskog and Sorbom, 2001).

5.2 Model identification
Before analyzing the structural model, it is desirable to check the model identification to obtain the correct estimate of the parameter values. The SEM is over-identified with 40 observed variables – there are \(\frac{40 \times 41}{2} \approx 820\) observations. The number of parameters to be estimated is 86, including the variances of 40 variables (22 exogenous and 18 indigenous variables that are the disturbance), with 5 direct loading on each latent variable. Furthermore, one error co-variance was set to free between the exogenous variables. Thus, the model degrees of freedom are \(820 - 85 - 1 = 734\) (see Figure 2, df = 734). Since the number of observations is greater than the number of parameters to be estimated, we conclude that the research model is over-identified and can be tested statistically.

5.3 Model goodness fit
The literature mentions a number of goodness of fit statistics to check the fitness of the model with the research data. The three most commonly used indices are RMSEA, CFI and NFI. Another goodness of fit statistic is \(\chi^2\), which was used in many studies but has severe limitations because it is affected by the sample size; when the sample size goes beyond 200 cases, it usually gives a significant result. The fitted SEM obtained from LISREL software accepted for the research model is illustrated in Figure 2. The research model presented in Figure 2 shows a good fit of SE, SC, SSCM and OP to the empirical data. The observed \(\chi^2\) was \(\chi^2 = 1,524.33\), degree of freedom df = 734 and RMSEA = 0.035. According to Table V, the fit values of indexes such as the RMSEA, CFI and GFI show that the full model fits the sample data.

5.4 Regression analyses results
Based on the results of the regression analyses presented in Figure 2 and summarized in Table VI, all of the hypothesized relationships were supported. The results indicate that SC and SE are positively related to SSCM and the OP. All of the correlations were significant at different levels and were in the predicted direction. The hypotheses for the relationships were tested using their associated \(t\)-statistics. \(t\)-values greater than 1.65, 1.98 or 2.576 were considered to be significant at the 0.10, 0.05 and 0.01 levels, respectively (Hair et al., 2006).

In case of \(H1\), the path coefficient for SE and SSCM was 0.16 and statistically significant at the 10 percent level. Although SSCM coefficient was small but it is a positive predictor of SE. For \(H2\), the path coefficient for SE and OP was 0.42 and statistically significant at the 5 percent level. OP was a significant predictor of SE. For hypothesis \(H3\), the path coefficient

<table>
<thead>
<tr>
<th>Goodness-of-fit measure</th>
<th>Recommended value</th>
<th>Model fit index</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\chi^2) (degree of freedom)</td>
<td>(\leq 3)</td>
<td>2.076</td>
</tr>
<tr>
<td>Goodness-of-fit index (GFI)</td>
<td>(\geq 0.90)</td>
<td>0.95</td>
</tr>
<tr>
<td>Adjusted goodness-of-fit index (AGFI)</td>
<td>(\geq 0.80)</td>
<td>0.88</td>
</tr>
<tr>
<td>Tucker–Lewis index (TLI)</td>
<td>(\geq 0.90)</td>
<td>0.92</td>
</tr>
<tr>
<td>Comparative fit index (CFI)</td>
<td>(\geq 0.90)</td>
<td>0.98</td>
</tr>
<tr>
<td>Root mean square error approximation (RMSEA)</td>
<td>(\leq 0.08)</td>
<td>0.035</td>
</tr>
</tbody>
</table>

Table V. Goodness-of-fit measures of the research model
for SC and SSCM was 0.08 and statistically significant at the 10 percent level. In case of H4, the path coefficient for SC and OP was 0.36 and statistically significant at the 5 percent level. OP was a significant predictor of SC. Finally, for H5, the path coefficient for SSCM and OP was 0.20 and statistically significant at the 10 percent level.

Clearly, in case of H1 and H3 which dealt with the direct relationship between SC, SE and SSCM indicate a positive but weak association of SC and SE with SSCM. Empirically, this finding is partially supported by several studies that consistently found a positive impact of SC on SSCM. For example, Vachon and Klassen (2008) find that a plant’s environmental performance increases as a result of more environmental collaboration. Similarly, Vachon and Klassen (2008) also find that the environmental collaboration and monitoring is positively linked to the technological integration among primary suppliers and major customers. The findings regarding the positive association between SE and SSCM were also supported by previous empirical research. For example, Bowen et al. (2001) find that the corporate environmental proactivity predicts the implementation of product-based green supply. Moreover, H5 indicating a moderate relationship between SSCM and OP reveals deficiency in SSCM execution. This finding is not fully consistent with the existing literature which shows that GSCM (Alvarez et al., 2001) or sustainable supply chain (Markley and Davis, 2007) positively relates to profitability.

H2 and H4, which explore the direct relationship between SC and SE and OP, reveal that OP has a positive correlation with both SC and SE. Empirically, this finding is supported by several studies that consistently report a positive impact of SC on OP. For example, Cooke and Wills (1999) reveal that SC improves OP in terms of turnover, employment, market share, or profit. Bernardes (2010) reports the influence of SC or embeddedness on OP. The findings regarding the positive correlation between SE and OP are also supported by previous empirical findings. For instance, Zahra (1986) finds that a firm’s emphasis on corporate entrepreneurship in terms of innovation, risk taking and futurity predicts net income-to-sales ratio. Rauch et al. (2009) also contend that EO influences OP.

6. Discussion
It is pertinent to employ the RBV for understanding how organizations could employ their resources and capabilities to achieve sustained performance. Changes in the competitive environment can either increase or decrease the commercial value of the firm’s resources (Miller and Shamsie, 1996). This is relevant in the context of a developing country such as Pakistan in which organizations operate in a very challenging economic and sociocultural environment. Under these circumstances, the resource portfolio of the firm requires managerial actions to strategically bundle resources in order to form capabilities for achieving higher level of performance (Ireland et al., 2003).

A positive but weak association of SC and SE with SSCM shows that there is room to better leverage the SC and SE to achieve sustained competitive advantage at the inter-supply chain level in Pakistan. The supply chain partners build relationships based on trust and reciprocity (Lawson et al., 2008) in order to enhance OP. However, there is a

<table>
<thead>
<tr>
<th>Direct effect</th>
<th>H1 SE-SSCM</th>
<th>H2 SE-OP</th>
<th>H3 SC-SSCM</th>
<th>H4 SC-OP</th>
<th>H5 SSCM-OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path coefficient</td>
<td>0.16</td>
<td>0.42</td>
<td>0.08</td>
<td>0.36</td>
<td>0.20</td>
</tr>
<tr>
<td>SE (t-statistics)</td>
<td>(0.085)</td>
<td>(0.19)</td>
<td>(0.045)</td>
<td>(0.16)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Notes: N = 242. ✓Accepted. *,<strong>,</strong> Significant at 10, 5 and 1 percent levels, respectively</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
possibility to better employ the resources and capabilities to achieve SSCM which could be a source of sustainable competitive advantage (Barney, 1991). Inter-organizational resources and capabilities are difficult to replicate due to their idiosyncrasies (Gold et al., 2010). Supply chain partners are likely to have sustained performance at the inter-supply chain level due to the sharing of complementary resources and capabilities. This helps in achieving inter-organizational competitive advantage (Dyer and Singh, 1998).

The findings of the current study reveal that in Pakistani organizations, SSCM has a moderate relationship with OP which shows the need to improve SSCM for enhancing the financial and non-financial OP. This implies that they struggle to prioritize sustainability performance over economic performance. Businesses need to pursue social, economic and environmental sustainability goals along with their financial bottom line (Walker and Jones, 2012).

The findings of the current study emphasize the need for more robust implementation of SSCM initiatives in Pakistan. The contextual setting in a developing country provides the explanation for this finding. Due to the lack of policy and institutional provisions for sustainable initiatives in a developing country, there is minimal participation from the formal sector of the economy (Rathore et al., 2011). In addition, another reason that industries give low priority to the environment is the low enforcement of a legal framework to support environmental protection (Shukla et al., 2009). The lack of willingness to adopt, financial constraints, lack of technology and lack of awareness are some of the other barriers to GSCM (Shukla et al., 2009; Ravi and Shankar, 2005; Zhu et al., 2005). In addition, Luken and Stares (2005) report that SMEs in developing countries do not know how to respond to the rising social and environmental requirements of global buyers. Moreover, in developing countries, individuals and institutional buyers are also very cost conscious. This makes them unwilling to pay extra for eco-friendly products (Green et al., 1998). This highlights that businesses operating in developing countries do not get much support from the society for undertaking SSCM initiatives. SSCM is a relatively new concept in Pakistan, given that organizations take plant-level initiatives. This underlines the need to develop more integrated system for the whole supply chain.

The current study reports high relationships (regression coefficients) of SC and SE with OP. This underpins that in Pakistan, organizations’ quest to develop SC and strategic orientation is more geared toward achieving a higher level of OP. One possible explanation of this orientation is that firms in developing countries face a challenging business environment due to power shortages, macroeconomic and political instability and bureaucratic corruption (Hussain et al., 2012). Under these adverse circumstances, organizations strive for their survival which results in focusing more on OP. However, in line with the RBV, organizations need to be more proactive, undertaking a higher level of risk and introducing innovations in order to remain competitive and survive in the marketplace. Moreover, strong relationships with suppliers enable an organization to achieve higher level of performance (Krause et al., 2007). Organizations develop trust and increasingly rely on business networks in order to have a competitive edge on the competitors. In a developing country context, the current study reports that organizations employ SC and SE for achieving OP instead of achieving a higher level of SSCM.

The current study reports a positive and significant association of SE with OP. However, SE is not strongly related to SSCM which shows that the risk taking, innovativeness and proactiveness dimensions of SE in Pakistani supply chains do not heavily focus on enhancing the environmental, economic, and social aspects of SSCM. This warrants that organizations think differently with regard to business models, product technologies and processes. This underlines the need that SSCM should be the strategic goal of organizations striving to achieve competitive advantage derived from a sustainable supply chain. Isaksson et al. (2010) also contend that supply chains have significant potential to innovate
for sustainable development. However, supply chains in Pakistan could achieve this through system thinking and focusing on technological as well as organizational innovations. Thus, the innovative quest of Pakistani supply chains needs to be more clearly aligned to SSCM involving all supply chain partners.

All in all, the current study employs the RBV to explore how organizations operating in a weak institutional regime deploy their resources to achieve SSCM and OP.

7. Conclusions

The current study makes a contribution to the literature on SSCM by empirically examining the relationships of SE and SC to SSCM and OP. Future researchers may employ the proposed research model because its theoretical relationships are expected to apply in developed and developing countries. However, the levels of individual constructs may vary in different settings.

The current study offers some practical implications for supply chain executives. Firms may fail to adapt if the external environment changes quickly (Harreld et al., 2007). This is crucial as firms in a developing country such as Pakistan face serious domestic issues and constraints including power shortages, macroeconomic and political instability and bureaucratic corruption (Hussain et al., 2012). This underlines the need to be more entrepreneurial in adapting, integrating, and reconfiguring internal and external organizational resources and functional capabilities (Teece and Pisano, 1994). In this context, developing entrepreneurial skills rather than managerial skills are more important in order to develop expertise to innovate (Penrose, 1959).

Another practical implication relates to the role of the top management. Managers could restrain the growth of the firm due to their limited knowledge about the firm resource base and their understanding of the external environment (Penrose, 1959). Managers need to accurately sense the changes in the environment and seize the opportunities by reconfiguring both tangible and intangible assets (Harreld et al., 2007). This highlights the advantage seeking and opportunity seeking behaviors (Ireland and Webb, 2007) enacted through the integration of a firm’s resources and capabilities.

Given the positive but weak association (regression coefficients) of SE and SC with SSCM, there is a need for aligning strategic initiatives with SSCM in collaboration with strategic partners. For this purpose, education and training could be instrumental for creating awareness among supply chain and logistics managers in Pakistan. In this context, the role of top management is critical in making SSCM a crucial component of an organization’s strategy. The successful implementation of GSCM requires coordination between all level of employees from bottom-line workforce to top management (Govindan et al., 2014).

It is estimated that a majority of the world’s manufacturing will be completed in Asia in the next two decades (Zhu et al., 2008). Thus, in order to remain competitive, it is imperative for Pakistani industries to undertake SSCM initiatives in order to show their commitment to international buyers. In this regard, top management needs to develop a strategic posture along with supply chain partners for fully integrating SSCM into their strategy. This is also critical for those firms which are part of a manufacturing cluster linked to a global value chain.

Luken and Stares (2005) find that many of the SMEs in developing countries perceive that responding to the environmental and social requirements of global buyers will make them lose their competitive advantage. However, they contend that the supply chain pressures could be an opportunity for export-oriented firms in Pakistan to enhance their social and environmental performance. Therefore, the top management and global supply chain managers need to have a more strategic perspective of global supply chain initiatives and work on developing the strategic capability for harnessing the benefits of SSCM.

Organizations as part of supply chains in Pakistan need to view compliance as an opportunity to innovate. Supply chain managers need to have the skills to work with supply
chain partners including other organizations, suppliers, and customers. Supply chain partners need to experiment with sustainable processes, materials and technologies. This could be instrumental to enhance SSCM (Nidumolu et al., 2009).

The study was designed to make a contribution to the literature on SSCM by employing the RBV in order to explore how a firm’s assets and capabilities such as SC and SE relate to SSCM and OP in a developing country context. In Pakistan, organizations are more likely to employ SC and SE for achieving OP. However, relatively less emphasis is placed on linking SE and SC to SSCM. Pakistani organizations need to integrate SSCM into their business strategies. Furthermore, the aforementioned results demonstrate that although managers believe that SC and SE lead to SSCM, as well as to competitiveness and economic performance of the firm, the companies do not have well-established sustainable supply chain practices due to lack of overall environmental concerns. It is concluded that organizations in Pakistan though have some degree of involvement in SSCM but still face some challenges.

Due to various contextual constraints in Pakistan, organizations employ SC and SE to achieve a higher level of OP instead of improving SSCM. In order to compete in the international market, Pakistani organizations need to view SSCM as a key driver of competitiveness and more closely link SE and SC with a supply chain-wide drive for SSCM.

Min and Kim (2012) report that the literature on GSCM is mainly focused on developed countries and relatively few studies examine SSCM in a developing country context. Govindan et al. (2014) also contend that little is known with regard to the barriers to GSCM in a developing country context. Future studies may attempt to link SE and SC to SSCM instead of merely investigating environmental or sustainability issues related to a supply chain. Further research involving broader samples across other developing countries is clearly needed. Future researchers can replicate the current study for comparative purpose as organizations operating in developing countries face similar challenges. Given that the supply chains in developing countries could be different from developed countries (Tipu and Fantazy, 2014), future comparative studies can also examine the relationships of SE and SC to SSCM in developed as well as developing countries. These comparative studies are pertinent, given that the developed economies also face serious, but different, domestic issues such as low-cost competition from abroad.

The current study has some possible limitations. First, all possible efforts were made to ensure that the questionnaires were distributed to the respondents who were involved in the operations and supply chain management activities. Also, it was assumed that the respondents have some knowledge of the study variables relevant to their organization. However, in some cases, this assumption may not hold due to the limited knowledge of the respondents. Second, with regard to the research variables, the assessment items were not exhaustive. Therefore, additional items could be considered in future research. Third, the generalizability of results needs to be considered with caution, given that the snowball technique was employed.

References


Further reading


Construct 1: strategic entrepreneurship (SE) (Covin and Slevin, 1990)

In general, the top managers of my firm favor...

<table>
<thead>
<tr>
<th>Items</th>
<th>A strong emphasis on the marketing of tried-and-true products or services</th>
<th>A strong emphasis on R&amp;D, technological leadership, and innovations</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>SE2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

How many new lines of products or services has your firm marketed in the past 5 years?

<table>
<thead>
<tr>
<th>Items</th>
<th>How many new lines of products or services</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE3</td>
<td>No new lines of products or services</td>
<td>5.06</td>
<td>1.32</td>
</tr>
<tr>
<td>SE4</td>
<td>Many new lines of products or services</td>
<td>5.13</td>
<td>1.35</td>
</tr>
</tbody>
</table>

In dealing with its competitors, my firm...

<table>
<thead>
<tr>
<th>Items</th>
<th>Typically seeks to avoid competitive clashes, preferring a &quot;live-and-let-live&quot; posture</th>
<th>Typically adopts a very competitive, &quot;undo-the-competitors&quot; posture live</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>SE6</td>
<td>Is very seldom the first business to introduce new product/services, administrative techniques, operating technologies, etc.?</td>
<td>Is very often the first business to introduce new product/services, administrative techniques, operating technologies, etc.?</td>
<td>4.98</td>
<td>1.34</td>
</tr>
</tbody>
</table>

In general, the top managers of my firm have...

<table>
<thead>
<tr>
<th>Items</th>
<th>A weak tendency to be ahead of others in introducing novel ideas or products</th>
<th>A strong tendency to be ahead of others in introducing novel ideas or products</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>SE9</td>
<td>A strong proclivity (liking) for low-risk projects (with normal and certain rates of return)</td>
<td>A strong proclivity (liking) for high risk projects (with chances of very high returns)</td>
<td>4.92</td>
<td>1.35</td>
</tr>
<tr>
<td>SE10</td>
<td>Willingness to commit resources to projects where cost of failure may be low</td>
<td>Willingness to commit resources to projects where cost of failure may be high</td>
<td>4.97</td>
<td>1.29</td>
</tr>
</tbody>
</table>

Construct 2: social capital (SC) (Lawson et al., 2008; Krause et al., 2007)

While answering the following questions, recall your business relationship with one of your supply chain partners

<table>
<thead>
<tr>
<th>Items</th>
<th>Strongly disagree</th>
<th>Strongly agree</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC1</td>
<td>Both firms share the same business values</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>SC2</td>
<td>The parties often agree on what is in the best interest of the relationship</td>
<td>4.82</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td>SC3</td>
<td>This supplier shares our goals for this business</td>
<td>4.67</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td>SC4</td>
<td>Both firms have a shared vision and a common understanding</td>
<td>4.57</td>
<td>1.33</td>
<td></td>
</tr>
<tr>
<td>SC5</td>
<td>The relationship with key suppliers is characterized by close, personal interaction at multiple levels</td>
<td>4.77</td>
<td>1.22</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
SC6 The relationship with key suppliers is characterized by mutual respect at multiple levels 4.69 1.15
SC7 The relationship with key suppliers is characterized by mutual trust at multiple levels 4.95 1.13
SC8 The relationship with key suppliers is characterized by personal friendship at multiple levels 4.95 1.26
SC9 There is high corporate level communication on important issues with key suppliers 4.95 1.26
SC10 We have very frequent face-to-face planning with key suppliers 4.84 1.15
SC11 Our engineers and sales staff have a close relationship with our suppliers’ staff 4.81 1.26
SC12 In the development process, direct communication is bilateral rather than unilateral 4.81 1.24
SC13 Frequent contact between our key suppliers and our engineers is important 4.78 1.19
SC14 Communication with our key suppliers often begins to occur earlier in the development process 4.89 1.15
SC15 Through informal discussion, our key suppliers often communicate important engineering information to us 5.01 1.25

Construct 3: sustainable supply chain management (Paulraj, 2011)
Our firm and supply chain partners work closely to achieve the followings

<table>
<thead>
<tr>
<th>Items</th>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSCM1</td>
<td>Decrease in cost of material purchased</td>
<td>5.09 1.28</td>
</tr>
<tr>
<td>SSCM2</td>
<td>Decrease in cost of energy consumption</td>
<td>5.09 1.21</td>
</tr>
<tr>
<td>SSCM3</td>
<td>Decrease in fee for waste discharge</td>
<td>5.01 1.14</td>
</tr>
<tr>
<td>SSCM4</td>
<td>Improvement in return on investment</td>
<td>4.91 1.22</td>
</tr>
<tr>
<td>SSCM5</td>
<td>Reduction in air emission</td>
<td>4.86 1.14</td>
</tr>
<tr>
<td>SSCM6</td>
<td>Reduction in waste (water and/or solid)</td>
<td>5.07 1.17</td>
</tr>
<tr>
<td>SSCM7</td>
<td>Decrease in consumption of hazardous/harmful/toxic materials</td>
<td>5.04 1.30</td>
</tr>
<tr>
<td>SSCM8</td>
<td>Decrease in frequency for environmental accidents</td>
<td>5.07 1.15</td>
</tr>
<tr>
<td>SSCM9</td>
<td>Increase in energy saved due to conservation and efficiency improvements</td>
<td>4.92 1.14</td>
</tr>
<tr>
<td>SSCM10</td>
<td>Improvement in overall stakeholder welfare or betterment</td>
<td>4.96 1.06</td>
</tr>
<tr>
<td>SSCM11</td>
<td>Improvement in community health and service</td>
<td>4.91 1.23</td>
</tr>
<tr>
<td>SSCM12</td>
<td>Reduction in environmental impacts and risks to general public</td>
<td>4.98 1.19</td>
</tr>
<tr>
<td>SSCM13</td>
<td>Improvement in occupational health and safety of employees</td>
<td>4.90 1.18</td>
</tr>
<tr>
<td>SSCM14</td>
<td>Improved awareness and protection of the claims and rights of people in community served</td>
<td>4.93 1.18</td>
</tr>
</tbody>
</table>

Construct 4: organizational performance (OP) (Fantazy et al., 2010)
Please indicate how each of the following performance measures compare with those of your major competitors

<table>
<thead>
<tr>
<th>Items</th>
<th>Very weak</th>
<th>Very strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP1</td>
<td>Average net profit</td>
<td>5.14 1.31</td>
</tr>
<tr>
<td>OP2</td>
<td>Average sales growth rate</td>
<td>5.15 1.27</td>
</tr>
<tr>
<td>OP3</td>
<td>Order lead-time</td>
<td>5.27 1.06</td>
</tr>
<tr>
<td>OP4</td>
<td>Response time to customer query time</td>
<td>5.21 1.10</td>
</tr>
<tr>
<td>OP5</td>
<td>Level of customer perceived value of product</td>
<td>5.25 1.01</td>
</tr>
<tr>
<td>OP6</td>
<td>Level of service systems to meet particular customer need</td>
<td>5.30 1.08</td>
</tr>
</tbody>
</table>
About the authors

Dr Syed Awais Ahmad Tipu holds Doctor of Science and Master of Science degrees from George Washington University, USA. He is currently working as Associate Professor in the College of Business Administration, University of Sharjah, United Arab Emirates. He has published in leading journals such as *Journal of Business Research, Journal of Business Ethics* and *International Journal of Logistics Management*. His research interests include sustainable supply chain, entrepreneurial cognition, work values of aspiring entrepreneurs, and necessity drive entrepreneurship. Dr Syed Awais Ahmad Tipu is the corresponding author and can be contacted at: Stipu@sharjah.ac.ae

Kamel Fantazy received his PhD in Operations Management from the Sprott School of Business at Carleton University, Ottawa, Canada. He worked as Research Fellow and taught in Canada. His research interests include supply chain management, improving performance of production and operation systems, international technology transfer and supply chain sustainability. He has 14 years of industrial experience in various managerial positions. He is currently Assistant Professor in the Faculty of Business and Economics, the University of Winnipeg, Canada. He has several research papers and was the recipient of the Best Research Paper Awards in the POM division at the ASAC conferences held in Halifax, Canada, 2008, Memorial University, 2012, and Nipissing University, 2104. He received the Best Paper Award in the Seventh Annual Supply Chain Symposium in Toronto, Canada, October 28–30, 2009.

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Authentic leadership and creativity

The intervening role of psychological meaningfulness, safety and work engagement

Richa Chaudhary and Chinmay Panda
Department of Humanities and Social Sciences, Indian Institute of Technology Patna, Patna, India

Abstract

Purpose – The purpose of this paper is to develop and test a conceptual model which explains whether and how authentic leadership, through psychological meaningfulness, safety and work engagement, influences employee creativity. The authors delineate two pathways from authentic leadership to work engagement, which in turn foster creativity. The first pathway illuminates how authentic leadership generates psychological meaningfulness which enhances work engagement and creativity subsequently, while the second elucidates the indirect impact of authentic leadership on creativity via psychological safety and work engagement.

Design/methodology/approach – Data were collected from 300 employees working in heavy engineering and automobile industry in India. Regression analysis was used to analyze data with the help of SPSS 24. Serial mediation effects were tested with the help of bootstrapping procedures using SPSS process macro (Hayes, 2013).

Findings – Psychological meaningfulness and work engagement were found to mediate the relationship of authentic leadership with creativity both independently and in series. Psychological safety failed to transfer the effect of authentic leadership on both work engagement and creativity.

Practical implications – Organizations may reap the benefits of a creative and engaged workforce by selecting, nurturing and developing authentic leaders.

Keywords India, Creativity, Work engagement, Authentic leadership, Psychological safety, Psychological meaningfulness

Paper type Research paper

1. Introduction

With the increased instances of corporate fraud and scams being reported across the globe, the need for more positive, genuine and value-based form of leadership has come to the fore. Stakeholders expect a high degree of integrity in the actions of their leader. In this context, authentic leadership has received a great deal of attention in the corporate sphere followed by an upsurge in the academic research. The proponents of authentic leadership emphasize upon its potential in reinstating confidence, trust, hope, resilience and optimism among stakeholders (Avolio et al., 2004). Walumbwa et al. (2008) have defined authentic leadership as “a pattern of leader behavior that draws upon and promotes both positive psychological capacities and a positive ethical climate, to foster greater self-awareness, an internalized moral perspective, balanced processing of information and relational transparency on the part of leaders working with followers, fostering positive self-development” (p. 94). Here, self-awareness refers to how one views oneself and the degree to which one is aware of one’s strengths/weaknesses and one’s impact on other people. Internalized moral perspective symbolizes self-regulation where one’s behaviors and actions are guided by one’s personal values and moral standards. Relational transparency relates to the extent to which one shares information openly and displays one’s true thoughts and emotions. Balanced processing refers to the degree to which a leader takes into account all the available information before arriving at a decision. Thus, authentic leaders exhibit high
ethical and moral standards by demonstrating a high level of consistency between their values and actions and unwillingness to compromise on internalized values in response to situational demands (Avolio et al., 2004; Barke and Cooper, 2006).

Due to the novelty of the construct, the research on authentic leadership is in very nascent stages. Though both academicians and practitioners theoretically propose that authentic leadership encourages positive employee attitudes and behaviors in the organizations, the empirical evidence to test the validity of above assertion is limited (Gardner et al., 2011). Moreover, the underlying psychological mechanisms through which authentic leadership influences follower outcomes are not understood clearly (Avolio and Mhatre, 2012; Gardner et al., 2011; Rego et al., 2014). Although recent studies have highlighted psychological capital and perceived organizational support as the possible mediators (Penger and Černe, 2014; Rego et al., 2012) of the relationship between authentic leadership and employee attitudes and behaviors, exploring alternative underlying mechanisms and processes facilitates better understanding of the nature of relationships. This is even more important for an emerging construct like authentic leadership which is in its early stages of development.

The present paper is an attempt to address the above identified gaps in the literature by providing a conceptual framework and empirical test of the relationship between authentic leadership and creativity. Studying creativity as an outcome variable bears merit as creative workforce can provide organizations with the needed competitive advantage in this dynamic business environment. Creativity involves the creation of novel and valuable ideas about products, services, methods and processes (Amabile, 1988; Shalley and Zhou, 2008). Creativity has been proposed to be a function of both personal and contextual factors (Amabile, 1988). Leadership as an important contextual factor has been recognized to play an essential role in stimulating creativity and innovation among followers (Oke et al., 2009). Different leadership styles such as servant leadership (Yoshida et al., 2014), transformational leadership (Shin and Zhou, 2003; Gumushoglu and Isev, 2009) and empowering leadership (Zhang and Bartol, 2010) have been related to creativity in the past. However, the ideal mix of leadership behaviors and creativity is yet to be discovered (Černe et al., 2013). At the same time, the underlying psychological mechanisms explaining the linkage between leadership and creativity are not known (Crossan and Apaydin, 2010; Černe et al., 2013).

The goal of the present study is to develop an integrated model to examine how authentic leadership affects the creativity of employees working in heavy engineering and automobile organizations in India. Using arguments from self-determination theory (Ryan and Deci, 2000), the paper shows how authentic leadership fosters employee creativity via shaping their perceptions of psychological meaningfulness, psychological safety and work engagement. Here, psychological meaningfulness refers to the experience of feeling worthwhile, useful and valuable (Kahn, 1990). People experienced meaningfulness when they feel that they made a difference and were not taken for granted (Kahn, 1990). Psychological safety is defined as “feeling able to show and employ one’s self without fear of negative consequences to self-image, status, or career” (Kahn 1990, p. 708). People feel safe in the situations when they believe that they will not bear the consequences of bringing their true selves to work roles (Kahn, 1990). Kahn (1990) proposed that the level of engagement/representation of self in a role depends upon three psychological conditions: meaningfulness, safety and availability. Work engagement refers to “a positive, satisfying, job-related state of mind that is characterized by vigor, dedication, and absorption” (Schaufeli et al., 2002). We delineate two pathways from authentic leadership to work engagement, which in turn foster creativity. The first pathway illuminates how authentic leadership generates psychological meaningfulness which enhances work engagement and creativity subsequently, while the second explicates the indirect influence of authentic leadership on creativity through psychological safety and work engagement.
The present study in Indian context is justified in that the national culture has a significant influence on the perceptions of authenticity (Cooper et al., 2005), which makes it important to gather evidences on positive outcomes of authentic leadership in diverse cultures to strengthen the theory building by expanding the nomological network of authentic leadership. This study offers several contributions to the literature. First, by exploring work engagement and creativity as outcomes, the study adds to the limited knowledge on benefits of authentic leadership for organizations. Second, by exploring psychological states of meaningfulness, safety and work engagement as sequential mediators, the study contributes to the understanding of the complex psychological mechanisms underlying the relationship between authentic leadership and employee creativity. Third, the study advances the theory building on authentic leadership, work engagement and creativity by providing empirical evidence on the nature of relationships among unique combination of variables from a unique cultural context of India. Finally, the study may provide practitioners with a tool to foster employee engagement and creativity, the key determinants of competitive edge in this competitive business era.

The remainder of the paper has been structured as follows. Section 2 presents the theoretical framework and research hypotheses. Sections 3 and 4 discuss the research methodology and results. Sections 5 and 6 present the discussion and implications. Section 7 covers study limitations and directions for future research.

2. Theoretical framework and research hypotheses

Research suggests that attitudes and beliefs of individuals are significantly influenced by information embedded in the social context (Salancik and Pfeffer, 1978). As leaders form a salient component of organizational environment, they may be expected to exercise significant influence on individuals’ attitudinal and behavioral outcomes. The present study builds on the theoretical framework of self-determination theory to link authentic leadership with work engagement and creativity through psychological meaningfulness and safety. Self-determination theory states that when employees’ intrinsic psychological needs of competence, relatedness and autonomy are satisfied, they are likely to put in more effort and bring more of their selves into work (Ryan and Deci, 2000). We propose that authentic leaders by demonstrating integrity, transparency and openness enhance the sense of competence, relatedness and autonomy resulting in high intrinsic motivation and creativity. As ethics forms an important component of authentic leadership, and considering the fact that research evidence on authentic leadership per se is limited, evidences relating other ethical/oral and value-based forms of leadership such as ethical leadership, transformational leadership, and charismatic leadership have been used to frame study hypotheses.

2.1 Authentic leadership and creativity

Authentic leadership by satisfying employees’ need for self-determination enhances intrinsic motivation (Zhou and Ren, 2012), which has been identified as an essential precursor to creativity (Amabile, 1997). Intrinsically motivated employees are likely to be more creative due to their cognitive flexibility, learning orientation, persistence and willingness to take risk (Amabile, 1997; Rego et al., 2014; Zhou and Ren, 2012). With regard to the relationship of authentic leadership with creativity, Cerne et al. (2013) in a study among team leaders and their followers in a Slovenian manufacturing and processing company reported both direct and indirect influence of authentic leadership on employees’ creativity via perception of support for innovation. Rego et al. (2012) in a study among employees and supervisors of commerce organizations in Portugal reported significant association between authentic leadership and creativity. They established partial mediation effect of employees’ psychological capital in the above relationship. Further, Rego et al. (2014) in a study among
retail organizations in Portugal reported the sequential mediation effect of positive affect and hope on the relationship between authentic leadership and creativity. Thus, the research evidence on the relationship between authentic leadership and creativity is limited with only three studies exploring the linkage, as discussed above. Therefore, we propose:

**H1.** Authentic leadership relates positively to employee creativity.

### 2.2 Psychological meaningfulness as mediator

Bass (1999) defined meaning as “perceived fit between the individual’s values and beliefs, and the requirements of a work role.” People derive meaning from consonance between their self-concept and behaviors (Gecas, 1982). Avolio et al. (2004) proposed that authentic leaders may promote the perceptions of meaning at work by encouraging the followers to act authentically by displaying consistency in their actions and moral beliefs. In this direction, charismatic leaders have been reported to increase the meaningfulness of goals by articulating vision and mission congruent with followers’ self-concept (Shamir et al., 1993). Similarly, transformational leadership has been reported to enhance “meaningfulness by infusing work and organization with moral purpose and commitment” (Shamir et al., 1993, p. 578). Since authentic leadership inspires authentic followership through positive modeling (Gardner et al., 2005), it is likely to enhance experienced meaningfulness at work by inducing self-awareness and self-regulation among followers. Bailey and Madden (2017) asserted that meaningfulness arises “when an individual perceives an authentic connection between work and a broader transcendent life purpose beyond the self.” Higher self-awareness develops perspective taking capacity among individuals (Kegan, 1982) and helps them draw meaning in the world around them through introspection and self-reflection (Gardner et al., 2005). Authentic leaders give meaning to tasks, efforts and goals by relating them to followers’ values and beliefs. They make employees internalize moral values and identity and link it with their self-concept, thereby enhancing work meaningfulness (Zhu et al., 2016).

Fostering a sense of belongingness has been reported to be an important mechanism to enhance work meaningfulness (Rosso et al., 2010; Lips-Wiersma and Wright, 2012). Authentic leaders may also generate experiences of meaningfulness by increasing followers’ identification with the collective through the articulation of common vision and goals (Dutton and Hapathy, 2003). Authentic leaders facilitate strong sense of belongingness by providing social support and promoting strong identification with the organization (Lips-Wiersma and Wright, 2012; Walumbwa et al., 2011; Hansen et al., 2013; Zhu et al., 2016). Again, authentic leaders may foster experienced meaningfulness at work by promoting high-quality connections with the coworkers and the organization (May et al., 2004; Kurtessis et al., 2015). Wang and Xu (2017) provided support for the above arguments by demonstrating a strong influence of ethical leadership on work meaningfulness among employees from the diverse nature of organizations in China.

Authentic leaders empower their followers by identifying their potentials and building their strengths which enhances the sense of purposefulness and future and leads to greater psychological meaningfulness (George, 2003). The balanced processing component of authentic leadership suggests that authentic leaders solicit information from the followers, even if it challenges their existing positions, before making any decision. Inviting inputs from the followers and considering them before arriving at a decision is likely to instill confidence among the followers that their opinion counts. This will make employees feel that their work is valued and that they are respected members of the organization, which will make them sense greater meaningfulness in and at work (Stephens and Carmeli, 2017). By expressing confidence in followers, authentic leaders are likely to augment the self-esteem and self-worth of the followers which consequently leads to higher task meaningfulness (Shamir et al., 1993). Although Avolio et al. (2004) in their proposed
theoretical framework linking authentic leadership with followers’ attitudes and behaviors proposed an association between authentic leadership and meaningfulness at work, no empirical work has tested the above proposition till date. Recognizing this dearth of empirical evidence, May et al. (2004) also underscored the need for more research aimed at exploring the relationship of authentic leadership with meaningfulness at work.

Further, the feeling of being valued by others is likely to make employees feel confident of their competencies to generate creative ideas to solve problems and engage in creative behaviors (Cohen-Meitar et al., 2009). This sense of positive regard, when others believe in one’s competence and worthiness, makes individuals feel safe in experimenting in innovative ways to solve organizational problems and hence, display higher creativity. Very limited attempt has been made in the literature to explore the relationship of meaningfulness and creativity. In the only study in the literature, as far as we know, Cohen-Meitar et al. (2009) provided empirical evidence in support of the positive linkage between meaningfulness (at work and in work) and creativity. They reported that experienced meaningfulness in and at work enhances employee creativity through the sequential mediation effects of organizational identity and positive psychological experiences (vitality, positive regard and mutuality, and organization-based self-esteem).

Based on the above arguments and literature support, we propose that authentic leadership will lead to enhanced creativity among the followers by increasing their perceptions of psychological meaningfulness:

\[ H2. \text{ Psychological meaningfulness mediates the relationship between authentic leadership and creativity.} \]

2.3 Psychological safety as mediator

Leaders have been identified to shape the perceptions of psychological safety through the social exchange process which provides key information regarding trust, competence and support to employees (Kahn, 1990; Edmondson, 2004). Several leadership styles such as transformational leadership (Detert and Burris, 2007), servant leadership (Schaubroeck et al., 2011) and ethical leadership (Walumbwa and Schaubroeck, 2009) have been examined as antecedents to psychological safety in past. Although Rego et al. (2014) proposed positive relation between authentic leadership and creativity with psychologically safe climate created by authentic leaders as the prime theoretical argument, the role of psychological safety as the underlying process was not explored empirically. We argue that authentic leader’s openness and transparent exchanges with the followers will enhance followers’ perceptions of psychological safety at work which makes the employees feel safe and secure in bringing their true selves to work without fear of negative consequences to their status, image and career (Kahn, 1990). This lack of fear will also make employees come up and present their ideas freely and hence, enhance their creativity.

Guided by internal moral standards, authentic leaders through the objective analysis of information and open exchanges with employees build a relationship of trust and respect with the followers (Avolio et al., 2004). Authentic leaders by being straightforward in their communication are likely to disclose reasons and motives behind all the decisions which may foster the perceptions of psychological safety (Han et al., 2017). Also, relational transparency on the part of leader is likely to result in greater information being available to employees through regular feedbacks from the leader, leading to enhanced concentration in work (Han et al., 2017). Han et al. (2017) in a study among employees of a telecommunications firm in China reported that leaders’ transparency influences employees’ creativity through affective (psychological safety) and cognitive mechanisms (ability to focus attention). Regular information exchanges between leader and follower breeds
interpersonal trust which leads to creativity (Edmondson, 1999). The leader-member relations characterized by trust and openness make employees feel safe to take risk and express dissenting views without apprehension. This perception of psychological safety is likely to result in employees being able to question, discuss and express unconventional perspectives freely (Han et al., 2017). In support, Liu et al. (2015) reported that the relationship of authentic leadership with whistle blowing was mediated by psychological safety. Interpersonal environment characterized by safety triggers employee motivation to innovate (Edmondson, 1999).

Therefore, we propose:

**H3.** Psychological safety mediates the relationship between authentic leadership and creativity.

2.4 Authentic leadership and work engagement

An authentic leader by being true to his values and showing a highest level of moral conduct will inspire followers to put in more of their selves to work thereby increasing their work engagement levels. By ensuring transparency in interpersonal transactions, authentic leaders build a relationship of trust with the followers and help them realize their true potentials. Authentic leaders augment work engagement by making followers discover their talent, assisting them in using this talent to create a better fit with the work goals, enriching their work and providing opportunities for developing rewarding relationships with coworkers (Gardner et al., 2005; May et al., 2004). Research evidence on the relationship between authentic leadership and work engagement is sparse with only three studies, to the best of our knowledge, providing evidence in the support of the relationship (see Giallonardo et al., 2010; Alok and Israel, 2012; Penger and Cerne, 2014). Giallonardo et al. (2010) in a study among new graduate nurses reported work engagement as an importance mediating mechanism which links authentic leadership with job satisfaction. Alok and Israel (2012) in a study among working professionals in India reported a positive impact of authentic leadership on work engagement through the mediating influence of psychological ownership. Penger and Cerne (2014) in a multilevel study among employees and managers of a Slovenian manufacturing company found significant positive connection between authentic leadership and work engagement. They also reported that the above relation was partially mediated by perceived organizational support. We hypothesize:

**H4.** Authentic leadership relates positively to work engagement.

2.5 Work engagement as mediator

According to broaden and build theory (Fredrickson, 2001), engaged employees more often experience positive emotions which broadens their intellectual and psychological resources prompting them to explore novel and unconventional ways of doing their job. These positive psychological experiences motivate employees to learn, assimilate new information and produce ingenious solutions and ideas (Cohen-Meitar et al., 2009). Further, engaged employees possess higher energy required to pursue creative paths than non-engaged ones (Atwater and Carmeli, 2009). In support, Demerouti et al. (2015) reported a positive linkage between work engagement and creativity among employees in the Netherlands. Bakker and Xanthopoulou (2013) also demonstrated a significant relationship between work engagement and creativity in a study among school principals and teachers. As authentic leadership enhances work engagement, which, in turn, encourages employees to exhibit creative behaviors, we propose:

**H5.** Work engagement mediates the relationship of authentic leadership with creativity.
2.6 Sequential mediation

As discussed, psychological meaningfulness and psychological safety have been advocated as one of the essential conditions for the employees to invest themselves physically, cognitively and emotionally into work (Kahn, 1990; May et al., 2004; Edmondson and Lei, 2014; Frazier et al., 2017). Experienced meaningfulness enhances work passion and makes individuals expend more effort in their work to produce better workplace outcomes (Hodson, 2004). In support, Demirtas et al. (2017) in a study among aviation sector employees in Turkey reported an indirect influence of ethical leadership on employees’ engagement through meaningful work. In addition, the perception of psychological safety increases employee engagement by overcoming their fear of bringing original selves to work (Kahn, 1990; May et al., 2004). May et al. (2004) provided empirical support for the positive relation between employee engagement and psychological safety. Also, Frazier et al. (2017) in a meta-analytical review on psychological safety reported engagement as one of the most important outcomes of psychological safety. Enhanced work engagement resulting from the feeling of psychological meaningfulness and psychological safety at work provides the needed motivation for the individuals to engage in creative pursuits.

As authentic leadership promotes a sense of psychological meaningfulness and safety which determine the level of work engagement and work engagement in turn leads to creativity, we expect that the relationship of authentic leadership with creativity will be sequentially mediated by psychological meaningfulness/psychological safety and work engagement. Thus, we hypothesize:

H6. Psychological meaningfulness and work engagement sequentially mediate the relationship between authentic leadership and creativity.

H7. Psychological safety and work engagement sequentially mediate the relationship between authentic leadership and creativity.

Figure 1 shows the hypothesized research model.

3. Research methodology

3.1 Participants and procedure

The sample for the present study consisted of employees from heavy engineering and automobile industry in India. A non-experimental correlational field study design was used for the purpose of data collection. A total of 550 questionnaires were distributed to both the employees and supervisors. Data were collected by second author of the paper via personal visits to the organizations using paper and pencil questionnaires during May–July 2017. The sampled organizations were located in Gurugram and Ranchi regions of
India. In total, 318 questionnaires were received out of which 18 were discarded due to incomplete information and extreme responses. The final sample for the study consisted of 300 employees (response rate 54.5 percent), where employees reported their immediate supervisor’s authentic leadership and their own psychological meaningfulness, safety and work engagement. The supervisors provided ratings on employee creativity. The individuals were also requested to furnish information on their gender, age, education, position in organizational hierarchy and tenure. With regard to sample characteristics, there were 247 males and 53 females of which 52 percent were graduates, 28 percent were post graduates and rest had other educational qualifications like diploma and PhD. In terms of age, 54.6 percent of the respondents were below the age of 30 years, 38.7 percent were in the age group of 31–40 years and the rest 6.6 percent were above 40 years. In total, 44 percent of the employees had worked with their current employer for 1–5 years, while 28.7 percent of the respondents had been associated with their current employer for 10–15 years. As to position in the organizational hierarchy, 48 percent (majority) of the employees belonged to middle management level, 44 percent worked at junior management level and rest 8 percent belonged to senior management level in the sampled organizations.

3.2 Measures
Authentic leadership was assessed using 16-item authentic leadership inventory by Neider and Schriesheim (2011). Employees responded to the authentic leadership of their immediate supervisors on a five-point Likert scale where 1 = strongly disagree and 5 = strongly agree. Sample scale items include “My leader openly shares information with others” and “My leader resists pressures on him/her to do things contrary to his/her beliefs.” In line with the previous studies, we use authentic leadership as a common core higher order construct than its individual dimensions due to significant overlap among them as they all form the part of leaders’ self-regulatory processes (Rego et al., 2012; Cerne et al., 2013). The Cronbach’s α value for the instrument was found to be 0.83.

Work engagement. Nine-item shortened version of Utrecht work engagement scale by Schaufeli et al. (2006) was used to assess work engagement. The items of the scale represent three dimensions – vigor, dedication and absorption. Each dimension consists of three items measured on a five-point Likert scale where 1 = strongly disagree and 5 = strongly agree. The sample items for each dimension are: “When I get up in the morning, I feel like going to work (vigor),” “I am enthusiastic about my job (dedication)” and “I feel happy when I am working intensely (absorption).” The instrument was reliable with a Cronbach’s α value for the scale was 0.64.

Creativity. In total, 12 items taken Zhou and George’s (2001) creativity scale were used to measure employee creativity. The creativity of employees was rated by their supervisors on a five-point Likert scale (1 = not at all characteristics to 5 = very characteristics). The typical sample items are “Often has a fresh approach to problems” and “Searches out new technologies, processes, techniques, and/or product ideas.” The instrument was reliable with a Cronbach’s α value of 0.83.

Psychological meaningfulness. Psychological meaningfulness was measured using six-item scale by May et al. (2004). The items were rated on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The sample scale items include “The work I do on this job is very important to me” and “My job activities are personally meaningful to me.” The reliability value for the scale was found to be 0.63.

Psychological safety. Psychological safety was measured using three-item scale by May et al. (2004). “I am not afraid to be myself at work” and “I am afraid to express my opinions at work (R)” represent sample scale items. The Cronbach’s α value for the scale was found to be 0.61, which was above the minimum acceptable level of 0.60 (Hair et al., 2010).
Control variables. We statistically controlled for the influence of demographic variables such as gender, age, education, position and tenure, as they have been reported to influence employee creativity (Furnham and Nederstrom, 2010).

Data analysis. The hypotheses were tested using regression analysis with the help of SPSS 24. The serial mediation effects were examined using bootstrapping procedures with the aid of SPSS process macro (Hayes, 2013). This method estimates direct and indirect effects using ordinary least square procedures. Effect size estimates of indirect effects were drawn using bootstrap confidence intervals. Bootstrapping is a robust non-parametric technique for hypotheses testing and effect size estimation without making any assumptions about the distribution of the statistic.

Since the data were collected from employees working in different heavy engineering and automobile organizations from various regions in India, intra-class correlation coefficient (ICC) values was assessed to detect any systematic variance in the sample. ICC (1) values for the study variables were found to be insignificant (0.03–0.10) and were below the minimum acceptable level (0.12, James, 1982) to justify aggregation at the organizational level. Since no significant group level differences exist to warrant group level comparison, there was no need to carry out multilevel analysis (Bliese, 2000).

4. Results
Means, standard deviations and inter-correlations among study variables are shown in Table I.

As can be seen, mean score for all the study variables except psychological safety was above 4. Authentic leadership exhibited significant correlation with psychological meaningfulness, work engagement and creativity. Psychological safety displayed significant correlation with creativity. However, authentic leadership and work engagement failed to show any correlation with psychological safety.

4.1 Hypotheses testing
To assess the indirect effect of authentic leadership on creativity, two alternate models were proposed. The first model tests the effect of authentic leadership on creativity via psychological meaningfulness and work engagement. The second model involves psychological safety and work engagement as serial mediators of the relationship between authentic leadership and creativity.

Model 1: psychological meaningfulness and work engagement as mediators. Full process model showing all path coefficients is shown in Figure 2. Total, direct and indirect effects are shown in Table II.

The total effect of authentic leadership on creativity was found to be significant ($c_1 = 0.5283$, $p < 0.001$). The direct effect model shows that authentic leadership related significantly with psychological meaningfulness ($a_1 = 0.3681$, $p < 0.001$) and work engagement ($a_2 = 0.4268$, $p < 0.001$). Psychological meaningfulness also showed

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Variables</th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AL</td>
<td>4.09</td>
<td>0.462</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>MNG</td>
<td>4.31</td>
<td>0.449</td>
<td>0.379***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>SFT</td>
<td>3.21</td>
<td>0.923</td>
<td>0.048</td>
<td>-0.085</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>WE</td>
<td>4.10</td>
<td>0.44</td>
<td>0.538***</td>
<td>0.409**</td>
<td>0.034</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>CRVTY</td>
<td>4.28</td>
<td>0.433</td>
<td>0.563***</td>
<td>0.465**</td>
<td>0.115**</td>
<td>0.603**</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: AL, authentic leadership; MNG, psychological meaningfulness; SFT, psychological safety; WE, work engagement; CRVTY, creativity. **Significant at the 0.01 level (two-tailed)
significant relationship with work engagement ($a_3 = 0.2346, p < 0.001$) and creativity ($b_1 = 0.1276, p < 0.01$). Further, the effect of work engagement on creativity was significant ($b_2 = 0.3783, p < 0.001$). Thus, $H_1$ and $H_4$ were supported.

The examination of direct effect model unveils that after including psychological meaningfulness and work engagement as mediators, the effect of authentic leadership on creativity was reduced ($c_1 = 0.2872, p < 0.001$), without any drop in the significance level. This suggests partial mediation, which was confirmed through the analysis of indirect effects. As shown in Table II, the indirect effect of authentic leadership, independently through psychological meaningfulness (Indirect 1: 0.0470, CI (0.0087, 0.1044)) and work engagement (Indirect 3: 0.1615, CI (0.1036, 0.1446)), was significant. Thus, $H_2$ and $H_5$ were supported.

Further, the sequential mediation model was supported as the indirect effect of authentic leadership on creativity through psychological meaningfulness and work engagement in series was significant (Indirect 2: $a_1a_3b_2 = 0.0327, BC CI (0.0156, 0.0663)$). The overall model with psychological meaningfulness and work engagement as mediators was significant ($F$-value = 83.22, $p < 0.001$) and explained 45.76 percent of the variance in creativity. Thus, $H_6$ was supported.
**Model 2: psychological safety and work engagement as mediators.** Full process model showing all path coefficients is shown in Figure 3. Total, direct and indirect effects are shown in Table III.

Figure 3 shows that authentic leadership failed to exercise any significant effect on psychological safety ($a_1 = 0.0957$, ns). The relationship of psychological safety with work engagement was insignificant ($a_3 = 0.004$, ns) and its relationship with creativity reached significance only at the 0.05 level of significance. Further, authentic leadership related significantly with work engagement ($a_2 = 0.5128$, $p < 0.001$) and, work engagement in turn related significantly with creativity ($b_2 = 0.4148$, $p < 0.001$).

As discussed in the above section, the total effect of authentic leadership on creativity was significant. However, the analysis of direct effect model shows that after including psychological safety and work engagement in the model, the effect of authentic leadership on creativity was reduced ($c_1 = 0.3116$, $p < 0.001$), without any drop in the significance level. This indicates partial mediation. To confirm the partial mediation, indirect effects were examined.

From Table III, it can be seen that the indirect effect of authentic leadership on creativity was significant only through work engagement (Indirect 3: 0.2127, CI (0.1443, 3004)).

![Figure 3. Parameter estimates for a serial mediation effect of psychological safety and work engagement](image)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Point estimate</th>
<th>SE</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect 1: AL $\rightarrow$ SFT $\rightarrow$ CRVTY</td>
<td>0.0003</td>
<td>0.0056</td>
<td>-0.0038</td>
<td>0.0022</td>
</tr>
<tr>
<td>Indirect 2: AL $\rightarrow$ SFT $\rightarrow$ WE $\rightarrow$ CRVTY</td>
<td>0.0002</td>
<td>0.0016</td>
<td>-0.0022</td>
<td>0.0053</td>
</tr>
<tr>
<td>Indirect 3: AL $\rightarrow$ WE $\rightarrow$ CRVTY</td>
<td>0.2127</td>
<td>0.0396</td>
<td>0.1443</td>
<td>0.3004</td>
</tr>
<tr>
<td>Total indirect effect</td>
<td>0.2167</td>
<td>0.0404</td>
<td>0.1463</td>
<td>0.3078</td>
</tr>
<tr>
<td>Direct effect of AL on CRVTY ($c'$)</td>
<td>0.3116</td>
<td>0.048</td>
<td>0.2172</td>
<td>0.4061</td>
</tr>
<tr>
<td>Total effect of AL on CRVTY ($c$)</td>
<td>0.5283</td>
<td>0.0449</td>
<td>0.4399</td>
<td>0.6167</td>
</tr>
</tbody>
</table>

**Table III.** Total, direct and indirect effects of authentic leadership on creativity

**Notes:** BC, bias corrected. 5,000 bootstrap samples were requested
The indirect effect through psychological safety was insignificant as bootstrap confidence interval was found to contain 0 (Indirect 1: 0.0038, CI (−0.0038, 0.0202)). Thus, H3 was not supported.

Although the overall model with psychological meaningfulness and work engagement as mediators was significant (F-value = 81.01, p < 0.001) and explained 45.09 percent of the variance in creativity, the sequential mediation effect of psychological safety and work engagement was not supported (Indirect 2: a1a3b2 = 0.0002; BC CI (−0.0022, 0.0053)). Therefore, H7 was not supported.

5. Discussion and theoretical contributions
In the present study, we developed and tested a conceptual model which explains how authentic leadership influences employee creativity. For the purpose of the study, two models were conceptualized. In the first model, psychological meaningfulness and work engagement were theorized as the serial mediators of authentic leadership–creativity relationship. In the second model, psychological safety and work engagement were proposed to sequentially mediate the aforementioned relationship. The results provided support for the first model while second model was not supported. Authentic leadership was found to affect creativity both directly and indirectly via psychological meaningfulness and work engagement, independently and in sequence. The employees who rated their supervisor high on authentic leadership found meaning in their work and exhibited higher work engagement. Psychological meaningfulness promoted by the authentic leaders was found to make employees think and act creatively. The psychological condition of experienced meaningfulness also improved intrinsic motivation and made employees devote themselves completely to work (May et al., 2004). The engaged employees felt the need to express themselves in more creative and innovative ways to improve their performance. They were enthusiastic and willing to put in extra-effort in their work toward better performance. Engaged employees possess the energy and vigor required to follow creative paths (Atwater and Carmeli, 2009) and more frequent experience of positive affect (Fredrickson, 2001) aroused them to explore unconventional paths and display higher creativity (Filipowitz, 2006). This supports our argument that authentic leaders by making employees participate in the decision-making process and maintaining the relationship of trust through relational transparency enhance the perceptions of psychological meaningfulness. This feeling of being worthwhile and valuable enhances an individual’s belief in his capabilities to generate creative ideas and engage in creative pursuits. Further, the feeling of being valued increases psychological availability to engage in their creative activities (Kahn, 1990).

On the other hand, psychological safety was not affected by authentic leadership and consequently, it failed to mediate the linkage of authentic leadership with creativity. This was in contradiction with our expectations that authentic leaders generate the climate of psychological safety which makes employees feel safe in experimenting and taking risks. This finding suggests that psychological safety may not carry the influence of authentic leadership on followers’ creativity. It also signifies that psychological safety may not be a direct outcome of authentic leadership. There is a possibility that authentic leaders may promote safety through its effect on some intermediate variables. Another possible reason for these unanticipated findings could be the low reliability of three-item psychological safety scale. A review of psychological safety research also highlights measurement issues in accurately capturing the construct (Newman et al., 2017). Thus, we recommend that the study be replicated in future with more rigorous measures of psychological safety to confirm the study findings.

The results add to and extend the findings of sparse studies conducted in the arena of authentic leadership demonstrating its association with work engagement (Penger and Černe, 2014; Alok and Israel, 2012; Giallonardo et al., 2010) and creativity (Rego et al., 2012, 2014;
Cerne et al., 2013; Semedo et al., 2017) from different countries. The findings emphasize upon the need of authentic leaders in Indian heavy engineering and automobile organizations for nurturing positive employee attitudes and behaviors. These findings from India hold special significance in the light of recent reports of corporate irresponsibility and governance failures. These findings from India are also important as authentic leadership is a developing construct and evidences on its positive outcomes are needed from different countries in the journey toward strong theory development. Thus, this study strengthens the existing body of knowledge and contributes to authentic leadership theory building by providing much needed empirical evidence on its consequences from a unique cultural context.

Past studies have theorized psychological conditions of meaningfulness and safety as the possible mechanisms which can explain the linkage between authentic leadership and follower attitudes and behaviors (May et al., 2004; Rego et al., 2014), without exploring the relationships empirically. This study by simultaneously considering the above variables in the model provides an empirical analysis of the above relationships and extends the findings of all previous works on authentic leadership. The findings establish psychological meaningfulness and work engagement as independent mechanisms linking authentic leadership with follower creativity. Further, by establishing sequential mediation effects of meaningfulness and work engagement, the study provided additional insights on complex dynamics of the relationship among study variables. In doing so, the study addresses the need for more research to illuminate the socio-psychological processes through which this new leadership style influences employee level outcomes (Cerne et al., 2013). The study makes a novel contribution by advancing the understanding of previously unexplored psychological mechanisms that explain the connection between authentic leadership and creativity.

By enriching the understanding of the motivational, social and psychological processes underlying the relationship between authentic leadership and creativity, the study also explicates the less understood nature of the relationship between leadership and creativity (Cerne et al., 2013) and advances the creativity literature by identifying the antecedent conditions for promoting creativity among employees. The study finally contributes to the leadership theory as the findings of the study feed into the stock of knowledge in the area of authentic leadership. It establishes the merit of authentic leadership over existing leadership styles in determining employees' engagement and creativity. This opens up the avenues for future research toward designing the interventions targeted at developing authentic leadership among executives which have not received much attention due to a lack of evidences in the support of positive outcomes of authentic leadership.

6. Practical implications
The study also carries significant implications for organizational managers. The findings provide important insights to organizational managers for developing a creative workforce. The results establish that authentic leadership fosters creativity by enhancing the perceptions of psychological meaningfulness and work engagement. Not only creativity, but authentic leadership will also help the organizations build a workforce which is vigorous, dedicated and absorbed in work. Due to mushrooming of the multinational organizations in India and all competing for the same talent, keeping employees engaged has figured as the biggest HR challenge facing companies. Our study advances better understanding of the drivers of work engagement. By establishing authentic leadership as the predictor of work engagement and creativity, the study provides an important tool to organizational manager for cultivating an engaged and creative workforce.

Since engagement and creativity are the keys to organizational excellence, the organizations can gain competitive edge in the market by promoting authenticity in leaders’ thoughts and actions. We recommend that the organizations may reap the benefits of a
creative and engaged workforce by selecting, nurturing and developing authentic leaders. Interventions directed at selecting authentic leaders may include focusing on self-esteem, integrity, self-monitoring and emotional intelligence as selection criteria as these traits have been reported to be significant determinants of authentic leadership emergence and effectiveness (Ilies et al., 2005). Interventions aimed at developing authentic leadership may include using coaches and mentors providing feedback to improve self-awareness, training on quality leader-member exchanges (Scandura and Graen, 1984), role modeling, etc. (Ilies et al., 2005). Further, efforts may be directed to enhance experienced psychological meaningfulness, which, in turn, may enhance work engagement and creativity. Perceptions of meaningfulness in work may be developed through job enrichment, improving the work role fit, facilitating rewarding interpersonal relationships and encouraging trustworthy and supportive relationships at work (May et al., 2004).

7. Limitations and scope for future research

It is important to take a note of limitations of this study which may provide directions for future research. First, cross-sectional research design limits our ability to make causal inferences. Therefore, experimental and longitudinal research design should be adopted in future research to establish causality. Second, though the possibility of common method bias was reduced by collecting data from multiple sources, i.e. the employees and their supervisors, caution should be exercised while interpreting the results. Third, the use of convenience sampling restricts the extent to which study findings can be generalized. Fourth, the relationship of authentic leadership with follower outcomes may not be assumed to be universal. Therefore, future research should look out for the contingencies of the relationships examined in the present study. For example, demographic and individual difference variables such as gender, personality and self-esteem can be investigated as moderators of the authentic leadership and creativity relationship. This may provide researchers with the valuable insights on boundary conditions of the relationship. Fifth, future research should utilize better measures of psychological safety to add confidence to the findings of present study. It is also recommended to replicate the proposed research model in different organizational and country contexts for deeper understanding of dynamics of the relationships. Though single industry focus in the present study did not result in significant organizational level differences eliminating the possibility of conducting multilevel study, multilevel studies should be undertaken in future by collecting data from different industries.

References


Wang, Z. and Xu, H. (2017), “When and for whom ethical leadership is more effective in eliciting work meaningfulness and positive attitudes: the moderating roles of core self-evaluation and perceived organizational support”, Journal of Business Ethics, available at: https://doi.org/10.1007/s10551-017-3563-x


About the authors
Dr Richa Chaudhary is Assistant Professor of Human Resource Management and Organizational Behaviour in the Department of Humanities and Social Sciences, IIT Patna. She received a PhD degree from the Department of Management Studies, IIT Roorkee in 2013. She has published a number research papers with international journals of repute like *Personnel Review, International Journal of Human Resource Management, Social Responsibility Journal, Evidence Based Human Resource Management, Education + Training, The Psychologist Manager, Team Performance Management*, etc. Her current research interests include work engagement, corporate social responsibility, authentic leadership, multilevel studies and green human resource management. Dr Richa Chaudhary is the corresponding author and can be contacted at richa.chaudhary18@gmail.com

Chinmay Panda is PhD Research Scholar in the Department of Humanities and Social sciences, IIT Patna. His research interests include authentic leadership, work engagement, creativity, psychological safety and meaningfulness.
The relationship between managerial entrenchment, earnings management and firm innovation

Mahdi Salehi
Ferdowsi University of Mashhad, Mashhad, Iran, and
Mahbubeh Mahmoudabadi and Mohammad Sadegh Adibian
Faculty of Economics and Business Administration, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

Purpose – The purpose of this paper is to evaluate the qualitative effect of corporate governance components, in the form of managerial entrenchment index, on earnings management and innovation.

Design/methodology/approach – In this study, the variable of managerial entrenchment, which includes the variables of management independence, dual role of management, management tenure, the board compensation and the board ownership percentage, was initially estimated through the exploratory factor analysis and its effect was evaluated on the dependent variables of the study using the test of multivariable regressions. Hence, a total of 103 listed companies on the Tehran Stock Exchange were selected and analyzed during 2011–2016. In this paper, the Jones model is used as the variable of accrued earnings management and for calculating the real earnings management, the models of abnormal operational cash flows, abnormal production costs and abnormal optional costs are employed. Moreover, the research and development cost to total costs ratio is used for calculating the innovation.

Findings – The results indicate a negative and significant relationship between managerial entrenchment and accrual-based earnings management; moreover, the entrenched managers are less likely to engage in manipulating the real activities accruals in Iran context. Furthermore, the findings show that there is a positive and significant relationship between managerial entrenchment and firm innovation.

Originality/value – What really sets this paper apart from other studies is that this research will make aware investors and stakeholders of this fact that managerial entrenchment will be a good way to diminish the manipulation of financial reporting and improve the corporate situation in emerging markets, particularly those bazaars facing with economic sanctions such as Iran. Undeniably, the study results will complete the knowledge gap between the developed economies and the emerging markets.

Keywords Innovation, Earnings management, Corporate governance, Managerial entrenchment

Paper type Research paper

1. Introduction

Currently, earnings management is one of the central issues in accounting studies. Since the amount of profit is one main factor of decision making for investors, from a behavioral perspective these studies have their own significance. The results of these projects showed that low volatility and profit sustainability are two signs of quality. Hence, investors invest in those projects more confidently that their profit trend is more stable. When companies in unpropitious economic status are under an increasing pressure, managers ask the accounting department to improve the benefit and change the information content. Accounting despite all its flexibilities does not seem to be able to provide useful information for management in such circumstances (Hope and Hope, 1996). Earnings management usually occurs through manipulating the discretionary accruals or the real activities. Real activity-based earnings management includes the manipulation of real operation of a business unit in order to distinguish the reported profit of a current period (Mitra et al., 2013). The manipulation of discretionary accruals bears some costs. When the earnings
management goes up, managers face the risk of investigation by auditors and lawmakers and are prone to litigation. The question raises here is that by the presence of agency problems, including information asymmetry and the ethical risk of investors and regulatory authorities, how we could be ensured of the quality of reported profit by the managers and in a broader sense the quality of the financial reporting?

The concept of “corporate governance” is proposed given the necessity of combating against the ominous phenomenon of earnings management (that sometimes the scope of which is kept out of sight in lower layers of accounting standards that even if being caught by the sharp eyes the auditors, we could reject them because they are working under the flag of accounting standards), aligning different spectrums of stakeholders, and directing the objectives of the organization by all stakeholders. Since corporate governance is actually concerned about the necessity of close monitoring of management, regular firm auditing, separating an economic firm from its ownership and finally protecting the rights of all investors and stakeholders, it is a response to the agency problem that is derived from the separation of management from ownership. In other words, longer lifespan of an organization is the ultimate goal of the corporate governance, in that it is the corporate governance that determines how and by whom the company should be governed and how the favorable trend of accountability and disclosure of information to stakeholders should be managed. Thus, we could say that companies with more appropriate corporate governance are less likely to be faced with the conflict of interests and its subsequent consequences (Ghodrati and Feizi, 2015). By establishing an efficient strategic system, we would be able to align the interests of managers and owners, to improve the operational performance of the firm, and to develop the companies. Effective corporate governance is a managerial mechanism for the firm resources (Gill and Shah, 2012).

On the other hand, corporate innovation contributes significantly to the creation, preservation and addition of firm value. However, investing in innovative activities relative to other investments with higher rate failure is known as a high-risk investment (Bhagat and Welch, 1995; Holmstrom, 1989). Since the costs of research and development (R&D) cannot turn into capital, investing in innovation could lessen the short-term accounting profit and lower the compensation-based accounting for managers. Therefore, risk aversion managers may not be interested enough in innovation investment. However, companies with no innovation would lose their market value and consequently endanger the long-term interests of their shareholders. Entrenchment or what is called management opportunism is the survival instinct of the managers. It seems that there are some methods for managers who like to increase their power, job security and payment. Few studies are conducted on the quality of corporate governance components in Iran (Hassa Yeganeh et al., 2009; Baratiyan and Salehi, 2013; Salehi et al., 2017; Salehi, Tarighi and Saldari, 2018). Hence, performing some studies to investigate the significance of the quality of corporate governance mechanisms in stock companies of Iran is the matter of the utmost importance. In most of the conducted studies in Iran on the relationship between corporate governance and earnings management, each component of corporate governance entered the models and its effect on firm variables was calculated, separately. In this paper, however, we tried to investigate the simultaneous effect of a set of corporate governance mechanisms, including the board structure (management tenure, the board independence, CEO duality) and some management motivational factors (percentage of share available to the board and the amount of the board compensation) in the form of contributing factors to management power, earnings management and innovation. So from now on, the set of these variables is referred to as managerial entrenchment.

Therefore, we expect that corporate governance mechanisms to be able to mitigate the agency problems and help the company’s primary goal, which is to increase the value of shareholders’ wealth. Corporate governance, above all, aims at the survival of the firm in the
long run and aims to protect the interests of shareholders against the management of the organizations. The most important point about Iran market is that the majority of Iranian firms had many financial problems due to Iran was faced with severe economic sanctions during the study period between 2011 and 2016 (Salehi, Tarighi and Safdari, 2018). Iran’s capital market has experienced strong fluctuations in recent years due to international sanctions, exchange rate change and economic downturn, so that the information published on the official website of Tehran Stock Exchange indicates that the total index of Tehran Stock Exchange in 2012 was 26,502 and it reached 89,532 in 2013. This means that the total index of Tehran Stock Exchange has grown by 237 percent during almost one year. At the end of 2014, it dropped by more than 45 percent to 61,532 units. These intense fluctuations show the need to pay attention to the general market conditions in the research, as the company’s performance is affected by these fluctuations. In such economic situation, it seems that companies try to find ways to get rid of these financial problems using the firm innovation. In fact, we are going to know if corporate governance mechanisms have been able to reduce the agency problems and improve the corporate innovation in Iran context. Undoubtedly, the research results will complete the knowledge gap between the developed economies and the emerging markets.

The remainder of this study proceeds as follow. The next section presents a theoretical framework, hypothesis development and a literature review. Section 3 provides the research methodology and outlines where data are obtained and the sample selection procedure. Section 4 describes the main results and statistical analyses. Finally, the last section provides the conclusion.

2. Theoretical principles, hypothesis development and literature

Prior to the development of capital markets, business firms were operated traditionally. These firms supplied their required capital through the properties of people and their entrepreneur relatives and the approach of founders was based on cooperation not investment. The factor of unlimited responsibility was a barrier to the partnership of individuals in investments. Joint-stock companies proposed from 1,855 upward. The advent of such companies is among the world’s major economic evolutions and the separation of ownership from control is one of the consequences of the phenomenon (Sajadi, 2009). The process of separating ownership from control takes place when stock ownership dispersion occurs due to the growth of a firm, the result of which is the decline of shareholders’ power and the incremental enhancement of managers’ authorities. The separation of ownership from control created the problem of manager accountability, as the agent of owners, and caused the shareholders not to influence the managerial section of the firm. This issue formed the basis of the agency theory (Sajadi, 2009). The agency relation based on the definition of Jensen and Meckling (1976) is a contract through which the employer or owner appoints an agent on their behalf and delegate the decision-making authority. Through the agency relation, it is assumed that each party is trying to maximize its own interests. According to the agency theory, the separation of ownership role from management leads to broker–agent demand, because it is likely that managers pursue their own interests even to the detriment of brokers (Mustapha and Che Ahmad, 2011). Corporate governance system is one of the available mechanisms to curtail the agency problem and information asymmetry between shareholders and managers in the capital markets (Shleifer and Vishny, 1986; Salehi, Tarighi and Safdari, 2018). This study evaluates the mutual effect of a set of corporate governance components and motivational factors on earnings management and firm innovation. Since the components and strategic corporate factors are analyzed in numerous studies, we are concerned about the relationship between these variables, in form of entrenchment indexes, and dependent variables of earnings management and innovative production of firms.
2.1 The relationship between managerial entrenchment and earnings management

When managers hold little equity and stockholders are too dispersed to take action against non-value maximization behavior, insiders might organize corporate actions to gain personal benefits, like shirking and privilege consumption (Farinha, 2003). When ownership and control are divided within a firm, agency costs arise. Nevertheless, agency costs drop if the ownership of the firm rises as executives are responsible for a larger cost of these shares. Conversely, given ownership to a manager within a company may convert into greater voting power which makes the manager’s workplace more secure. Henceforth, they gain protection against takeover dangers and the current managerial market. Banko et al. (2013) proposed two approaches concerning the effect of managerial entrenchment on earnings management. The first approach indicates that the entrenched managers are motivated by earnings management. Several studies are available on this issue. Zhao and Chen (2008) revealed that companies with entrenched managers (the metric for measuring entrenchment is the board dispersion) have fewer earnings management. Stein (1988) also stated that threats related to expropriation are great incentives for short-sighted managers. Since the entrenched managers obliged to lessen the threats related to expropriation, they concentrated on long-term strategic policies rather than short-term ones, like earnings management. In line with such an approach, Pugh et al. (1992) realized that managers adopted a long-term approach, like capital as well as R&D costs, to prevent from anti-takeover amendments. The second approach shows that the entrenched managers are actually more motivated for earnings management. The experimental evidence signifies that entrenched managers are those who have weak performance (Gompers et al., 2003; Bebchuk et al., 2009). Moreover, the topical literature shows that personal financial motivations exist among managers for higher income. For example, Healy (1985) and Holthausen et al. (1995) noticed that managers manipulate the benefit for compensation plans. Based on the strong evidence from Iran, Salehi, Tarighi and Safdari (2018) investigated the impact of corporate governance mechanisms on audit fees. They classified the board compensation into Delta (Board’ cash rewards) and Vega (Board’ non-cash rewards). The results showed that there is a positive association between audit fees and delta, but not Vega; this implies that a fee premium is linked to CEO Delta incentives. Their outcomes demonstrated that Iranian firms pay more audit fees when they give managers more rewards. Furthermore, they found that there is not a significant relationship between fees resulting from audit risk and Delta and Vega incentives of the board. Exactly inconsistent with agency theory, they realized that the independence of board members did not affect audit fees. Bergstresser and Philippon (2006) indicated a linear relationship between CEO motivations and the manipulation of discretionary accruals. These studies illustrate that personal motivations for all CEOs and lack of sufficient supervision enable the entrenched managers in search of personal interests to perform the earnings management more freely (Banko et al., 2013). Dechow et al. (1996) declared that earnings manipulation is systematically related to the weakness of internal and external supervisions and companies with earnings management are more likely to have managers who influence the board of directors, have CEOs with a dual role and have CEOs who are the firm founder. By evaluating the impact of personality traits of innovative managers on real earnings management, Kouaib and Jarboui (2016) perceived that the characteristics of managers of such firms have a positive relationship with the probability of their earnings management. Ali and Zhang (2015) assessed the changes of CEO motivation during their tenure for earnings management and found that the increasing earnings management is higher during the first years of tenure than the upcoming years because new managers are trying to show the favorable results and to influence the market perception of their competencies during the first years of tenure.
In contrast, Dechow and Sloan (1991) declared that within the last years of CEO tenure, managers cut the costs of R&D to raise the short-term benefits. When managers embark on aggressive earnings management, they would be faced with the increase of risk of meticulous investigation of auditors and legal authorities and are more vulnerable to legal claims. Desai et al. (2006) noticed that companies with profit restatement may incur some penalties like losing jobs or decline of job opportunities. Within a study entitled “The relationship between market return, corporate governance and earnings management,” Farooque et al. (2014) stated that the earnings management has a considerable adverse effect on the market return and in preventing the earnings management the board size has more impact than institutional ownership. Given the impact of managerial entrenchment on earnings management, we could formulate the following hypotheses:

**H1.** There is a significant relationship between managerial entrenchment and accrual-based earnings management.

**H2.** There is a significant relationship between managerial entrenchment and real earnings management.

### 2.2 The relationship between managerial entrenchment and firm innovation

R&D costs are known as a proxy for corporate innovation. The significance of R&D is to the point that companies classify into developed and undeveloped ones based on the ratio of R&D funds received from gross national income. R&D costs, in addition to growth incentives and economic development of the community, increase the benefit of the firms. Today, R&D activities are the basis of innovation and pave the way for new demands constantly (Varamesh et al., 2014). Several definitions proposed so far for innovation and opinion-leaders and scholars take into account the topic of innovation from different aspects. By reviewing these definitions, a clean-cut image will be formed in the reader’s mind about the issue. Schumpeter was the first who defined the innovation in form of a scientific concept. In fact, he attempted to realize the contributing factors to the economic growth of companies and in this way, he became familiar with the role and significance of innovation in the growth of communities. Schumpeter in his definition of innovation expressed that as a disturbing factor to the existing order and an economic balance. The term “creative destruction” signifies the same issue. According to this theory, innovation is defined in five significant dimensions, all of which are indicative of the process of product development (Swedberg, 2007). These five aspects of innovation consist of new products, new manufacturing methods, the establishment of new markets, achieving novel raw materials and modern organizational methods. It seems that our study includes a combined set of innovation different aspects because the aim of R&D costs of a company is to go along these lines to grow the company in the future. Markides (2000) considered the organizational innovation as a completely different competitive method, which appears by violating the game rules. Broadly, we could define R&D as a planned and basic exploration, which is used with the aim of acquiring new knowledge to create a product and a modern service and to improve the products and manufacturing processes, effectively (Lipczynski et al., 2005). Some studies indicated that managers are able to play a vital and active role in the innovative strategy of a firm (Dooley and O’Sullivan, 2003). Passing and operating the innovative projects may signify the risk-taking capacity of management (Hirshleifer et al., 2012). The results of Helmers et al. (2017) show a positive relationship between a committed board and firm innovation. Gua and Zhang (2017) carried out a study on the impacts of Sarbanes–Oxley Act on the innovation of major corporations and noticed that this Act causes the growth of innovation of major companies, especially in high-tech industries. Kamoto (2016) indicated that instead of the
agency problem, due to the investors’ intolerance of failure, the motivational incentives of
management for innovation will decline in general companies. In addition, management
ownership leads to the improvement of the range of firm innovation. In this paper, the
mutual relationship between firm innovation, corporate governance and dividend policy is
studied and the experimental results proposed a positive relationship between control
stock purchase by management and the range of innovation. Ben-David et al. (2013)
showed that highly confident managers are able to influence the firm decisions about
investment policies. Hirshleifer et al. (2012) perceived that companies with highly
confident managers invest more in innovative projects. Innovative projects are a type of
investment in intangible assets. Such projects are at the same time the most perilous and
profitable activities and are a kind of commitment for highly confident managers, so they
are potentially important. Such investments are dubious from two sides. First, the risk of
failure is high in these projects due to nature of such activities; second, the real
manipulation of these projects may develop and increase (Hirshleifer et al., 2012). Ferreira
et al. (2014) analyzed the management motivations about innovative projects in public and
private companies and showed that failure intolerance in public companies lowers the
innovative motivations. Sapra et al. (2014) found out that there is a U shape relationship
between the range of managerial innovation and expropriation costs.

According to the managerial entrenchment theory, it is expected from the growth of
entrenchment to debilitate the regulatory effects of external control and lead to a lower
investment in innovation, so the increase of managerial entrenchment could have an inverse
effect on investors’ value (Chakraborty et al., 2014). On the other hand, Beyer et al. (2012)
declare that it is probable that managers invest a little on R&D due to the risk of failure in
projects, like decrease of compensation or job loss. Managers, however, may invest at a
higher level on innovative projects because of higher compensation and position and more
power. They perceived that there is an inverse relationship between the degree of
managerial ownership and the costs of R&D. When entrenchment occurs, managers are not
afraid of the determining effect of risk of innovative projects on occupation and position and
are willing to invest in innovation at a higher level. According to Chakraborty et al. (2014),
we believe that by intensifying the managerial entrenchment and failure intolerance in
innovative projects, the range of investment will decline in innovative activities:

\[ H3. \text{ There is a negative and significant relationship between managerial entrenchment}
\text{ and innovation.} \]

3. Research method
In terms of objective, the present study is practical and in terms of data-analysis method, the
study is cross-sectional correlational. Since the study is about the relationship between
managerial entrenchment, earnings management and product innovation among companies
listed on Tehran Stock Exchange, the desired method for hypothesis testing is a
retrospective correlation, so the present study is a type of retrospective research. The
information about independent, dependent and control variables was collected from the
financial statements of the companies listed on TSE via reliable resources. The time range of
the study was (2011–2016) as long as six years.

3.1 Research population and statistical sample
The target population included all companies listed on TSE which involved in the
productive goods, during the period 2011–2016. The reason for choosing these types of
companies is that accessibility to financial information for these companies is more than
other companies. Also, due to the regulations and standards of the Tehran Stock Exchange,
information on the financial statements of these companies is more homogeneous. Common features of the firms to determine the population are as follows:

(1) According to the research time period (2011–2016), the company is listed on TSE before the year 2011 and its name is not removed from the companies mentioned by the end of 2016;

(2) the fiscal periods of companies should be finished at the end of the solar year in order to enhance the comparability and homogeneity of companies in terms of time period;

(3) the company should be continuously active during the research period and its shares have been traded, and there is no trading halt; and

(4) the type of the company activity is productive and thus investment companies, leasing, credit and financial institutions and banks are not included in the sample due to their different natures.

Taking account of the above conditions, a sample size of 103 companies from firms listed on the TSE has been selected. See Table I.

Looking at the details, as regards sample industry distribution, we can find that our sample includes 618 firm-year observations that represent 21 industries and spans the years 2011–2016. In addition, it was found that the groups such as Production of metal products, Computer-related facilities and services, Agriculture and related services and Extraction of oil, gas and other services except for exploration have the lowest and the group of Pharmacy has the highest number of observation in our statistical sample.

3.2 Research models
3.2.1 The first research model. To test the research hypotheses, multivariate regression and data panel method have been used in this study. In the first research model, our purpose is to investigate the association between managerial entrenchment and accrual-based earnings

<table>
<thead>
<tr>
<th>Industry name</th>
<th>Firm-year obsv.</th>
<th>% of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacy</td>
<td>78</td>
<td>12.62</td>
</tr>
<tr>
<td>Automotive and the manufacture of automotive parts</td>
<td>72</td>
<td>11.65</td>
</tr>
<tr>
<td>Cement, lime and plaster</td>
<td>66</td>
<td>10.67</td>
</tr>
<tr>
<td>Other non-metallic mineral products</td>
<td>54</td>
<td>8.73</td>
</tr>
<tr>
<td>Basic metals</td>
<td>48</td>
<td>7.78</td>
</tr>
<tr>
<td>Chemical products</td>
<td>48</td>
<td>7.76</td>
</tr>
<tr>
<td>Food and beverage products except for sugar</td>
<td>42</td>
<td>6.79</td>
</tr>
<tr>
<td>Machinery and appliances</td>
<td>36</td>
<td>5.82</td>
</tr>
<tr>
<td>Rubber and plastic</td>
<td>30</td>
<td>4.85</td>
</tr>
<tr>
<td>Extraction of metal ores</td>
<td>30</td>
<td>4.85</td>
</tr>
<tr>
<td>Electric machines and appliances</td>
<td>24</td>
<td>3.88</td>
</tr>
<tr>
<td>Textiles</td>
<td>12</td>
<td>1.94</td>
</tr>
<tr>
<td>Transportation, warehousing, and communications</td>
<td>12</td>
<td>1.94</td>
</tr>
<tr>
<td>Sugar</td>
<td>12</td>
<td>1.94</td>
</tr>
<tr>
<td>Tile and ceramic</td>
<td>12</td>
<td>1.94</td>
</tr>
<tr>
<td>Petroleum products, coke and nuclear fuel</td>
<td>12</td>
<td>1.94</td>
</tr>
<tr>
<td>Production of metal products</td>
<td>6</td>
<td>0.97</td>
</tr>
<tr>
<td>Computer-related facilities and services</td>
<td>6</td>
<td>0.97</td>
</tr>
<tr>
<td>Technical services</td>
<td>6</td>
<td>0.97</td>
</tr>
<tr>
<td>Agriculture and related services</td>
<td>6</td>
<td>0.97</td>
</tr>
<tr>
<td>Extraction of oil, gas and other services except for exploration</td>
<td>6</td>
<td>0.97</td>
</tr>
<tr>
<td>Total</td>
<td>618</td>
<td>≈100</td>
</tr>
</tbody>
</table>

Table I. Firm-year observations distributed across the industry sectors
management among Iranian firms. In order to test the first hypothesis, the following regression model was used. In the below model, Accrual Earning Management (AEM) is defined as a dependent variable. In addition, we consider the Entrenchment variable as an independent variable and the rest of the variables are variable controls.

Model 1:

\[
\text{Accrual Earning Management}_{it+1} = \beta_1 + \beta_2 \text{Entrenchment}_{it} + \beta_3 \text{Institutional Ownership}_{it} + \beta_4 \text{Ownership Concentration}_{it} + \beta_5 \text{Debt}_{it} + \beta_6 \text{Dividend}_{it} + \beta_7 \text{Size}_{it} + \beta_8 \text{ROA}_{it} + \beta_9 \text{Investment}_{it} + \epsilon_{it}.
\]

Dependent variable. It should be noted that one of the most important issues regarding earnings management is the methods of discovery of profit management in business units, the most important of which is the method of discovery of accrual-based earnings management. Discretionary accruals indicate the difference between accounting profit and its related cash. The general interpretation of accruals is that accrued accruals are subject to the management’s perception of financial events. The Generally Accepted Accounting Principles have given companies freedom of action relative to the time of identification and the amount of revenue and expense. In fact, when managers identify the accounting profit as an amount more than the resultant cash, the discretionary accruals will be created (Piri and Tobreh Rizi, 2014). The Jones model is one of the accepted models in this field, which is formulated as follows:

\[
\frac{\text{TA}_{it}}{A_{it-1}} = a_0 + a_1 \frac{1}{A_{it-1}} + a_2 \frac{\Delta S_{it}}{A_{it-1}} + a_3 \frac{\text{PPE}_{it}}{A_{it-1}} + \epsilon_{it},
\]

where TA is total discretionary accruals (the difference between operational profit and operating cash flow); \(A_{it-1}\) the total assets at the beginning of \(t\) period; \(\Delta S\) the sales changes; and PPE is the property, plant and & equipment value.

Independent variable. In this study, managerial entrenchment is the independent variable. By emphasizing on managerial measurements which influence the interests of shareholders, it is one of the most expensive factors that include several structures, among which we could refer to the following items: issues related to limited-voting shares, golden parachutes, earnings smoothing, poison pills, major amendments, etc. Various methods are proposed to analyze the entrenched behavior of the CEO in resources, the most salient of which is the entrenchment index introduced by Bebchuk et al. (2009). This index, known as the E index in references, is the total of six variables, four of which limit the voting power of the shareholders. The other two variables, named poison pills and Golden parachutes, are indicative of anti-takeover actions of the management. Lin and Liu (2013) used the CEO tenure as the entrenchment behavior index. To measure the CEO entrenchment, Lin et al. (2014) used the main factor analysis based on the following six characteristics of the CEO:

1. Shares available to the CEO: CEOs with a higher proportion of shares have more control over firms and this could enhance their capabilities for following their personal interests. Thus, the higher the proportion of shares available to the CEO, the more probable is the emergence of entrenchment phenomenon.

2. CEO duality: CEO duality could weaken the board independence and its related performance, which in turn increases the CEO power. If the manager is the board director at the same time, the range of entrenchment goes up.
(3) Payment: the higher proportion of compensation to the total payment to the manager is indicative of the lower level of CEO entrenchment.

(4) CEO compensation: the higher proportion of compensation to the total payment to the manager is indicative of the lower level of CEO entrenchment.

(5) Purchase option for the manager: the higher proportion of compensation to the total payment to the manager is indicative of the lower level of CEO entrenchment. On the other hand, the stock option may be an incentive to the manager.

(6) CEO tenure: CEOs with longer tenure are more entrenched because they may establish some subgroups and debilitate the quality of internal control.

In this research, similar to Lin et al. (2014), Florackis and Ozkan (2009) and also based on the information of the Iranian capital market, we combine a set of corporate governance mechanisms that are probably related to the interest and capability of management to influence the interests of shareholders and evaluate them in the form of managerial entrenchment index. In this paper, we use the explanatory factor analysis (in the form of principal component analysis) to calculate the managerial entrenchment variable. In multivariate statistics, exploratory factor analysis (EFA) is a statistical method used to uncover the underlying structure of a relatively large set of variables. EFA is a technique within factor analysis whose overarching goal is to identify the underlying relationships between measured variables (Norris and Lecavalier, 2010). Mainly, this statistical method is used for two reasons; first, the explanatory factor analysis enables us to combine an extensive set of corporate governance variables to create a managerial entrenchment proxy. However, in the previous studies, a limited set of corporate governance factors were considered as the managerial entrenchment or ignored the multilinearity problem, which could be due to the existence of several corporate governance variables as control or independent variable in the experimental models. On the other hand, controlling the mutually potential relations of the variable is not an easy task. Second, one of the characteristics of explanatory factor analysis is that each variable included in managerial entrenchment has a weight based on the output of correlation coefficient matrix and this method is in contrast with the previous studies, in which the effect of each factor of corporate governance is equal.

Control variables. According to the review of the text, in order to control the impact of other variables that somehow affect research analysis, the variables of institutional ownership, ownership concentration, debt, dividend, company size, ROA and investment will be controlled. Institutional ownership is the total stocks available to the legal entities of firms as the percentage of institutional ownership. Ownership concentration is the number of stocks in the possession of major shareholders, who are among the main shareholders in the firm’s financial statement reports or the report of the board to the assembly and possess more than 5 percent of the firm stocks (Hassas Yeganeh et al., 2009). In this paper, total ownership percentage of three major shareholders who possess more than 5 percent of the firm stock is called ownership concentration. The debt ratio is indicative of the amount of debts in the capital structure of the firm, which is achieved by dividing total debts to total assets. The following equation is used to achieve the paid dividend: Dividend = DPS/EPS. Firm size equals the total sale of a firm; in fact, in order to measure the size of the company and also the homogeneity of the data, the logarithm of the company’s total sales is calculated. ROA shows information about return on assets. Finally, investment variable is the ratio of the fixed assets to total assets. What is worth mentioning is that control variables are the same for all study models.

3.2.2 The second, third and fourth research models. In this paper, we are going to know if there is a significant relationship between managerial entrenchment and Real Earnings
Management using three measures of abnormal operating cash flow, abnormal production costs and abnormal optional costs. In order to achieve this goal, the research model is defined as follows.

Models 2–4:

\[
\text{Real Earning Management}_{it+1} = \beta_1 + \beta_2 \text{Entrenchment}_{it} + \beta_3 \text{Institutional Ownership}_{it} \\
+ \beta_4 \text{Ownership Concentration}_{it} + \beta_5 \text{Debt}_{it} \\
+ \beta_6 \text{Dividend}_{it} + \beta_7 \text{Size}_{it} + \beta_8 \text{ROA}_{it} + \beta_9 \text{Investment}_{it} + \epsilon_{it},
\]

With respect to the second model, REM based on CFO is defined as a dependent variable. In this model, the purpose of the second hypothesis is to investigate the effect of real earnings management on real earnings management using the abnormal operating cash flow measure. In this regard, the abnormal operating cash flow is as follows:

\[
\text{CFO}_t = a + b_1 \text{Sales}_t + b_2 \Delta \text{Sales}_t + \epsilon_t,
\]

where CFO is operational cash flow ratio to the assets at the beginning of the period; Sales are the sales to the assets ratio at the beginning of period ratio; and \(\Delta\)Sales are sales changes ratio to the assets at the beginning of period ratio.

In the third model, the research was aimed at examining the second hypothesis of the study in terms of abnormal production costs:

\[
\text{PROD}_t = a + b_1 \text{Sales}_t + b_2 \Delta \text{Sales}_t + b_2 \Delta \text{Sales}_{t-1} + \epsilon_t,
\]

where PROD is the total ratio of production costs (which is equal to the final price of goods sold plus the change of the inventory) to the assets at the beginning of period ratio; Sales are the ratio of sales of the assets at the beginning of the period; and \(\Delta\)Sales describe sales changes ratio to the assets at the beginning of period ratio.

Finally, we considered that the third hypothesis of this study was examined in terms of optional abnormal costs:

\[
\text{DISCEXP}_t = a + b_1 \text{Sales}_{t-1} + \epsilon_t,
\]

where DISCEXP is general, R&D and advertisement cost ratio to the assets at the beginning of the period; and Sales are the ratio of sales to the assets at the beginning of the period.

3.2.3 The fifth research model. In the last study model, we have a tendency to know if managerial entrenchment affects firm innovation. So as to reach this goal, we designed this model as follows.

Model 5:

\[
\text{Innovation}_{it+1} = \beta_1 + \beta_2 \text{Entrenchment}_{it} + \beta_3 \text{Institutional Ownership}_{it} \\
+ \beta_4 \text{Ownership Concentration}_{it} + \beta_5 \text{Debt}_{it} + \beta_6 \text{Dividend}_{it} \\
+ \beta_7 \text{Size}_{it} + \beta_8 \text{ROA}_{it} + \beta_9 \text{Investment}_{it} + \epsilon_{it}.
\]

Here, innovation variable is identified as a dependent variable. As a matter of fact, the innovation of a company is calculated based on the ratio of the R&D costs to total company costs.
4. Results

4.1 Descriptive statistics

Descriptive statistics are used to describe the basic features of the data in a research (Salehi, Tarighi and Safdari, 2018). To evaluate the data, the descriptive statistics including minimum, maximum, mean, median and standard deviation are calculated and presented in Table II.

The results of Table II are extracted from 103 listed companies on the Tehran Stock Exchange during six financial years. According to the results of descriptive statistics related to the innovation variable and the closeness of the descriptive index to zero (mean, median and Standard deviation of 0.002, 0 and 0.012, respectively), it can be deduced that in general innovation in most Iranian companies is negligible and it is close to zero. In addition, regarding the control variables of this study, it can be stated that on average, three major shareholders owned 73 percent of the shares of the companies under study, in other words, the concentration of ownership in the investigated companies was high and 71 percent of the shareholders of the companies were institutional ones. On average, 82 percent of companies disclosed their intangible assets. In the following, the descriptive statistics related to management entrenchment components are depicted in Table III.

The results of Table III showed that on average, the board holds 63 percent of the shares of their managed companies, and 67 percent of the board are non-executive members. In general, CEO duality does not exist in most companies, and more than half of the companies have a CEO tenure over three years.

4.2 Kaiser–Meyer–Olkin (KMO) test

The KMO measure of sampling adequacy is a statistic that indicates the proportion of variance in research variables that might be caused by underlying factors. High values (close to 1.0) generally indicate that a factor analysis may be helpful with your data. If the

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEM</td>
<td>0.1125</td>
<td>0.0827</td>
<td>0.1072</td>
<td>0.0000</td>
<td>0.8585</td>
</tr>
<tr>
<td>REM1</td>
<td>0.1236</td>
<td>0.0828</td>
<td>0.1355</td>
<td>0.0000</td>
<td>1.5338</td>
</tr>
<tr>
<td>REM2</td>
<td>0.1399</td>
<td>0.0989</td>
<td>0.1346</td>
<td>0.0000</td>
<td>1.3138</td>
</tr>
<tr>
<td>REM3</td>
<td>0.0464</td>
<td>0.0381</td>
<td>0.0480</td>
<td>2.04E-05</td>
<td>0.732107</td>
</tr>
<tr>
<td>Innovation</td>
<td>0.0026</td>
<td>0.0000</td>
<td>0.0127</td>
<td>0.0000</td>
<td>0.1389</td>
</tr>
<tr>
<td>Debt</td>
<td>0.6005</td>
<td>0.6040</td>
<td>0.4526</td>
<td>0.0127</td>
<td>10.4133</td>
</tr>
<tr>
<td>Dividend</td>
<td>0.5887</td>
<td>0.6884</td>
<td>0.6883</td>
<td>0.0000</td>
<td>11.1354</td>
</tr>
<tr>
<td>Entrenchment</td>
<td>0.0738</td>
<td>0.1074</td>
<td>0.9600</td>
<td>-2.9806</td>
<td>2.0737</td>
</tr>
<tr>
<td>Institutional Own</td>
<td>72.7375</td>
<td>81.3050</td>
<td>24.9314</td>
<td>0.0000</td>
<td>100.0000</td>
</tr>
<tr>
<td>Investment</td>
<td>0.2482</td>
<td>0.2124</td>
<td>0.2006</td>
<td>0.0006</td>
<td>2.4797</td>
</tr>
<tr>
<td>Ownership Con</td>
<td>70.8238</td>
<td>73.9000</td>
<td>19.6077</td>
<td>0.0000</td>
<td>99.9950</td>
</tr>
<tr>
<td>ROA</td>
<td>0.1486</td>
<td>0.1321</td>
<td>0.1611</td>
<td>-0.7896</td>
<td>1.0451</td>
</tr>
<tr>
<td>SIZE</td>
<td>13.6633</td>
<td>13.5020</td>
<td>1.7588</td>
<td>7.1017</td>
<td>19.2816</td>
</tr>
</tbody>
</table>

Table II. Descriptive statistics of research variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duality</td>
<td>0.010</td>
<td>0</td>
<td>0.101</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Independence</td>
<td>0.669</td>
<td>0.8</td>
<td>0.213</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.514</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Share</td>
<td>62.990</td>
<td>69.295</td>
<td>25.883</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Compensation</td>
<td>7.011</td>
<td>7.09</td>
<td>1.946</td>
<td>2.944</td>
<td>9.384</td>
</tr>
</tbody>
</table>

Table III. Descriptive statistics of management entrenchment components
value is less than 0.50, the results of the factor analysis probably would not be very useful. In our study, the value of KMO test statistic is 0.59, which means we reasonably can use the factor analysis.

4.3 Factor loadings of exploratory analysis

Factor loadings are numerical values that indicate the strength and direction of a factor on a measured variable. Factor loadings indicate how strongly the factor influences the measured variable. In order to label the factors in the model, researchers should examine the factor pattern to see which items load highly on which factors and then determine what those items have in common (Fabrigar et al., 1999). Whatever the items have in common will indicate the meaning of the factor. In order to calculate the managerial entrenchment variable, EFA based on the matrix of correlation coefficient among five variables, the percentage of board shares, board independence, board tenure, board compensation and the board duality, is performed annually. The new variable called management entrenchment is estimated using the linear combination of five variables of corporate governance based on the following formula:

\[ F_j = \sum W_j X_i = W_{j1}X_1 + W_{j2}X_2 + \cdots + W_{jp}X_p, \]

where \( W \) denotes factor load coefficients and \( P \) represents the number of variables. Generally speaking, the results of factor loadings are presented in Table IV.

According to the above results, for example, in 2011, tenure, independence, duality and compensation of the board have a positive relationship with management entrenchment. Among the four variables mentioned above, the compensation of the board has the greatest impact, and the share owned by the board of directors has a negative relationship with management entrenchment.

4.4 F-Limer test

Apparently in accounting studies, when data are collected for several firms over a specific time period, in this case, we are faced with longitudinal data (pooled or panel). Hence, when data are longitudinal, the type of estimation of a model must first be determined. Based on econometric science, first of all, it is necessary to specify whether the model is fitted to the ordinary least squares (OLS) or panel data method. The F-Limer test (Chow) is used for reaching this purpose. In this test, the non-acceptance of the null hypothesis means that the model must be estimated with a panel data pattern and OLS model otherwise (Salehi, Tarighi and Safdari, 2018).

According to the results of Table V, it can be concluded that since the probability value of the \( H_0 \) test that is less than 0.05 for all research models, the preference of the OLS method is rejected, while the panel data method is accepted.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Duality</td>
<td>0.024</td>
<td>0.006</td>
<td>0.021</td>
<td>0.001</td>
<td>0.106</td>
<td>0.000</td>
</tr>
<tr>
<td>Independence</td>
<td>0.160</td>
<td>0.254</td>
<td>0.216</td>
<td>0.997</td>
<td>0.376</td>
<td>0.334</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.024</td>
<td>0.000</td>
<td>0.102</td>
<td>0.005</td>
<td>0.084</td>
<td>0.001</td>
</tr>
<tr>
<td>Share</td>
<td>−0.196</td>
<td>−0.173</td>
<td>−0.144</td>
<td>−0.078</td>
<td>0.093</td>
<td>0.240</td>
</tr>
<tr>
<td>Compensation</td>
<td>0.997</td>
<td>0.997</td>
<td>0.997</td>
<td>0.159</td>
<td>0.616</td>
<td>0.997</td>
</tr>
</tbody>
</table>

Table IV. Exploratory analysis factor loadings of management entrenchment variable
4.5 Hausman test

Following confirming the use of the panel data method in all research models, the Hausman test is used to determine whether a panel data with fixed effects should be used or a panel data with random effect (Salehi, Tarighi and Safdari, 2018). The Hausman test is an important factor in identifying the presence or absence of correlation between the error of regression and independent variables. Random effects model is used if such a relationship exists (the acceptance of $H_0$), and if it does not, fixed effects model is be used.

What stands out from Table VI is that since the probability value of $H_0$ is less than 0.05 for the first four research models, the preference of the fixed effects model is accepted and the random effects model is rejected. However, the results of the last model went in opposite direction. Hence, the preference of the random effects model is accepted for the fifth study models.

4.6 Variance inflation factor (VIF)

In statistics, the VIF measures the severity of multicollinearity in an OLS regression analysis. It provides an index that measures how much the variance of an estimated regression coefficient is increased due to collinearity (Salehi, Tahervafaei and Tarighi, 2018). If the linearity is high within a regression equation, this means that there is a high correlation between the independent variable and it is probable that the model lacks high reliability due to the high coefficient of determination. The VIF test is used for the test of linearity. Since the value of calculated VIF in this study is less than 10, there is no linearity among the independent variables. In addition, we annualized the two-by-two correlation of variables and found no high correlation which shows there is no linearity among the variables.

4.7 The results of research models

After performing various statistical tests and identifying its results, the findings of the hypotheses of this research are shown in Table VII.

It is necessary to test the significance of the model before variables examination, approval or rejection of the hypothesis. This can be done by calculating the $F$-statistic and $p$-value of this statistic. Since $p$-value calculated for this statistic is less than 0.05, the

<table>
<thead>
<tr>
<th>Model</th>
<th>Null hypothesis</th>
<th>Statistic</th>
<th>$p$-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Preferred OLS</td>
<td>1.95</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
<tr>
<td>Model 2</td>
<td>Preferred OLS</td>
<td>2.21</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
<tr>
<td>Model 3</td>
<td>Preferred OLS</td>
<td>1.98</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
<tr>
<td>Model 4</td>
<td>Preferred OLS</td>
<td>3.36</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
<tr>
<td>Model 5</td>
<td>Preferred OLS</td>
<td>3.19</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
</tbody>
</table>

Table V. $F$-Limer test

<table>
<thead>
<tr>
<th>Model</th>
<th>Null hypothesis</th>
<th>Statistic</th>
<th>$p$-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Preferred random effects model</td>
<td>27.95</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
<tr>
<td>Model 2</td>
<td>Preferred random effects model</td>
<td>61.08</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
<tr>
<td>Model 3</td>
<td>Preferred random effects model</td>
<td>42.32</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
<tr>
<td>Model 4</td>
<td>Preferred random effects model</td>
<td>70.82</td>
<td>0.000</td>
<td>$H_0$ denial</td>
</tr>
<tr>
<td>Model 5</td>
<td>Preferred random effects model</td>
<td>4.32</td>
<td>0.823</td>
<td>$H_0$ support</td>
</tr>
</tbody>
</table>

Table VI. Hausman test
significance of all models can be confirmed at the five-percent error level. According to the result, the high value of $R^2$ shows all our models will fit better our data. One of the important assumptions in the classical regression model states that the regression residuals should have the lack of serial autocorrelation. Since the amount of Durbin–Watson state for all models is between 1.5 and 2.5, this provides strong evidence of the lack of serial autocorrelation in the residuals.

The first hypothesis deals with investigating the relationship between managerial entrenchment and accrual-based earnings management. Given the obtained results of the first study model, the level of possibility of managerial entrenchment is less than 5 percent. Consequently, $H_0$, namely the insignificance of the obtained coefficient is rejected and $H1$ is accepted and the obtained coefficient is statistically significant. The results of this hypothesis showed a negative and significant connection between managerial entrenchment and accrual-based earnings management. Regarding the control variables of this model, it can be mentioned that the variables of Investment, firm size and ROA are positively connected with managerial entrenchment; however, the relationship between the Dividends variable and dependent variables is significantly negative.

With respect to the second hypothesis, we decided to evaluate the effect of managerial entrenchment on real earnings management using the different measures of abnormal operating cash flow (CFO), abnormal production costs (PROD) and abnormal optional costs (DISCEXP) by three separate models. The results suggested that there is a negative and significant association between managerial entrenchment and Real Earnings Management (REM) in Iran market. The outcomes also indicated that Real Earnings Management is high when the amount of debt decreases; in addition, companies with a better financial position are more likely to engage in real earnings management activities. We also found strong evidence that firms in which the ratio of the fixed assets to total assets is high have less tendency to involve in real earnings management behavior. Finally, according to the outcomes of the fifth model, we realized that the probability level of managerial entrenchment is less than 5 percent, which means the relationship between managerial entrenchment and firm innovation is significant. The positive sign of managerial entrenchment coefficient signifies a positive and significant relationship between these two variables.
5. Discussion
Consistent with agency theory, different corporate governance mechanisms in the form of managerial entrenchment led to reducing the manipulation of financial reporting. In this regard, our findings will warn investors and stakeholders that managerial entrenchment might be the best way in decreasing the earnings management and the agency problems in emerging economies, especially those markets struggling with financial sanctions like Iran.

Inconsistent with the managerial entrenchment theory suggesting the growth of entrenchment leads to a lower investment in innovation, and managers may have little investment in R&D activities due to the risk associated with project failures (Chakraborty et al., 2014; Beyer et al., 2012), the results showed that the managerial entrenchment contribute to the firm value when the firms are full of financial problems. As previously mentioned, the majority of Iranian firms experienced financial distress because of severe economic sanctions during the study period. In such economic environment, our findings showed that the manager who has a higher proportion of shares, the manager who is the board director at the same time and those with longer tenure have a better power to improve the financial situation of a firm. In other words, the entrenched managers were not afraid of the determining effect of risk of innovative projects on occupation and position and had tendency to invest in innovation at a higher level in order to get rid of financial problems.

It is strongly recommended that creditors, analysts, investors and other stakeholders to place more emphasis on the quality of corporate governance and the characteristics of the board of directors in their economic decisions. Considering the importance of R&D costs and the possibility of manipulating them, it is suggested to the authorities of the stock exchanges in developing countries to provide the mandatory requirements for better and more accurate disclosure of such costs for investors to examine.

What is worth mentioning is that researchers always have limitations in their research. Of course, it should be noted that research constraints do not mean the research failure in the stages of development, implementation, analysis and explanation of the results. As with other studies, this research has encountered some limitations:

- The limitation of access to financial information of companies led to the use of public corporations listed to Tehran Stock Exchange for research purposes. Therefore, research results for other companies will not be generalized.
- Financial statements are not subject to moderation due to the effects of inflation, and this may affect the results of the research.
- Many companies prevent disclosure information about salary, management-owned shares and CEO compensation as the highest executive, so there is no possibility to carefully check the management entrenchment index for all firms listed on TSE.

6. Conclusion
Due to the importance of the existence of the necessary mechanisms and the desirable quality of profit, most societies have made efforts to provide the appropriate corporate governance system. In recent decades, there have been different theories about the relationship between corporate governance mechanisms and profit management, and various empirical research studies have been carried out on these theories that led to the confirmation or rejection of the aforementioned theories. This research also explores the relationship between the quality of corporate governance in the form of the “managerial entrenchment” and earnings management among the firms listed on Tehran Stock Exchange. The results of the study indicate a negative and significant relationship between managerial entrenchment and accrual-based earnings management. This means that the entrenched managers are less likely to engage in manipulating discretionary accruals in the
Iranian context. The second purpose of this study was to examine the link between managerial entrenchment and real earnings management using three different measures. In this paper, the measures of abnormal operating cash flow (CFO), abnormal production costs (PROD) and abnormal optional costs (DISCEXP) were used as a proxy for evaluating the real activities manipulation. The results witnessed that sanctuary managers are really interested in decreasing the level of activities related to real profit management in Iranian market. Hence, our findings are consistent with the studies of Pugh et al. (1992), Stein (1988), Gompers et al. (2003), Zhao and Chen (2008), Bebchuk et al. (2009) and Banko et al. (2013).

In the second step in this paper, our goal was investigating the influence of managerial entrenchment on firm innovation. The evidence experiences a positive association between management entrenchment and firm value. This implies that Iranian managers who have more executive power in the decision-making process of a company have been able to drive the corporate innovation. In keeping with this notion, Hirshleifer et al. (2012) believed that the provision and operation of innovative projects can be a sign of management risk-taking capacity. When entrenchment takes place, since management has the power to decide, it is not afraid of the determining effect of risk in innovative projects in that it is willing to invest in such projects. Kamoto (2016) indicated that there is a positive relationship between the purchase of control stock by the management and the range of innovation. On the other hand, Chakraborty et al. (2014) concluded that by the increase of entrenchment the regulative effects of external control are declined and this could cause a lower investment in innovation.

References


About the authors
Mahbubeh Mahmoudabadi obtained her MA in Accounting from the Ferdowsi University of Mashhad, Mashhad, Iran. Her research interest focuses on auditing, accounting and finance.

Mohammad Sadegh Adibian is doing his PhD in Economics at the Ferdowsi University of Mashhad, Mashhad, Iran. His research interest focuses on general economics and micro economics.
Proposing an improved economic value model for human resource valuation

Sasan Ghasemi, Arash Shahin and Ali Safari
Department of Management,
University of Isfahan, Isfahan, Iran

Abstract

Purpose – The purpose of this paper is to propose an improved economic value model for human resource valuation. For this purpose, the probability of promoting people to a higher position and the satisfaction coefficient of employees and customers have been computed by competency and Kano models.

Design/methodology/approach – In order to calculate the probability of promoting people to a higher position, competency model and questionnaire have been used. The satisfaction coefficient of employees and customers has been calculated as a factor influencing the survival of an individual in the organization by using the Kano model. For this purpose, two questionnaires have been designed; one was completed by the employees and the other by the customers. The proposed model has been examined in a consulting company in Iran.

Findings – The human resource value of the company under study has been estimated over 29bn rials (Iranian currency). The obtained results indicate that the proposed approach as an integrative monetary and nonmonetary measure can remove the limitations of the economic value model.

Practical implications – The proposed model helps organizations in managing their human capitals more effectively.

Originality/value – In this study, the Kano and competency models have been integrated with one of the common models of human resource valuation, i.e. the economic value model. The proposed integrated model seems more effective compared to the basic model of economic value model. Application of the proposed model within the context of Iran for the first time would constitute as potential for contribution to the knowledge of human resource management in the developing countries.

Keywords Human resource accounting, Kano model, Competency model, Economic value model, Human resource valuation

1. Introduction

There is no doubt that in the present era when the global economy has been rapidly changed from an industrial economy to a knowledge-based economy (Sahoo, 2016) as a result of the emergence of new technologies, including the internet, human capital has been of particular importance to the organizations (Monday, 2017). Therefore, those organizations can survive that are aware of the role of their human capital in the flexibility of organization to adapt to environmental conditions. Such organizations can apply appropriate management measures to use this valuable capital, effectively and efficiently. Human resource accounting is known as a new branch of accounting (Islam et al., 2013) by which, the required information is provided to managers and can help them in using resources more efficiently and effectively. In fact, human resource accounting attempts to determine the quantity of intangible requirements that individuals bring with themselves to the organization (Monday, 2017). Human resource accounting is not just an accounting system for determining the human resource value, but it is a thinking approach about human management in organizations (Rolla, 2017).

While many knowledge-based companies have been established and developed in Iran in recent years, currently those companies are facing critical challenges in managing their human resources. The most important competitive advantage of such companies is not their financial asset or even high price equipment, it is their knowledge workers that differentiate...
them from their competitors. One of the challenges is the lack of managers’ awareness about the importance of human resources compared to other organizational resources. While human resources have been developed in many Iranian organizations in recent years, the subject has not been recognized by senior managers sufficiently, compared to financial resources, which has been always considered as the first priority. Therefore, proposing an approach for determining the value of human resources based on economic value model can help managers in better recognition of human resources compared to other organizational resources. This in turn will help managers in making more accurate and effective decisions in managing their human resources.

With regard to human dynamics, the human resource value assessment method has been always one of the challenges for human resource managers and accounting specialists. Hence, in human resource accounting literature, numerous approaches and models have been offered for measuring the human resource value, each of which has its own specific limitations, which has made it impossible and difficult to offer a true and real valuation of the human resources. Some of the models and methods that have been developed so far for human resource accounting include Individual Value Indicators Model (Flameholtz model), economic value (present value), replacement cost, multiplier, competitive building, historical cost and Markov model.

In this study, an attempt has been made to add new variables in order to eliminate the limitations of the economic value model of Lev and Schwartz (1971). Among the limitations of the economic value model is that this model assumes that employees will work at the same previous position until the retirement age or death and will not be promoted. Also, in this model, the factors of leaving organization by individuals include merely death and retirement. Therefore, in the proposed model, it is tried to remove these two important limitations. For this purpose, the probability of leaving and dismissing employees are added to the model by using employee satisfaction coefficient ($\alpha$) and customer satisfaction coefficient ($\beta$). These coefficients are calculated using the Kano model. Another added variable is the probability of promoting an individual to a higher position, which is calculated based on the competency model.

The proposed model has been examined in a management consultancy company. The results of this study help organizations including management consultancy companies in managing their human resources more effectively and efficiently based on quantitative and accurate information.

In the following, human resource accounting, the Kano model and the competency model are briefly described and the proposed method for human resource valuation is proposed. Then, in a case study, the proposed model is examined. Finally, findings are discussed and theoretical and practical suggestions are addressed.

2. Human resource accounting

The view at human resource as a capital cost has been provided by accounting theorists for many years, since such costs develop the economic benefit of organization after a period of time (Stovall and Neill, 2017). In 1973, the American Committee of Human Resource Accounting defined the concept of human resource accounting as “a series of steps that helps identifying and providing data required and related to human resources and results in transferring this data to the stakeholders.” Human resource accounting is the process of identifying and measuring human resource information and reporting this information to the interested and stakeholder individuals. In fact, human resource accounting is the transformation of the qualitative and subjective concept of human resource value in a quantitative and objective form, by using accounting knowledge. It deals with three important and basic parts of human resources including identifying its quantity and quality, evaluating and measuring economic value and associated financial reporting (Rana and Maheshawari, 2005).
Researchers of human resource accounting, by using human capital-based theory, try to provide desirability patterns of human resource management (Bamberger et al., 2014). In Table I, some of the definitions of human resource accounting are addressed.

The major objectives of human resource accounting are to identify the human resource value, to measure the cost and value of individuals for the organization and to determine the cognitive and behavioral impact of such information on the organization (Islam et al., 2013). Human resource accounting determines three main duties for human resource professionals: it provides a framework for facilitating the right decision making about human resources, provides required information about the cost and value of human beings as the assets of organizations; and motivates the queue management with regard to human resource perspective in decision making about the individuals of organization (Flamholtz, 1998).

Almost all of the methods that have been proposed for human resource accounting so far have limitations (Monday, 2017). In this study, the economic value model has been considered as the base model of research and attempt is made to enter other variables into this model. The reason for regarding this model as the base is the study of Talebnia and Ghorbani (2010) who found that the economic value model is more acceptable than the other models. In Table II, common approaches used for the valuation of human resources are addressed and their weaknesses and strengths are explained.

In the following, the economic value approach is described.

3. Economic value model

The economic value model is based on the theory of capital and has been presented by Lev and Schwartz (1971). In order to measure the value of an individual in an organization, they considered expected benefits based on services provided by individuals. In this model, the human resource value is measured considering salary and benefits paid to the employees (Flamholtz, 2001). Based on this model, future revenues obtained from human resource services can be discounted and considered as human resource value. The following equation indicates the model of human resource valuation based on the economic value model:

$$E(V_i^*) = \sum_{t=1}^{T} P_t(t+1) \sum_{t=1}^{T} I_t \left( \frac{1}{1+r} \right)^{T-t},$$

where $E(V_i^*)$ is the expected value of an individual human capital; $P_t(t)$ is the probability of an individual’s death at age $t$; $t$ is the current age of an individual; $I_t$ is the annual income of the individual until retirement time; $r$ is the discount rate in each period; and $T$ is the retirement age.

<table>
<thead>
<tr>
<th>Author</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flamholtz (1973)</td>
<td>Measuring the cost and value of the organization individuals; the process of specifying and measuring data about human resources and reporting this information to the interested people and stakeholders</td>
</tr>
<tr>
<td>Flamholtz (1998)</td>
<td>Specifying, quantifying and reporting the abilities of the employees of organization</td>
</tr>
<tr>
<td>Flamholtz et al. (2002)</td>
<td>The accounting for individuals as human assets</td>
</tr>
<tr>
<td>Rahaman et al. (2013)</td>
<td>The process of cost measurement that organizations incur for the recruitment, selection, employment, training and development of their human assets</td>
</tr>
<tr>
<td>Dhanabhakyam and Mufliha (2016)</td>
<td>The provider of monetary data related to the human resources of organization, which in addition to being used by management in decision making, is also used by analysts and even employees</td>
</tr>
</tbody>
</table>

Table I. Some definitions of human resource accounting
4. Kano model
Since product requirements do not play equal role in customer satisfaction, identifying factors which determine satisfaction is essential for the success of any organization (Tontini and Silverira, 2007; Chen and Su, 2006). In 1980, Kano’s two-dimensional model became one of the most famous quality models (Kano et al., 1984). Kano model (Figure 1) has been used for recognizing specific requirements which potentially create customer satisfaction or dissatisfaction (Mikulic and Prebezac, 2011; Shahin and Mohammadi Shahiverdi, 2015; Shahin et al., 2017). Classified quality requirements of the Kano model are located in five groups including must-be, one-dimensional, indifferent, attractive and reverse. Such requirements have different impacts on customer satisfaction (Shahin, 2004; Shahin and Zairi, 2009; Shahin and Nekue, 2011; Shahin et al., 2013).

4.1 Must-be requirements
These are customer’s expected or “must” requirements, and are not regarded as a chance for product better performance. Increasing performance of these requirements decreases
customer dissatisfaction and its decrease results in customer intensive dissatisfaction. An example of these requirements is brake in automobiles.

4.2 One-dimensional requirements
As much as these requirements are increased, customer satisfaction will be higher. On the other hand, lack of these requirements results in customer dissatisfaction. A part of those requirements stated orally by customer are often one-dimensional requirements.

4.3 Attractive requirements
These requirements are unstated or unexpected requirements of customer, but might result in increasing customer excessive satisfaction. However, the absence of these requirements does not result in customer dissatisfaction. In competitive market the fact that product manufacturers offer equal performance, meeting attractive requirements that meet unknown requirements, creates a competitive advantage. After a while, these requirements usually change into one-dimensional requirements and then into basic (i.e. must-be) requirements.

4.4 Indifferent requirements
These are characteristics which are not considered by customers. Offering or not offering them is not important.

4.5 Reverse requirements
If these requirements are fulfilled, they will have reverse impact on customer satisfaction (Chen, 2012).
Another group of requirements is questionable requirements. Such requirements are related to the suspected and contradictory responses which need more investigation and are not entered into calculation and analysis.
Having used Kano questionnaire, the information concerning customers’ expectations of product desired requirements are collected and then they are evaluated by the Kano evaluation table (Table III).

With regard to the Kano evaluation table, and having specified the type of each requirement, the satisfaction and dissatisfaction coefficients are determined according to Equations (6) and (7). Satisfaction indicator (SI) implies to what extent fulfillment of the requirements impact on customer satisfaction, and dissatisfaction indicator implies to what extent customer is dissatisfied if the requirements are not fulfilled. These indicators are calculated as follows:

\[ SI = \frac{(A + O)}{(A + O + M + I)} \]

\[ DI = \frac{(A + O)}{(A + O + M + I)} \]

5. Competency model

Competency is a set of personal capabilities that enables a person to perform duties he is responsible for, successfully (Kessler, 2006). Competency descriptors include the identification of knowledge, skills, attitudes, values and judgment that are dynamic, developmental and evolutionary and represent a learning chain (Sahoo, 2016; Loufrani-Fedida and Saglietto, 2016). While the concept of competency has been widely used, competency has various meanings for various individuals, and currently there is also no single academic definition for it (Takey and Carvalho, 2015).

A competency is the related knowledge, skills and attitudes that impact on a major part of the individual’s occupation (including a role or responsibility), relevant to the performance in that occupation. It can be measured according to the accepted standards and can be developed and improved by practice and exercise (Hsieh et al., 2012). In front of job description that lists only the responsibilities or expectations of the job, competencies include many requirements and factors that affect job success. For example, being optimist or pessimist are attitudinal concepts which are not mentioned in job description (Hsieh et al., 2012).

A model of competency has been defined as an explanation and description of the specific behavior of skills and attributes needed to be effective in performing a job (Goldman and Scott, 2016). Today, competency models are widely used as the basis for talent.
The competency model in this study is that in the knowledge era, what should be considered in human resource accounting is the ratio of knowledge and ability of individuals (Monday, 2017). In this study, it has been tried to consider knowledge, skills and ability of individuals in calculating the human resource value using the competency model and determining the ratio of individual promotion.

6. Research methodology

As previously mentioned, in this study an attempt has been made to remove the limitations of the economic value model of Lev and Schwartz (1971). Among the limitations of economic value model is that in this model it is assumed that employees will be working in the same previous position until the end of retirement age or death and will not be promoted. Also, in this model only an individual’s death and retirement have been considered as the factors of an individual’s leaving the organization. Therefore, in the proposed model, it is tried to remove these two important limitations. For this purpose, the probability of leaving and dismissing employees has been added to the model by using employee satisfaction coefficients (α coefficient) and customer satisfaction coefficient (β coefficient). These coefficients are calculated using the Kano model. Another variable added to the model is the ratio of an individual’s promotion to a higher position, which is calculated based on the competency model. Therefore, the following equation is proposed for calculating the human resource value:

$$E(V_t) = \sum_{t=1}^{T_i} \frac{P(t+1) \sum_{i=1}^{n} I_i T_i}{(1+r)^{t-1}}.$$  

where $P(t)$ is the probability of a person’s death; $\alpha$ is the individual’s satisfaction with the job; $\beta$ is customer satisfaction with the organization; $I_i$ is the expected income of the person at age $t$ and at the job position of $i$; $T_i$ is the ratio of a person’s promotion to the position $i$; $r$ is the degradation rate per period; $t$ is the age of a person at any time; $r$ is the current age of a person; $n$ is the number of upgradeable positions for the individual or groups of people in their job path until the retirement age; and $T$ is the retirement age.

In order to calculate the human resource value of organization under study, each of the variables required for the proposed model is calculated according to Figure 2. The steps are described in the following.

1. Computing an individual’s ratio of promotion to the position $i$ based on the competency model ($T_i$)
2. Computing the ratios of employee job satisfaction and customer satisfaction based on the Kano model ($\alpha$ and $\beta$)
3. Computing the probability of an individual’s death at age $t$ ($P(t)$)
4. Computing the expected income of an individual at age $t$ and at the job position ($I_i$)
5. Valuating the human resources of company based on the proposed approach
6.1 Calculating the ratio of individual’s promotion to a higher position based on the competency model

The research population included 15 employees of a consultancy company. Purposive sampling (Holloway and Wheeler, 2010) has been used for data collection. This type of sampling is used when a unit of research is selected with a specific objective and is not selected randomly (Tashakkori and Teddlie, 2003). Therefore, eight employees of the company were introduced by its executive manager for valuating human resources. They had the key job positions and filled both of the questionnaires of competency and job satisfaction. In addition, 11 key customers of the company were selected in a similar way (using purposive sampling) to fill the customer satisfaction questionnaire. They were key customers, since they have been in direct communication with the company during the past three years.

In order to calculate the ratio of each individual’s promotion to a higher position, the competency model is used. The competency model of the organization under study consists of three main parts of professional competencies (competencies required to carry out professional duties of job), general competencies (including competencies that all people working in the organization require to carry out their job successfully) and management competencies (those that people who have some kind of managerial duties should have). Each of these main parts includes the criteria that a questionnaire is prepared for assessing the current competency of each individual. The questionnaire has a five-point spectrum and should be filled by the individual’s manager or supervisor, his/her colleagues and himself/herself. After completing the questionnaires, the average value of competency of each person is calculated. By comparing the competency level of each individual with the minimum value of competency required for a higher position, the ratio of an individual’s promotion to a higher position is calculated. For this purpose, the following equation is used:

\[
\text{Ratio of individual’s promotion} = \frac{\text{Current competency of individual}}{\text{Higher position competency}}.
\]

6.2 Calculating the satisfaction coefficient of employees and customers based on the Kano model

As previously mentioned, for human resource valuation based on the proposed approach, two satisfaction coefficients are needed. One is the employee satisfaction coefficient and another one is the customer satisfaction coefficient. For this purpose, two questionnaires are designed based on the Kano model, one of which is completed by the employees and another by the customers.

In the Kano model, to identify qualitative requirements, a simple method consisting of two positive and negative questions is used. To this end, for each product or service requirement, a pair of questions is designed that the customer can respond to by selecting one of the five options available. The first question (functional form) is related to the customer reaction if the product fulfills that requirement, and the second question (dysfunctional form) is related to the customer reaction in the absence of that requirement. Based on the answer given by the customer to the above two questions, the product requirements are classified into six categories: A (Attractive requirement), M (Must-be requirement), O (One-dimensional requirement), I (Indifferent requirement), R (Reverse requirement) and Q (Questionable requirement) (Table III).

Minnesota Job Satisfaction Questionnaire is used to compile job satisfaction questionnaire in two short and long forms. The short form introduces 20 effective requirements in job satisfaction. Considering that the Kano questionnaire has two questions
for each requirement, hence, the job satisfaction questionnaire consists of 40 questions based on the Kano model.

The Servqual (abbreviation of the words Service and Quality) questionnaire is used to measure customer satisfaction ratio. The questionnaire has been compiled by Parasuraman et al. (1985) and has 22 questions. In this study, due to the fact that the questions of the Kano model are twofold, 11 important requirements of the questionnaire are extracted and a 22-question questionnaire is designed. The customers of organization under study are asked to answer the questions of this questionnaire based on their experience of relationship with the organization.

Cronbach’s $\alpha$ is used for reliability and internal consistency analyses. SPSS 22 software is used for the analyses. The $\alpha$ values are determined as 0.820 and 0.898 for the data collected via the questionnaires of employee satisfaction and customer satisfaction, respectively. This is satisfactory, since the values are higher than the minimum acceptable value, i.e. 0.7 (Taber, 2017).

6.3 Calculating the probability of an individual’s death at age $t$

In order to determine the probability of an individual’s death, a mechanism is used that insurance companies employ. For this purpose, the mortality table provided by the Central Insurance Agency of Iran is used. Based on this table, the probability of each individual’s death can be determined at any age.

6.4 Calculating the expected future income of individuals

To calculate the expected future income of individuals, the past trend of annual salary increase is investigated. For this purpose, the percentage of minimum salary changes for workers subjected to the labor law over the past ten years is investigated so that the average value of changes is computed. The basic salary information of the Department of Labor over the past ten years is presented in Table IV.

Based on the results obtained from the trend analysis of past years, it is assumed that the salary base in future years would have an increase of 18 percent.

6.5 Calculating human resource value using the proposed approach

Finally, by using the calculated values for the variables in the proposed approach, the human resource value of the company under study is calculated. In the following, the proposed approach of human resource valuation is examined in a case study.

<table>
<thead>
<tr>
<th>Year</th>
<th>Minimum salary at the Department of Labor (in rials)</th>
<th>Percentage of changes compared to previous year</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>1,500,000</td>
<td>–</td>
</tr>
<tr>
<td>2007</td>
<td>1,830,000</td>
<td>22</td>
</tr>
<tr>
<td>2008</td>
<td>2,196,000</td>
<td>20</td>
</tr>
<tr>
<td>2009</td>
<td>2,635,200</td>
<td>20</td>
</tr>
<tr>
<td>2010</td>
<td>3,030,000</td>
<td>15</td>
</tr>
<tr>
<td>2011</td>
<td>3,300,000</td>
<td>9</td>
</tr>
<tr>
<td>2012</td>
<td>3,897,000</td>
<td>18</td>
</tr>
<tr>
<td>2013</td>
<td>4,871,250</td>
<td>25</td>
</tr>
<tr>
<td>2014</td>
<td>6,089,100</td>
<td>25</td>
</tr>
<tr>
<td>2015</td>
<td>7,124,250</td>
<td>17</td>
</tr>
<tr>
<td>2016</td>
<td>8,121,660</td>
<td>14</td>
</tr>
<tr>
<td>2017</td>
<td>9,298,310</td>
<td>15</td>
</tr>
</tbody>
</table>

Table IV. The percentage of changes in the minimum salary of Iran over the past 10 years
Average percentage of change 18
7. Case study and findings

As a case study, the proposed approach was implemented in a management consulting
company in Iran. In management consultancy companies, the value of knowledge-oriented
human resources is much higher than their physical capital and assets value. Therefore,
the use of a model that could perform human resource valuation is recognized as an
important need in this sector. On the other hand, such companies have a close relationship
with their customer and prefer to consider customer satisfaction in their human
resource evaluation. This important subject and the weaknesses of other models imply the
necessity of developing a more comprehensive and effective model for calculating
the value human resource and help the managers of such companies in making
appropriate policies and strategies.

In fact, four main criteria have been used for selecting the case company, based on the
study of Dawes et al. (1992). They included: popularity of the company in the country;
transaction size of the company; number and popularity of the company’s customers; and
providing a wide range of services to customers.

At the time of conducting this study, there were 15 employees activating in different
positions. According to the human resource experts of the company, eight employees had
the key and important positions of the company who were chosen for human resources
valuation. In addition to completing questionnaires related to the competency, they also
completed job satisfaction questionnaire. A total of 11 key customers of the company
who had direct contact with the company also completed the customer satisfaction
questionnaire. In order to investigate the validity of research questionnaire, the content
validity index (CVI) and content validity ratio (CVR) were used. To calculate these two
criteria, the views of 11 university professors were investigated. They had expertise in
management area, especially organizational behavior and human resource management.
To calculate the CVR index, they were asked to select an option out of a three-point
scale. The scale included options of “1: the item is necessary”; “2: the item is useful but not
necessary”; and “3: the item is not necessary.” Subsequently, the CVI of CVR was
calculated based on the average value of opinions obtained for each item and the
following equation:

$$\text{CVR} = \frac{n_c - n/2}{n/2}$$

(6)

In this equation, $n$ is the total number of specialists; and $n_c$ is the number of specialists who
have chosen the necessary option (Ruhl et al., 2015).

For calculating CVI, people were also asked to score the items based on Likert
four-point spectrum. It included options of “1: it is not relevant”; “2: it is relatively
relevant”; “3: it is relevant”; and “4: it is completely relevant.” Then, based on the average
value of opinions obtained for each item and the following equation, the CVI was
calculated as follows:

$$\text{CVI} = \frac{n_b}{n}$$

(7)

In this equation, $n$ is the total number of experts and $n_b$ is the number of experts who gave
scores of 3 and 4 to the item. Since the minimum acceptable values for the CVI and CVR
criteria for 11 experts are 0.62 and 0.59, it was concluded that according to the experts’
opinion, the questionnaire used had acceptable content validity (Ruhl et al., 2015).

In the following, the proposed approach of human resource valuation in the company
under study is described based on the steps outlined in the previous sections.
7.1 Computing the ratio of experts’ promotion to a higher position based on the competency model

As previously mentioned, eight people were selected as the key company experts. These experts were working as chief executive officer, marketing executive, marketing senior, human resource executive, human resource senior, management consulting executive, management consulting senior and management consulting assistant. To calculate the ratio of the experts’ promotion to a higher position by the use of Equation (1), the values obtained for the mentioned experts are addressed in Table V, regarding the four aspects of competency.

The minimum competency scores required for the chief executive officer, manager, senior expert and expert positions were considered as 5, 4, 3 and 2, respectively. Accordingly, and by using Equation (1), the ratio of each expert’s promotion to a higher position was calculated and addressed in Table VI.

7.2 Computing the satisfaction coefficient of employees and customers based on the Kano model

After completing the satisfaction questionnaire by the employees, responses were analyzed using the Kano evaluation table. By the use of this table, based on customers’ opinion (here employees), the type of presented product or service (here the job) requirement could be determined. After data analysis regarding the selected positions, the satisfaction and dissatisfaction coefficients were calculated for each individual, as presented in Table VII.

After customers completed the questionnaires, the responses were analyzed by forming the Kano evaluation table. It should be noted that for customers, the frequency of responses for each requirement was calculated as the total response of customers and therefore a satisfactory coefficient and a dissatisfaction coefficient were calculated based on all customers. The results are presented in Table VIII.

7.3 Computing the probability of an individual’s death at age t

By using the mortality table and based on the age of each expert of the organization under study, the probability of the death of individuals can be calculated. The obtained values are presented in Table IX. It should be noted, however, that due to the limited length of the paper, only information about the current age of individuals is provided, while in the proposed model, the probability of the death of individuals should be calculated annually and up to the retirement age. The computation method is similar to Table IX.

<table>
<thead>
<tr>
<th>Position</th>
<th>Management consulting assistant</th>
<th>Management consulting senior</th>
<th>Management consulting executive</th>
<th>Human resource executive</th>
<th>Human resource senior</th>
<th>Marketing senior</th>
<th>Marketing executive</th>
<th>Chief executive officer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management consulting competency</td>
<td>1.63</td>
<td>2.81</td>
<td>3.48</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3.59</td>
</tr>
<tr>
<td>Human resource competency</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.85</td>
<td>2.29</td>
<td>–</td>
<td>–</td>
<td>3.11</td>
</tr>
<tr>
<td>Marketing competency</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.25</td>
<td>3.05</td>
<td>3.24</td>
</tr>
<tr>
<td>General competency</td>
<td>2.45</td>
<td>3.29</td>
<td>3.63</td>
<td>2.96</td>
<td>3.51</td>
<td>2.31</td>
<td>3.79</td>
<td>3.74</td>
</tr>
<tr>
<td>Managerial competency</td>
<td>1.65</td>
<td>2.77</td>
<td>3.27</td>
<td>2.63</td>
<td>3.04</td>
<td>1.56</td>
<td>3.58</td>
<td>3.85</td>
</tr>
<tr>
<td>Average</td>
<td>1.91</td>
<td>2.96</td>
<td>3.46</td>
<td>2.48</td>
<td>2.95</td>
<td>1.71</td>
<td>3.48</td>
<td>3.51</td>
</tr>
</tbody>
</table>

Table V. Average values of competencies for each expert working in the selected key positions
7.4 Computing the expected future income of individuals

Based on the results obtained from the trend analysis of past years, it is assumed in this study that the salary base in the future will have an increase of 18 percent. This increase is usually performed proportional to the inflation rate in Iran, and is high because of the high inflation rate in Iran. The salary base ratio of each employee in the company is presented in Table X and it is assumed that 18 percent is added to the value, every year.
7.5 Computing the human resource value using the proposed approach

Based on the proposed approach and also based on the computations, the value of each expert of the company under study was calculated and presented in Table XI. The total value of the company's human resources was also calculated.

8. Discussion

The findings indicate that the current value of the human resources of the company under study is over 29bn rials. Therefore, it is concluded that a company's human capital's value might be higher than the value of its tangible assets. It is particularly true regarding the fact that the company is knowledge based and its most important capital is its human resources. It is important to note that knowing the value of a company's human resources can help decision makers in taking effective decisions, significantly. It seems that using the proposed integrated approach of competency and Kano models can contribute to an improvement in the existing models of human resource valuation.

Literature review indicated that performing and implementing human resource accounting in the organization can increase the performance and profitability of organization (Okpako et al., 2014). In addition, information obtained from human resource accounting helps managers in making better decisions in human resource management area, which in turn affects organizational performance, indirectly. The question that managers and human resource professionals have been dealing with is how this work should be done, in other words, what is the best method for human resource accounting? Although accounting has specific importance for the organization, the lack of a more effective method and also the weakness of approaches that have been proposed so far (Lakshmi Rao, 2014) have caused the authors of this paper to develop an improved approach. For this purpose, the limitations of one of the most prominent human resource accounting approaches, i.e. the economic value model have been removed by integrating the competency model and the Kano model.

The findings also indicated that the Kano and the competency models can strengthen the economic value model and turn it into a suitable technique for valuing human resources based on the monetary index. Providing human resource value as a monetary indicator provides an opportunity to the human resource experts and managers not only to investing in each human resource management system properly and understand the consequences of their decisions in this area based on clear and tangible indicators, but also to assess the risks of their decisions, accordingly. Of course, as it was stated, human resource accounting is not just a human resource valuation approach, it also provides a thinking method for the managers in the field of human resource management.

The human resource valuation models including the proposed approach can assist decision makers in recognizing the composition of employee grades; planning for increasing the productivity of human resources and employee development; more reasonable payments

<table>
<thead>
<tr>
<th>Position</th>
<th>Management consulting assistant</th>
<th>Management consulting senior</th>
<th>Management consulting executive</th>
<th>Human resource senior</th>
<th>Human resource executive</th>
<th>Marketing senior</th>
<th>Marketing executive</th>
<th>Chief executive officer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current age of individual</td>
<td>29</td>
<td>28</td>
<td>34</td>
<td>38</td>
<td>44</td>
<td>32</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>Life expectancy</td>
<td>0.969</td>
<td>0.970</td>
<td>0.960</td>
<td>0.952</td>
<td>0.935</td>
<td>0.964</td>
<td>0.941</td>
<td>0.941</td>
</tr>
<tr>
<td>Death probability (life expectancy−1)</td>
<td>0.031</td>
<td>0.030</td>
<td>0.040</td>
<td>0.048</td>
<td>0.066</td>
<td>0.036</td>
<td>0.059</td>
<td>0.059</td>
</tr>
</tbody>
</table>

Table IX. Computing the probability of the death of individuals
<table>
<thead>
<tr>
<th>Organizational position</th>
<th>Management consulting assistant</th>
<th>Management consulting senior</th>
<th>Management consulting executive</th>
<th>Human resource senior</th>
<th>Human resource executive</th>
<th>Marketing senior</th>
<th>Marketing executive</th>
<th>Chief executive officer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current monthly base salary of the individual</td>
<td>9,299,310</td>
<td>14,878,896</td>
<td>20,458,482</td>
<td>9,299,310</td>
<td>20,458,482</td>
<td>9,299,310</td>
<td>20,458,482</td>
<td>27,897,930</td>
</tr>
<tr>
<td>Annual salary and base salary of every individual in 2017</td>
<td>111,591,720</td>
<td>178,546,752</td>
<td>245,501,784</td>
<td>111,591,720</td>
<td>245,501,784</td>
<td>111,591,720</td>
<td>245,501,784</td>
<td>334,775,160</td>
</tr>
<tr>
<td>Work experience</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>18</td>
</tr>
</tbody>
</table>

Table X. Salary and wage data of individuals in the current year

An improved economic value model
to the employees and expenses on social welfare per employee; human asset valuation; and comparing human assets and total assets. Today the economic crisis and growth in Iran as a developing country is generating the pressure for evolution to the mental stage in which human beings increasingly seek greater fulfillment in harmonious relationships, psychological gratification and cultural enrichment, in addition to increasing material consumption. Thus, the progressive development of human resources made possible by the continuous evolution of human consciousness is the ultimate determinant of sustainability.

The proposed approach calls for a much more profound shift in thought and action to make the development of human capacities and fostering of human competencies together with increase in employee and customer satisfaction the center-piece of sustainable development strategy.

It is important to note that the valuation of human resources is more important for the human resource function and senior management than for line management and the accounting function of an organization. To facilitate the adoption of the proposed approach, it is important to note that human resource professionals might not have the necessary expertise to measure or value human resources and that the best way to make progress would be to use multifunctional teams including both human resource and accounting functions to improve the understanding and adoption of the new approach across the organization.

9. Conclusions
In this study, a new approach was proposed for improving human resource valuation based on economic value model. Compared to the economic value model, the proposed approach seems innovative since the competency model was used to calculate the ratio of an individual's promotion to higher positions every year and it was added to the model. Also, the Kano model was used for computing employee satisfaction and customer satisfaction coefficients, and the probability of individuals' leaving the organization for reasons other than death and retirement was added to the model.

In order to investigate the applicability of the proposed approach, it was implemented in a management consulting company. Questionnaire was used to determine the competency score of individuals, employee job satisfaction and customer satisfaction. The probability of people's death was obtained based on the mortality table and according to the current age of individuals. The expected income of individuals in the future was estimated based on past trend and by using regression method. Finally, the value of each employee was calculated using the proposed approach and the company's total human resource value was computed.

Human resource valuation includes monetary and nonmonetary measures. The monetary measures are divided into cost-based models and value-based models. The proposed approach in this paper was typically a value-based model. While the economic value model is classified as a monetary measure, the use of competency and Kano models improved this limitation. The proposed integrated model seems more effective compared to the basic model of economic value model, since the proposed approach can be recognized as an integration of both monetary and nonmonetary measures and this seems an important contribution to the knowledge of human resource valuation.

Table XI.
Computed value of human resources (in million rials)

<table>
<thead>
<tr>
<th>Organizational position</th>
<th>Management consulting assistant</th>
<th>Management consulting senior</th>
<th>Management consulting executive</th>
<th>Human resource executive</th>
<th>Human resource senior</th>
<th>Human resource executive</th>
<th>Marketing senior</th>
<th>Marketing executive</th>
<th>Chief executive officer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts' value</td>
<td>1,922</td>
<td>2,075.1</td>
<td>2,011.0</td>
<td>7,378.9</td>
<td>3,038</td>
<td>2,597.6</td>
<td>4,041.6</td>
<td>6,028.5</td>
<td></td>
</tr>
<tr>
<td>Note: Total value of human resources: 29,093.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Human resource value information can play a crucial role in internal managerial decision making, and its measures can be used to show that investments in a company’s human resources may result in long-term profit for the company. The proposed model of this study helps organizations in managing their human capitals more effectively and, in addition to increasing the efficiency and profitability of their respective organization, make decisions on their human resources based on the true competencies of individuals. On the other hand, since the proposed approach was developed on the basis of competency model, it helps employees in fair conditions to show their value to the organization and, accordingly, have a greater incentive to advance their career path.

As with any other research, this research also has some limitation. The proposed approach was implemented and investigated only in one organization and therefore, its applicability could be investigated in other organizations. In most of organizations, decisions about employee promotion are not usually made based on their competencies, and other issues are also involved in such decisions (e.g. the political behavior of organization), which is suggested to be considered in future studies.

In addition, it is suggested to the practitioners to apply the proposed approach in other organizations for human resource valuation and to compare the obtained results with the results of other human resource accounting models.

References


**Corresponding author**

Arash Shahin can be contacted at: shahin@ase.ui.ac.ir; arashshahin@hotmail.com

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldpublishing.com/licensing/reprints.htm](http://www.emeraldpublishing.com/licensing/reprints.htm)

Or contact us for further details: permissions@emeraldsight.com
Efficiency in European football teams using WindowDEA: analysis and evolution

Lucía Isabel García-Cebrián
Departamento de Dirección y Organización de Empresas,
Universidad de Zaragoza, Zaragoza, Spain, and
Fabiola Zambom-Ferraresi and Fernando Lera-López
Department of Economics,
Institute for Advanced Research in Business and Economics,
Universidad Pública de Navarra, Pamplona, Spain

Abstract
Purpose – The purpose of this paper is to analyze efficiency and its evolution in teams that played in the UEFA Champions League during nine seasons. The aim is to present a research procedure for determining the most accurate data envelopment analysis to estimate and compare the efficiency.

Design/methodology/approach – First, the authors analyzed the existence of a temporal trend using the S-statistic. The authors calculated the Kruskal–Wallis statistic to verify if there is stability in relative ranks. The results of the aforementioned tests have indicated that window analysis is an accurate methodology to apply to the sample. The authors analyzed 94 clubs with a sample of 288 observations, obtaining 768 efficiency ratios. They have been calculated using super-efficiency which enables to discriminate efficient units.

Findings – Results indicate that there is a low efficiency level in the nine seasons observed. There is a strong correlation between sports results and the efficiency of semifinalists. The authors conclude that improvement in a club’s efficiency could enhance its sports results. Finally, as practical implications, the authors highlight benchmark teams and alternative sports tactics to help clubs become more efficient and achieve better sports results.

Originality/value – This paper contributes to sports efficiency literature by presenting a research procedure to identify the most accurate methodology to be applied to panel data. To the best of the authors’ knowledge, this paper is the first empirical study on international football competitions applying WindowDEA to incomplete panel data.

Keywords Performance measurement, Benchmarking, Sports, Data envelopment analysis, UEFA Champions League

1. Introduction

The UEFA Champions League (UCL) is one of the most prestigious tournaments in the world and the most important club competition in European football. Nowadays, football, and specially the UCL, has a strong social impact. The Berlin final of the 2014/2015 tournament attained extraordinarily high television audience ratings as it aired in more than 200 countries to 400m viewers. Its social media presence has also exploded. The number of Twitter followers grew 51 percent during the 2014/2015 season, the UCL Facebook page is the world’s most popular league or association page on the platform, and during that season, overall page likes increased by 30 percent, reaching 45.6m likes (UEFA).

In the sports field, football is a very competitive sector. As with other organizations, professional football clubs must seek the best use of their resources. The financial-economic restrictions implemented by the FIFA Financial Fair Play have made the significance of the optimal use of available resources more apparent. Efficiency is even more essential in the UCL context, because clubs in small leagues must compete with Europe’s largest. Consequently, estimating efficiency as a means of analyzing clubs’ sports performance is an important approach if they need to consider the resources employed, not just the final
sporting result. The assessment of sports performance can also provide useful information to help team managers decide whether to hire players or make more investment on its own reserve of young players, help coaches design playing strategy or tactics and also help players train and improve their individual technical shortcomings.

Therefore, the main aim of this paper is to estimate and analyze the efficiency and evolution of clubs playing in the UCL using data envelopment analysis (DEA). The period under study covers 2004/2005 to 2012/2013, in other words, nine seasons forming a panel data set. DEA methodology does not need the specification of a production function and allows efficiency calculations in multi-input and multi-output organizations. Our second aim is to present a research procedure for determining the most accurate DEA methodology to estimate and compare the efficiency of panel data. The final selection of DEA methodology in this paper is based on the characteristics of the sample under study and not on the literature or research innovation bias. When testing for the existence of a temporal trend in our sample, window analysis emerged as the accurate methodology to analyze efficiency in this paper.

The remainder of the paper is as follows. Section 2 contains a review of the empirical evidence on efficiency in football. Section 3 describes the sample and variables we analyzed. Section 4 explains the research procedure we followed to discover the most accurate methodology and the methodology we applied. The results of the analysis of the temporal trend, the calculation of efficiency in DEA by means of window analysis and the results of efficiency are exposed in Section 5. In Section 6, we discuss our results and present our concluding thoughts.

2. Literature review

There is an extensive academic research on efficiency and its implications in the sports field. The efficiency analysis in football leagues is a well-established research line (Barros and Garcia-del-Barrio, 2011). Dawson et al. (2000) and Kulikova and Goshunova (2013) provide a comprehensive review of this literature.

From a methodological perspective, two main approaches have been used to measure the efficiency of sports: the econometric or parametric (stochastic frontier analysis) and the non-parametric frontier methodology (DEA). As Kulikova and Goshunova (2013) observed, DEA is the most popular and it measures technical efficiency. Technical efficiency refers to the ability of an organization or a decision-making unit (DMU in DEA literature) to obtain the maximum potential output from given amounts of factor inputs, or the minimum input required to obtain a given level of output. This concept involves physical quantities and technical relationships (Coelli et al., 2005).

Studies analyzing football efficiency include papers that observed regular national championships and eliminatory competitions. The main difference between them is that the outcome of eliminatory competitions is more uncertain when compared with regular leagues. The regular national football leagues reward the most stable performance and last longer, because teams usually play two rounds against all the others in the league. In contrast, eliminatory competitions require less time and random results are more likely. Regular competitions are normally leagues played in countries and eliminatory competitions involve nations or clubs in a large territory (i.e. the FIFA World Cup or the UCL).

There is more comprehensive empirical literature on regular leagues. The English Premier League (Haas, 2003; Barros and Leach, 2006; etc.) and the Spanish Liga (Espitia-Escuer and Garcia-Cebrián, 2004; González-Gómez and Picazo-Tadeo, 2010; etc.) are the most analyzed. Haas (2003), with a small sample size of 20 clubs (2000/2001 season), investigated how close the English Premier League clubs play to their potential. Barros and Leach (2006), combining financial and sports variables, applied DEA to measure the efficiency of the teams playing in the Premier League for five seasons (1998/1999 to
2002/2003) for the 12 clubs that participated in the competition in all the seasons studied. The authors made important recommendations at a managerial level, but, as a limitation, they highlighted the need for a more extensive panel data set to generalize the conclusions. Espitia-Escuer and García-Cebrián (2004) analyzed the efficiency of Spanish clubs playing in the national league for three seasons (1998/1999 to 2000/2001) and González-Gómez and Picazo-Tadeo (2010) compared the performance between the Spanish league, the Spanish cup and European competitions for six seasons (2001/2002 to 2006/2007). Other national leagues have also been analyzed, such as the Italian (Boscá Liern et al., 2009), German (Tiedemann et al., 2011), Portuguese (Ribeiro and Lima, 2012), Greek (Barros and Douvis, 2009), Brazilian (Barros et al., 2010) and Mexican (Torres-Dávila and García-Cebrián, 2012).

We would highlight the work by Tiedemann et al. (2011), which analyzed the football players in the German league for seven seasons (2002/2003 to 2008/2009), due to its different perspective. They found a clear positive relationship between a team’s average player efficiency score and its rank in the league table at the end of the season. Their results are corroborated by those of Sala-Garrido et al. (2009) which highlighted the very nature of football: the performance of the whole team is more important than that of its constituent parts.

The relation between estimated efficiency rankings and sports results is one of the most mentioned in the literature. Barros and Leach (2006) and Torres-Dávila and García-Cebrián (2012) found a statistical correlation between awarded points and estimated efficiency scores. Espitia-Escuer and García-Cebrián (2004) and Haas (2003), however, did not find a significant correlation between efficiency and sports results.

Concerning supranational competitions, the UCL is the world’s top competition at football club level, and efficiency studies are really scarce in this area (as examples we can find Espitia-Escuer and Garcia-Cebrián, 2010 and Zambom-Ferraresi et al., 2017). The fact that the sample comprises an unbalanced panel data set and the limited availability of data could be some reasons for this shortage.

Espitia-Escuer and García-Cebrián (2010) evaluated the efficiency of teams that played in the UCL for four seasons (2003/2004 to 2006/2007). First, they observed clubs’ efficiency season by season, and, in a second analysis, they estimated a frontier for the sample as a whole. The results indicated that in the four analyzed seasons there were no different dominating tactics, which means that there was no technological change in this period. In this case, inefficiency results from wasting resources. Zambom-Ferraresi et al. (2017) analyzed the sports performance of the same competition for ten seasons (2004/2005 to 2013/2014). They estimated season-by-season efficiency and found high inefficiency levels in the analyzed period, in contrast to the results of Espitia-Escuer and Garcia-Cebrián (2010).

As we have a panel of data in the present paper, our literature revision also focuses on the treatment found in former studies proposing DEA as a method for measuring the efficiency of football teams. Some of them have estimated its efficiency season by season (Tiedemann et al., 2011; Torres-Dávila and García-Cebrián, 2012; Zambom-Ferraresi et al., 2017) and/or have looked at the sample as a whole (Barros and Leach, 2006; Espitia-Escuer and Garcia-Cebrián, 2010; González-Gómez and Picazo-Tadeo, 2010). To the best of our knowledge, only Sala-Garrido et al. (2009) applied window analysis at a football league scope.

3. Data and variables
Considering existing literature, we propose estimating the efficiency of nine UCL seasons. This analysis will be based on DEA, because specifying a production function is unnecessary and it provides easily interpretable results. Our main interest is to know how efficiently teams use their sports resources, turning play styles and tactics into victories. Hence, the input measures are the sports statistics of the main actions on the field and the output measure is the sport result clubs achieve at the end of the competition.
Our sample comprises teams that played nine UCL seasons, from 2004/2005 to 2012/2013. As the clubs participating in the competition change from one year to the next, there are 32 clubs per season. Some clubs participated in more than one season and some of them participated in all the analyzed seasons. This leads us to have an unbalanced panel data set comprising 288 observations relating to 94 clubs.

In the present paper, we have taken the following variables as representatives of actions on the field: ball recoveries, crosses (open play), corners and total shots. All these inputs have been used in previous works (Sala-Garrido et al., 2009; Espitia-Escuer and Garcia-Cebrián, 2010; Torres-Dávila and García-Cebrián, 2012; Carmichael and Thomas, 2014) and they are the main resources employed by teams to try to score goals and win matches. The aggregated values of the playing statistics over the whole competition are provided by Opta Sports data. Table I shows the descriptive statistics of these variables.

The use of ball recovery as an indicator of actions on the field is relatively new and provides a large quantity of information on a team’s intention to dominate the game. Even when a team does not have much ball possession, if it tries to get the ball, this variable will measure the purpose of playing actively. The number of ball recoveries could be considered the outcome of defensive plays or the first step in attack plays (Carmichael and Thomas, 2014; Zambom-Ferraresi et al., 2017).

Crosses and corners are common plays and an important way to get close to the penalty area, where a major part of the plays turns into goals. Carmichael and Thomas (2014) included them in their model. Both inputs are indicators of one characteristic tactic. Crosses can be a quick way to go to the penalty area or a tactic variation that allows opening spaces in the central area. If a team mainly used these plays (crosses and corners), it could be characterized as a tactic style. On the other hand, if a team used them a great deal, but in the same amount as other plays, this could only indicate a variation in the type of play, which is very important to surprise opponent teams.

Total shots are one of the most important indicators of actions on the field in football. The main reasonable option to score a goal is shooting. Sala-Garrido et al. (2009), Espitia-Escuer and Garcia-Cebrián (2010), Torres-Dávila and García-Cebrián (2012) and Carmichael and Thomas (2014) used this input variable to estimate football teams’ efficiency.

The output used in this paper is a variable representing sports results during the Championship: the amount of the UCL financial retribution related to sports performance. They are the prizes clubs receive for advancing in the UCL phases. The ranking provided by this variable is the same as the ranking provided by points and played stages, but this output measure will allow us to differentiate between clubs that were eliminated in the same stage but have a different number of victories (Zambom-Ferraresi et al., 2017). For example, in the 2012/2013 season clubs received an €8.6m prize just for participating in the group stage, plus €1m for each victory and €0.5m for each draw. The clubs that passed into the round of 16 received €3.5m more; €3.9m for playing the quarterfinals, and €4.9m for playing the semifinals. The runner-up earned an extra prize of €6.5m and the champion earned €10.5m. Consequently, the minimum that a club could receive was €8.6m and if a club won

<table>
<thead>
<tr>
<th>Variable</th>
<th>Means</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total shots (I)</td>
<td>104.5</td>
<td>45.1</td>
<td>34</td>
<td>255</td>
<td>288</td>
</tr>
<tr>
<td>Crosses (I)</td>
<td>187.1</td>
<td>72.1</td>
<td>57</td>
<td>444</td>
<td>288</td>
</tr>
<tr>
<td>Corners (I)</td>
<td>38.8</td>
<td>16.5</td>
<td>12</td>
<td>108</td>
<td>288</td>
</tr>
<tr>
<td>Ball recoveries (I)</td>
<td>456.4</td>
<td>159.9</td>
<td>210</td>
<td>994</td>
<td>288</td>
</tr>
<tr>
<td>Sports results (O)</td>
<td>15.5</td>
<td>6.6</td>
<td>8.6</td>
<td>39.9</td>
<td>288</td>
</tr>
</tbody>
</table>

Table I. Descriptive statistics
all the matches in the group stage and won the competition, it could receive a prize of €37.4m for its sports performance. As prizes change every three years and our sample covers nine seasons, we have taken the values of last season’s prize as the reference year.

4. Calculation of efficiency using DEA with panel data
Several methodologies have been proposed in the literature to calculate efficiency. One approach is frontier methods, which consider the estimation of isoquant as the sample data envelope. Among frontier methods, deterministic non-parametric and stochastic methods have also been most often used in empirical papers. Deterministic non-parametric or DEA does not specify a functional form for production processes and this is its main advantage. Nevertheless, all deviations from the frontier are classified as inefficient.

Under the assumption of constant returns to scale, DEA proposes the solution to the following linear programming problem for every unit in the sample:

\[
\text{Min. } \lambda_i, \quad (1)
\]

s.t.: 

\[
\lambda_i x \geq z X \\
y \leq z y \\
z \in R^+,
\]

where \(x\) is the vector of consumed inputs by unit \(i\) under analysis, \(X\) is the matrix of consumed inputs for all units in the sample, \(y\) is the vector of obtained output by unit \(i\), \(Y\) is the matrix of obtained outputs for all units in the sample, \(z\) is a vector of parameters whose values are obtained in the resolution of the problem and \(\lambda_i\) is the ratio of efficiency for unit \(i\). In the orientation to input presented in this problem, \(\lambda_i\) represents the radial reduction to be applied to every input in unit \(i\) to become efficient.

In the original DEA formulation, efficient units present a ratio equal to 1 and their ratio for inefficient units is less than 1. This supposes that efficient units in a sample show the same value of efficiency and it is impossible to discriminate among them. To solve this inconvenience, Andersen and Petersen (1993) proposed the calculation of super-efficiency, which consists of calculating efficiency ratios on the basis of the aforementioned problem, but taking the whole sample except the unit under analysis as a comparison sample. Consequently, inefficient units present the same ratio of efficiency as solving the original DEA problem, but efficient units do not have their efficiency ratio limited by 1, so they can present ratios above this value and we can order them by the criteria of the highest value. Therefore, the efficiency ratios calculated for football clubs in the sample under study have been calculated using super-efficiency.

In an empirical work aiming to discover the efficiency of units forming a sample, the availability of information in the form of panel data is an improvement to be taken into consideration. Of course, calculations can be repeated in every single period in the time horizon, but the treatment of pooled data can be taken advantage of as a whole. If a researcher has panel data, several analysis possibilities exist. Tulkens and Vanden Eeckaut (1995), Asmild et al. (2004) and Cullinane and Wang (2006) propose four different calculations of efficiency using DEA from a panel data set:

1) Contemporaneous approach: efficiency is calculated for every DMU taking only the input and output data of DMUs in the same period as a reference set. As many efficiency ratios as periods in the whole sample are calculated for each DMU, each of them with a different frontier.
Sequential: efficiency is calculated for every DMU in one period taking input and output data of DMUs in the same period and all the precedents as a reference set.

Intertemporal: efficiency is calculated for every DMU in one period taking input and output data of DMUs in the whole time of the sample as a reference set. Efficiency is calculated taking all the data from the panel in a pooled manner. As many efficiency ratios as periods in the whole sample are calculated for each DMU, all of them with the same frontier.

Window analysis: efficiency is calculated for every DMU in one period taking input and output data of DMUs in the periods forming the window length as a reference set. Researchers should decide on window length depending on the characteristics of each empirical work. DMUs forming each window vary because the first period is suppressed to add the following one to form the next window.

The choice of methodology requires taking their characteristics into account. Following Tulkens and Vanden Eeckaut (1995) concerning changes in production technology, it could be argued that if the contemporaneous method is chosen, the efficient frontier is considered to change from period to period. Concerning the technology available at each moment, it is assumed that, with the sequential method, the way of producing in the past will also be available in the future; therefore, only shifts in the frontier reflecting technical progress are assumed. If the intertemporal frontier is chosen, it is assumed that no shift occurs.

If there is no temporal trend, the efficiency calculated year by year only considers DMU data in the same period as a reference sample and some organizations could be qualified as efficient, even though they could be qualified as inefficient if they were compared with the same units' performance in different periods. In this case of no temporal trend, the best way of exploiting information provided by panel data is to estimate an intertemporal frontier and calculate efficiency by taking it as a reference set. In fact, Brockett et al. (1998) establish that the use of a single efficiency frontier assumes that no technological changes affecting productive efficiency have occurred over the time periods. Wang et al. (2014) explain that, in the presence of technical change, evaluating efficiency using an intertemporal frontier formed by many periods can deem that DMUs observed in the technical improvement periods are efficient. Therefore, DMUs of periods with no technical improvement and considered inefficient would be recommended to become a benchmark impossible to reach given their technological context.

Where a temporal trend is detected, the year-by-year calculation of efficiency can lose information: it is difficult to analyze improvements in efficiency because the sample changes every year. In this case, two different situations can be applied. If the temporal trend is in fact the realization of an event at a single and recognizable moment that drastically changes technological conditions, efficiency could be calculated with two intertemporal frontiers, one before that event and the other afterwards. Brockett et al. (1998) proposed a procedure in the event of no drastic changes. These authors quoted that these trends can develop slowly, only in a few DMUs (they are not generalized), and can go unnoticed by managers. Brockett et al. (1998) also remarked that efficiency results calculated across time in contrast to year-by-year results are more useful to managers as they avoid the influence of extraordinary events in a single period.

Proposing recommendations that fit with the technological evolution in the sample under study require verification of this evolution before starting the efficiency calculation. Consequently, Brockett et al. (1998) suggested an intertemporal calculation of efficiency and the running of several tests on the obtained efficiency ratios. For the study of the eventual existence of trends in performance over time, following Brockett et al. (1998) and Ross and Droge (2002), we have used the S-statistic proposed by Brockett and Kemperman (1980). For the analysis of stability in relative ranks, we have used the Kruskal–Wallis statistic,

Efficiency in European football teams

2131
5. Results

5.1 Analysis of the existence of temporal trend

We are going to verify the two aspects proposed in Brockett and Golany (1996) and Brockett et al. (1998): analysis of the existence of a temporal trend in the whole sample and of stability in relative ranks.

Concerning the analysis of a temporal trend in the whole sample, Brockett et al. (1998) and Ross and Droge (2002) use the $S$-statistic proposed by Brockett and Kemperman (1980). These authors have a complete panel in their analysis. In our case, we have unbalanced panel data because not all the teams in the sample have played in the UCL throughout the entire period under analysis. Nevertheless, it is possible to apply the $S$-statistic because observations are accumulated by period and 32 teams play in the UCL every season. As we have previously obtained super-efficiency in our calculations, we have avoided ties in the application of the test. As we obtained a statistic value equal to 3.830, we found that we can reject the null hypothesis of no temporal trend with a significance level of 5 percent.

To analyze the stability in relative ranks, Brockett et al. (1998) and Sueyoshi and Aoki (2001) propose the use of the Kruskal–Wallis statistic. Brockett and Golany (1996) also quote this method. Brockett et al. (1998) and Sueyoshi and Aoki (2001) have a complete panel again, but the original paper of Kruskal and Wallis (1952) considers the possibility of a different number of observations for each individual in the statistic calculation. We have calculated this statistic for our sample using super-efficiency ratios, again to avoid ties, and we have obtained a value for the statistic equal to 101.66; therefore, we have accepted the null hypothesis of equivalent distribution of efficiency ranking with a level of significance of 5 percent.

To sum up, we have detected a temporal trend and teams have not maintained their relative positions over time. Therefore, intertemporal analysis cannot be conducted with our data to determine which teams are efficient because a temporal trend exists. The contemporaneous approach is also incomplete because the technical changes in football (players and coaches) are usually gradual and it is possible to find similar contexts in two consecutive periods. Although Brockett et al. (1998) do not use window analysis, they quote this method in their paper referring to it as a way of introducing new temporal data in the DEA calculation. These authors mention that studies applying WDEA do not use statistical tests, an issue that we try to solve in our paper. Furthermore, window analysis provides knowledge of robustness in efficiency ratios. This helps to identify really reliable results and, as a consequence, to identify which DMUs can be used as an efficiency benchmark. Due to statistically equivalent distribution of efficiency positions, we would not expect efficient teams to be the same throughout the period studied in this paper.

5.2 Calculation of efficiency in DEA using window analysis

Window analysis consists of calculating efficiency rates using DEA but forming the sample with data collected from a number of consecutive periods. This sample extracted from the whole panel is called a window, and the number of periods in a window is called the width of the window. Calculations are repeated by eliminating the data corresponding to the first period in the window and adding data for the following period (season in our case) after the last in the window. This process is repeated as many times as the length of panel provides data. Window analysis enables us to assess the robustness of the efficiency ratios and to come to more reliable conclusions on the evolution of efficiency.
As we have previously highlighted, only Sala-Garrido et al. (2009) applied window analysis methodology to the sports field. Traditionally, apart for a methodological approach (Tulkens and Vanden Eeckaut, 1995; Sueyoshi and Aoki, 2001), studies considering window analysis have been applied to various economic sectors: banking (Yue, 1992; Asmild et al., 2004); air forces (Charnes et al., 1984); brewing industry (Day et al., 1994); carbonated beverage industry (Charnes, Cooper, Golany, Learner, Phillips and Rousseau, 1994); container port (Cullinane and Wang, 2006); semiconductor manufacturer (Chung et al., 2008); telecommunications firms (Yang and Chang, 2009); science park (Sun and Lin, 2009); and coffee sector (Suárez and Mejía, 2010), for instance. Recent works have also applied window analysis to calculate efficiency. For example, Detotto et al. (2014) analyzed the productivity of the hospitality sector in Italian regions; Wang et al. (2014) and Epure and Lafuente (2015) studied efficiency in banking. Mariano et al. (2015) presented a revision of the literature on the application of DEA to analyze efficiency in human development processes and window analysis appears as one of the extensions of DEA if panel data are available.

When applying window analysis, one important decision is to establish the width of windows, which depends on assumptions about changes in the frontier (Tulkens and Vanden Eeckaut, 1995). Several approaches have been proposed in the literature. Papers by Boussofiane et al. (1991), Day et al. (1994), Ross and Droge (2002), Sala-Garrido et al. (2009) and Suárez and Mejía (2010) state that it is a decision taken by an analyst. Nevertheless, to provide some criteria, Paradi et al. (2004), Asmild et al. (2004), Charnes, Cooper, Golany, Learner, Phillips and Rousseau (1994) and Yang and Chang (2009) suggest that windows should be wide enough to have the necessary degrees of freedom and narrow enough to maintain the same context. Finally, more formal suggestions are those by Cooper et al. (2004), Charnes, Cooper, Lewin and Seiford (1994, p. 60), who propose the trial and error method, and Cooper et al. (2000), who propose a formula.

Following the formulae provided by Cooper et al. (2000), we should have used a width of five seasons in the present paper. Nevertheless, Sala-Garrido et al. (2009) suggest that most teams change either players or their coaches from one season to another, which means there is no sense analyzing windows of more than three seasons for football clubs. Regarding this consideration, we have also employed a width of three seasons, as this period is consistent with the average time coaches, players and teams remain in the championship.

Before exposing the efficiency results, it is important to remember that we do not have a complete panel of data. In this paper, efficiency for each team in each window is calculated with the sample formed by teams playing in some of the three seasons forming the window, although some of them did not participate in all three seasons.

The overall results of efficiency (WDEA) are presented in Table II. Due to a shortage of space, Table II is organized by season. For seasons that belong to more than one window, we present the efficiency scores for all the respective windows (two or three). The differences between the efficiency scores for the same season in different windows give us an idea of the robustness of the results. The coefficient of variation (CV) between efficiency scores is added to help observe robustness whenever there is more than one window. The results are sorted by the clubs’ sporting performance. The lines separating clubs in the results table represent the final stage in which clubs are ranked in each season. From top to bottom, these are final (champion and runner-up), semifinals, quarterfinals, knockout round and group stage. The results highlighted in gray are efficient observations and results underlined are not considered robust.

To interpret the results, it is important to remember that observations with an efficiency score of less than 1 are inefficient and of 1 or more are efficient. For general results, a CV above 10 percent was considered to mean substantial differences in efficiency values and not robust. On the other hand, all those observations that present a CV < 10 percent had estimated robust efficiency scores. As we have used super-efficiency in our calculations,
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Champion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liverpool05</td>
<td>Barcelona06</td>
<td>Milan07</td>
</tr>
<tr>
<td>1.16</td>
<td>1.06</td>
<td>0.92</td>
</tr>
<tr>
<td><strong>Runner-up</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milan05</td>
<td>Arsenal06</td>
<td>Liverpool07</td>
</tr>
<tr>
<td>0.91</td>
<td>1.06</td>
<td>0.93</td>
</tr>
<tr>
<td><strong>Semifinalist</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chelsea05</td>
<td>Villarrea06</td>
<td>Chelsea07</td>
</tr>
<tr>
<td>0.92</td>
<td>0.92</td>
<td>0.72</td>
</tr>
<tr>
<td>PSV05</td>
<td>Milan06</td>
<td>Manchester United07</td>
</tr>
<tr>
<td>0.73</td>
<td>0.79</td>
<td>0.73</td>
</tr>
<tr>
<td><strong>Quarter-finalists</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juventus05</td>
<td>Lyon06</td>
<td>Valencia07</td>
</tr>
<tr>
<td>0.77</td>
<td>0.81</td>
<td>0.75</td>
</tr>
<tr>
<td>Internazionale05</td>
<td>Juventus06</td>
<td>Bayern Munich07</td>
</tr>
<tr>
<td>0.82</td>
<td>0.74</td>
<td>0.73</td>
</tr>
<tr>
<td>Lyon06</td>
<td>Internazionale06</td>
<td>PSV Eindhoven07</td>
</tr>
<tr>
<td>0.82</td>
<td>0.82</td>
<td>0.75</td>
</tr>
<tr>
<td>Bayern Munich05</td>
<td>Benfica06</td>
<td>Roma07</td>
</tr>
<tr>
<td>0.88</td>
<td>0.77</td>
<td>0.75</td>
</tr>
<tr>
<td><strong>Knockout stages</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Werder Bremen05</td>
<td>Liverpool06</td>
<td>Lyon07</td>
</tr>
<tr>
<td>0.87</td>
<td>0.70</td>
<td>0.77</td>
</tr>
<tr>
<td>Monaco05</td>
<td>Bayern Munich06</td>
<td>Porto07</td>
</tr>
<tr>
<td>0.88</td>
<td>0.77</td>
<td>0.71</td>
</tr>
<tr>
<td>Manchester United05</td>
<td>Chelsea06</td>
<td>Real Madrid07</td>
</tr>
<tr>
<td>0.89</td>
<td>0.86</td>
<td>0.74</td>
</tr>
<tr>
<td>Arsenal05</td>
<td>Ajax06</td>
<td>Arsenal07</td>
</tr>
<tr>
<td>0.92</td>
<td>0.82</td>
<td>0.66</td>
</tr>
<tr>
<td>Real Madrid05</td>
<td>PSV Eindhoven06</td>
<td>Barcelona07</td>
</tr>
<tr>
<td>0.76</td>
<td>1.02</td>
<td>0.78</td>
</tr>
<tr>
<td>Bayer Leverkusen05</td>
<td>Real Madrid06</td>
<td>Internazionale07</td>
</tr>
<tr>
<td>0.75</td>
<td>0.72</td>
<td>0.95</td>
</tr>
<tr>
<td>Barcelona05</td>
<td>Rangers06</td>
<td>Lille07</td>
</tr>
<tr>
<td>0.82</td>
<td>0.91</td>
<td>0.66</td>
</tr>
<tr>
<td>Porto05</td>
<td>Werder Bremen06</td>
<td>Celtic07</td>
</tr>
<tr>
<td>0.87</td>
<td>0.66</td>
<td>0.88</td>
</tr>
<tr>
<td><strong>Group stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamo Kyiv05</td>
<td>Schalke06</td>
<td>Werder Bremen07</td>
</tr>
<tr>
<td>0.92</td>
<td>0.87</td>
<td>0.71</td>
</tr>
<tr>
<td>Olympiakos05</td>
<td>Club Brugge06</td>
<td>CSKA Moscow07</td>
</tr>
<tr>
<td>1.03</td>
<td>0.84</td>
<td>0.89</td>
</tr>
<tr>
<td>Panathinaikos05</td>
<td>Udinese06</td>
<td>AEK07</td>
</tr>
<tr>
<td>1.16</td>
<td>0.75</td>
<td>0.83</td>
</tr>
<tr>
<td>Fenerbahce05</td>
<td>Besiktas06</td>
<td>Copenhagen07</td>
</tr>
<tr>
<td>0.90</td>
<td>0.70</td>
<td>0.82</td>
</tr>
<tr>
<td>Valencia05</td>
<td>Lille06</td>
<td>Bordeaux07</td>
</tr>
<tr>
<td>0.84</td>
<td>0.72</td>
<td>0.70</td>
</tr>
<tr>
<td>CSKA Moscow05</td>
<td>Artmedia06</td>
<td>Benfica07</td>
</tr>
<tr>
<td>0.89</td>
<td>0.84</td>
<td>0.79</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Champion</strong></td>
<td><strong>Runner-up</strong></td>
</tr>
<tr>
<td>Manchester United 08</td>
<td>Chelsea 08 0.83 0.87 0.78 0.06</td>
</tr>
<tr>
<td><strong>Semifinalist</strong></td>
<td><strong>Quarter-finalists</strong></td>
</tr>
<tr>
<td>Barcelona 08 0.80 0.80 0.80 0.00</td>
<td>Chelsea 09 0.83 0.75 0.75 0.06</td>
</tr>
<tr>
<td>Liverpool 08 0.74 0.75 0.72 0.02</td>
<td>Arsenal 09 0.79 0.74 0.74 0.03</td>
</tr>
<tr>
<td><strong>Knockout stages</strong></td>
<td></td>
</tr>
<tr>
<td>Internazionale 08 0.92 0.96 0.88 0.04</td>
<td>Juventus 09 0.84 0.76 0.76 0.06</td>
</tr>
<tr>
<td>Sevilla 08 0.69 0.75 0.65 0.07</td>
<td>Atlético Madrid 09 0.94 0.86 0.86 0.05</td>
</tr>
<tr>
<td>Milan 08 0.75 0.82 0.71 0.07</td>
<td>Sporting Lisboa 09 0.94 0.84 0.84 0.06</td>
</tr>
<tr>
<td>Olympiacos 08 0.86 0.95 0.85 0.06</td>
<td>Roma 09 0.93 0.84 0.84 0.06</td>
</tr>
<tr>
<td><strong>AV</strong></td>
<td>0.84</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.13</td>
</tr>
</tbody>
</table>

**Table II. Efficiency in European football teams**
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Clubs/windows</td>
<td>Clubs/windows</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>CV</td>
<td>CV</td>
</tr>
</tbody>
</table>

**Champion**
- Internazionale10: 1.35, 1.22, 1.22, 0.06
- Barcelona11: 1.18, 1.05, 1.17, 0.06

**Runner-up**
- Bayern Munich10: 0.90, 0.90, 0.90, 0.00
- Manchester United11: 0.91, 0.84, 0.92, 0.05

**Semifinalist**
- Lyon10: 0.79, 0.79, 0.79, 0.00
- Real Madrid11: 0.71, 0.71, 0.72, 0.01

(continued)
<table>
<thead>
<tr>
<th>Quarter-finalists</th>
<th>Barcelona10</th>
<th>0.73</th>
<th>0.73</th>
<th>0.73</th>
<th>0.00</th>
<th>Schalke11</th>
<th>0.78</th>
<th>0.75</th>
<th>0.83</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bordeaux10</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.00</td>
<td>Shalhtar Donetsk11</td>
<td>0.87</td>
<td>0.80</td>
<td>0.88</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Manchester United10</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.00</td>
<td>Chelsea11</td>
<td>0.60</td>
<td>0.60</td>
<td>0.69</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Arseni10</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.00</td>
<td>Tottenham11</td>
<td>0.81</td>
<td>0.76</td>
<td>0.82</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>CSKA Moscow10</td>
<td>0.79</td>
<td>0.79</td>
<td>0.76</td>
<td>0.02</td>
<td>Internazionale11</td>
<td>0.72</td>
<td>0.71</td>
<td>0.71</td>
<td>0.01</td>
</tr>
</tbody>
</table>

|Knockout stages| Chelsea10 | 0.80  | 0.80 | 0.80 | 0.00 | Bayern Munich11 | 0.64  | 0.64 | 0.74 | 0.08 |
|              | Fiorentina10 | 0.84  | 0.84 | 0.84 | 0.00 | Valencia11 | 0.58  | 0.58 | 0.69 | 0.10 |
|              | Sevilia10 | 0.78  | 0.78 | 0.78 | 0.00 | Marseille11 | 0.55  | 0.55 | 0.66 | 0.10 |
|              | Real Madrid10 | 0.69  | 0.69 | 0.69 | 0.00 | Arsenal11 | 0.87  | 0.80 | 0.85 | 0.04 |
|              | Porto10 | 0.66  | 0.66 | 0.66 | 0.00 | Copenahgun11 | 0.85  | 0.77 | 0.85 | 0.05 |
|              | Olymipicos10 | 0.80  | 0.80 | 0.80 | 0.00 | Roma11 | 0.65  | 0.65 | 0.75 | 0.08 |
|              | Milan10 | 0.80  | 0.80 | 0.80 | 0.00 | Lyon11 | 0.55  | 0.55 | 0.64 | 0.09 |
|              | Stuttgart10 | 0.76  | 0.76 | 0.76 | 0.00 | Milan1 | 0.80  | 0.71 | 0.78 | 0.06 |

|Group stage| Unirea Urziensii10 | 0.92  | 0.92 | 0.90 | 0.01 | Sporting Braga11 | 0.95  | 0.95 | 0.97 | 0.01 |
|           | Juventus10 | 0.85  | 0.85 | 0.85 | 0.00 | Spartak Moscow11 | 0.68  | 0.66 | 0.74 | 0.06 |
|           | Liverpool10 | 0.80  | 0.80 | 0.80 | 0.00 | Ajax11 | 0.67  | 0.63 | 0.73 | 0.07 |
|           | Marseille10 | 0.69  | 0.69 | 0.69 | 0.00 | Rangers11 | 1.01  | 0.99 | 1.02 | 0.01 |
|           | Wolfsburg10 | 0.69  | 0.69 | 0.69 | 0.00 | Rubin Kazan11 | 0.81  | 0.70 | 0.75 | 0.07 |
|           | Rubin Kazan10 | 0.95  | 0.95 | 0.90 | 0.03 | Twente11 | 0.56  | 0.56 | 0.65 | 0.09 |
|           | Dynamo Kyiv10 | 0.72  | 0.72 | 0.72 | 0.00 | Benfica11 | 0.55  | 0.55 | 0.66 | 0.10 |
|           | Standard Liege10 | 0.79  | 0.79 | 0.79 | 0.00 | Hapoel Tel-Aviv11 | 0.92  | 0.81 | 0.85 | 0.06 |
|           | AZ Alkmaar10 | 0.75  | 0.75 | 0.75 | 0.00 | Werder Bremen11 | 0.52  | 0.52 | 0.62 | 0.10 |
|           | Apoel10 | 1.23  | 1.26 | 1.00 | 0.12 | Basel11 | 0.56  | 0.56 | 0.65 | 0.09 |
|           | Zurich10 | 0.93  | 0.93 | 0.91 | 0.02 | CFR Cluj11 | 0.70  | 0.67 | 0.75 | 0.05 |
|           | Besiktas10 | 0.67  | 0.67 | 0.67 | 0.00 | Panathinaikos11 | 0.92  | 0.80 | 0.82 | 0.08 |
|           | Atletico Madrid10 | 0.66  | 0.66 | 0.66 | 0.00 | AJ Auxerre11 | 0.61  | 0.57 | 0.62 | 0.05 |
|           | Rangers10 | 0.72  | 0.72 | 0.72 | 0.00 | Bursaspor11 | 0.61  | 0.59 | 0.65 | 0.05 |
|           | Maccabi Haifa10 | 0.69  | 0.69 | 0.69 | 0.00 | MSK Zilina11 | 0.52  | 0.52 | 0.59 | 0.07 |
|           | Debreceni VSC10 | 0.70  | 0.70 | 0.70 | 0.00 | Partizan11 | 0.57  | 0.49 | 0.53 | 0.07 |

(continued)
Table II.

<table>
<thead>
<tr>
<th>Season</th>
<th>Clubs/windows</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011/12</td>
<td>0.91</td>
<td>0.80</td>
<td>0.79</td>
<td>0.73</td>
<td>0.69</td>
</tr>
<tr>
<td>2012/13</td>
<td>0.14</td>
<td>0.11</td>
<td>0.08</td>
<td>0.09</td>
<td>0.07</td>
</tr>
</tbody>
</table>

**Champion**
- **Chelsea** 2012: 0.88
- **Bayern Munich** 2013: 0.95

**Runner-up**
- **Bayern Munich** 2012: 0.69
- **Borussia Dortmund** 2013: 0.89

**Semifinalists**
- **Real Madrid** 2012: 0.72
- **Barcelona** 2013: 0.85

**Quarter-finalists**
- **Benfica** 2012: 0.61
- **Paris Saint-Germain** 2013: 0.81

**Knockout stages**
- **Paris Saint-Germain** 2013: 0.81
- **Malaga** 2013: 0.82
- **Juventus** 2013: 0.70
- **Galatasaray** 2013: 0.73

**Group stage**
- **CFR Cluj** 2013: 0.91
- **Shakhtar Donetsk** 2013: 0.73
- **Olympiacos** 2013: 0.75

(continued)
<table>
<thead>
<tr>
<th>Team</th>
<th>CV</th>
<th>SD</th>
<th>Small</th>
<th>Medium</th>
<th>CV</th>
<th>SD</th>
<th>Large</th>
<th>CV</th>
<th>SD</th>
<th>Small</th>
<th>Medium</th>
<th>CV</th>
<th>SD</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valencia</td>
<td>0.56</td>
<td>0.67</td>
<td>0.70</td>
<td>0.11</td>
<td>Zenit</td>
<td>0.69</td>
<td>0.71</td>
<td>0.79</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Porto</td>
<td>0.51</td>
<td>0.59</td>
<td>0.66</td>
<td>0.12</td>
<td>Anderlecht</td>
<td>0.73</td>
<td>0.74</td>
<td>0.79</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trabzonpor</td>
<td>0.74</td>
<td>0.79</td>
<td>0.80</td>
<td>0.04</td>
<td>BATE Borisov</td>
<td>1.17</td>
<td>1.22</td>
<td>1.69</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lille</td>
<td>0.55</td>
<td>0.66</td>
<td>0.69</td>
<td>0.11</td>
<td>Dynamo Kyiv</td>
<td>0.72</td>
<td>0.73</td>
<td>0.83</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viktoria Plzen</td>
<td>0.67</td>
<td>0.76</td>
<td>0.78</td>
<td>0.08</td>
<td>Ajax</td>
<td>0.67</td>
<td>0.69</td>
<td>0.76</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shakhtar Donetsk</td>
<td>0.54</td>
<td>0.63</td>
<td>0.65</td>
<td>0.10</td>
<td>Manchester City</td>
<td>0.68</td>
<td>0.70</td>
<td>0.73</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KRC GenK</td>
<td>0.69</td>
<td>0.75</td>
<td>0.76</td>
<td>0.05</td>
<td>Spartak Moscow</td>
<td>0.70</td>
<td>0.71</td>
<td>0.80</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borussia Dortmund</td>
<td>0.52</td>
<td>0.57</td>
<td>0.59</td>
<td>0.07</td>
<td>Lille</td>
<td>0.62</td>
<td>0.64</td>
<td>0.67</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BATE Borisov</td>
<td>0.70</td>
<td>0.71</td>
<td>0.76</td>
<td>0.05</td>
<td>Sporting Braga</td>
<td>0.60</td>
<td>0.66</td>
<td>0.75</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dinamo Zagreb</td>
<td>0.66</td>
<td>0.71</td>
<td>0.72</td>
<td>0.05</td>
<td>Montpellier</td>
<td>0.55</td>
<td>0.57</td>
<td>0.63</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Villareal</td>
<td>0.61</td>
<td>0.65</td>
<td>0.66</td>
<td>0.04</td>
<td>Dinamo Zagreb</td>
<td>0.68</td>
<td>0.68</td>
<td>0.75</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OteluGalati</td>
<td>0.62</td>
<td>0.65</td>
<td>0.66</td>
<td>0.03</td>
<td>Nordsjaelland</td>
<td>0.80</td>
<td>0.81</td>
<td>0.96</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AV</td>
<td>0.68</td>
<td>0.75</td>
<td>0.78</td>
<td></td>
<td></td>
<td>0.76</td>
<td>0.79</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>0.16</td>
<td>0.15</td>
<td>0.15</td>
<td></td>
<td></td>
<td>0.12</td>
<td>0.13</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: CV, coefficient of variation. Values highlighted in gray = efficient; Italic values = non robust results.
efficient units present ratios above 1, independent of the exact ratio value. We will, therefore, consider efficient teams robust if the ratios in all the seasons in the same window are above 1, even if their CV is above 10 percent. Consequently, efficient teams at least once with CV less than 10 percent or with CV above 10 percent, but efficient in all seasons in a window, will be considered robust efficient clubs and will be used as reference units and benchmarking for the inefficient clubs in the sample.

The first highlighted result concerns the low efficiency level of the UCL for the 2004/2005 to 2012/2013 seasons. This result agrees with those found by Zambom-Ferraresi et al. (2017). Our sample comprises 288 observations and, due to the repetition in calculating efficiency in WindowDEA, we have 768 different ratios and only 6 percent of the entire sample is efficient. Looking at windows separately, these results are similar. On the one hand, the window with more efficient observations is window 3, with nine efficient observations. On the other, window 4 had only three efficient observations.

Clubs that had robust efficient scores in their respective windows will be a reference for the rest of the sample. Internazionale 2009/2010 and Barcelona 2010/2011 are the best benchmarking observations. Both clubs were champions and had relative technically efficient scores when compared with other teams that played in windows 4–6 and windows 5–7, respectively. Apoel’s 2011/2012 performance is an example and must be a reference for all the small and medium clubs in the sample. This club reached the quarterfinals of the UCL by being efficient when compared with the performance of the other clubs in the two periods before and after. The performance of Slavia Prague 2007/2008, Apoel 2009/2010 and BATE Borisov 2012/2013 is also worthy of note. These clubs did not advance past the knockout round, but as they did not waste their sports resources, they can be qualified as efficient. These robust efficient clubs achieved good sports results on the basis of the inputs they employed.


Finally, efficient teams are not the same throughout the entire time horizon in this paper, as expected after the statistical acceptation of null hypothesis of the equivalent distribution in efficiency ranking. This lack of continuity in efficient teams is not only due to the existence of teams not playing in all the seasons under study, as some of the efficient teams have a continuous presence in the UCL.

5.3 Evolution of efficiency and sports results

Given that WDEA was employed, we can draw more accurate conclusions about the evolution of efficiency. By observing the relation between DMU efficiency in the first season of a window and the variation in its efficiency during the window period, we can observe interesting issues. Figure 1(a)–(g) is easy to interpret and helps us to analyze the evolution of efficiency. The clubs on the right side of the horizontal axis have shown positive efficiency changes, and the clubs on the left side have had negative changes during the window cycle. Clubs in the top left quadrant are efficient in the first season of the windows, but cannot maintain this efficiency in the next two seasons. We find nine observations with this performance and four of them are the champions of the competition in the first season of the respective window (w): Barcelona 2005/2006 (w 2) and 2010/2011 (w 7), Internazionale 2009/2010 (w 7) and Liverpool 2004/2005 (w 1).

In the seven windows we studied, no case was found in the top right quadrant, which means that no clubs were capable of maintaining efficiency for three seasons. These results
corroborate those found by Zambom-Ferraresi et al. (2017), where teams had many problems maintaining their efficiency during the seasons.

In the bottom left quadrant, we can see those clubs that were inefficient in the first season of the window and cannot improve on this inefficiency during the window cycle. The most highlighted cases are Lyon and Manchester United, which are inefficient and do not change in five out of seven analyzed windows. Real Madrid was found in the same case in four out
of seven of the observed windows and this team is present in all of them. In contrast, Porto participated in the UCL in four out of these seven windows and was inefficient in all the seasons in which they played.

Porto’s situation captures our attention. In the last decade, exactly the period analyzed here, Porto was known as one of the best clubs at the time for hiring (mainly) and training young, talented and unknown players. This means that, when hiring players, Porto, (www.dailymail.co.uk/sport/football/article-3023048/Danilo-transfer-Real-Madrid-means-Porto-440m-player-sales-winning-Champions-League-2004.html) presents a satisfactory performance, but its resources on the field were used inefficiently. Pursuing efficiency in all areas is essential for medium clubs like Porto, which is competing on the field with the largest clubs in Europe.

Finally, the bottom right quadrant includes clubs that improved their efficiency scores during the window cycle, but their improvement did not suffice to be efficient when their performance is compared with all the clubs in the window. The clubs in this situation are Arsenal, Chelsea and Internazionale, four times out of the seven possible. Barcelona and Real Madrid also improved their efficiency three times during the seven observed windows. We could see that these are major clubs that participated in the UCL in all or almost all the analyzed windows. In some performances they had good sports results, but they wasted their resources, e.g. Arsenal reached the 2008/2009 semifinal and Chelsea was the 2011/2012 champion. Sometimes clubs present very poor sports results considering the number of inputs they employed and, in these cases, their inefficiency is evident. This is the case of Real Madrid in windows 4–6.

The relation between sports results and efficiency can be observed in Figure 2. Most observations are inefficient, which explains the few observations on the right side of the axis of the efficiency score. The observations of the first two competition stages (group and knockout) are very crowded, as was expected, because both concentrate the largest part of the sample. The highest sports prizes for these two groups could be 14.6 and 18.1 (million euros), respectively. A growing trend in Figure 2 can only be observed in the final stages of the competition.

![Figure 2. Relationship of efficiency evolution and sports results](image)

**Note:** n = 768
After analyzing the correlation between sports results and the efficiency score of all the 768 ratios in the sample in Table III, we have found a positive and significant correlation of almost 40 percent of the sample. However, the importance of efficiency is obvious when we look at the correlation for the clubs that reach the semifinal stage. The correlation between efficiency and the sports results in this phase is 72 percent. By observing these results, we can conclude that for all kinds of club, whether large or small, efficiency is significant, positive and highly correlated with reaching the final phases of the UCL.

Individually, one of the most highlighted performances when observing both efficiency and sports results was Internazionale in the 2009/2010 season (calculated in windows 4–6); a team characterized by Jose Mourinho’s stamp, playing mostly in counter attack, using one of the simplest and most known tactics in football, the numerical superiority of players. In a decade characterized by an ornate play style of the victorious Barcelona and Spanish selection, the ball possession style was considered one of the main outcomes of football. However, it is important to note that ball possession is not an important outcome. At the end of the match, it does not matter if a team has had a high percentage of ball possession if the team loses the match. Although fans would undoubtedly prefer to see their team controlling ball possession, they would obviously rather win the match. Internazionale and its coach proved that “alternative” tactics could and must be employed (including for major clubs) to surprise and win competitions. Based on the results calculated and obtained by means of DEA, Internazionale won the competition that season without wasting its resources. Looking for other performance indicators (not included in our model) of this team in 2009/2010, we could confirm a different style of play: only 45 percent of ball possession; the team that committed the most fouls and received the most yellow cards in the season; and also the one with more offsides. All these characteristics are the opposite of the norm in recent years. Most teams have tried to imitate Barcelona, and it is undeniable that they had a beautiful style of play. Fans have enjoyed watching Barcelona, especially as the team was winning. This seemed the perfect tactic, but it is not available to everyone. José Mourinho has noted this and made an intelligent and efficient use of this information and its resources.

Figure 3(a)–(g) shows results from the relation between the evolution of efficiency scores and the evolution of sports results. The evolution is calculated on the basis of the difference between the last and the first season of each analyzed window. Regular windows in this paper are composed of three seasons, so they contain 96 efficiency ratios. As we mentioned above, several clubs could not participate in the UCL on an ongoing basis. Therefore, we merely analyzed the evolution of those clubs that played in all three years in each window. From an overall view of all seven analyzed windows, we can see a clearly positive trend, except for windows 1 and 5. This trend means that clubs that improve their efficiency enhance their sports results and vice versa.

<table>
<thead>
<tr>
<th></th>
<th>Entire sample</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EffScore</td>
<td>SR</td>
<td>EffScore</td>
<td>SR</td>
<td></td>
</tr>
<tr>
<td><strong>EffScore</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson</td>
<td>1</td>
<td>0.393**</td>
<td>1</td>
<td>0.722**</td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>0.000</td>
<td></td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>768</td>
<td>768</td>
<td>96</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td><strong>SR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson</td>
<td>0.393**</td>
<td>1</td>
<td>0.722**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>0.000</td>
<td></td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>768</td>
<td>768</td>
<td>96</td>
<td>96</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** EffScore, efficiency score; SR, sports results. **Correlation is significant at 0.01 (2-tailed)**

**Table III.** Correlation between efficiency and sports results
6. Discussion of results and conclusions

The aim of this paper is to analyze efficiency and its evolution in teams that played in the UCL taking advantage of the panel data we have for nine sport seasons. This paper contributes to sports efficiency literature by presenting a research procedure to identify the most accurate methodology to be applied to panel data. First, we analyzed the

existence of a temporal trend using the S-statistic proposed by Brockett and Kemperman (1980). We have calculated the Kruskal–Wallis statistic to verify if there is stability in relative ranks. Our efficiency ratios have been calculated using super-efficiency (Andersen and Petersen, 1993), which enables us to discriminate efficient units among them.

The results of the aforementioned tests have indicated that window analysis is an accurate methodology to apply to the UCL sample. In comparison with other papers as Zambom-Ferraresi et al. (2017) and Espitia-Escuer and García-Cebrián (2010) which also calculated efficiency for football teams playing UCL, WDEA can be considered a superior methodology because it is the best adapted to the temporal evolution in ratios verified by statistical tests. Besides, this methodology enables us to assess the robustness of efficiency ratios to detect the best benchmark clubs.

To the best of our knowledge, this paper is the first empirical study on international football competitions applying WDEA to incomplete panel data. Previously, only Sala-Garrido et al. (2009) applied this methodology to eight seasons of the Spanish league, from 2000/2001 to 2007/2008. Nevertheless, to obtain a complete panel data, the authors only considered the most regular teams, those with a good performance.

Our general results show a low efficiency level in the analyzed sample: only 6 percent of the teams under analysis can be considered as efficient (with an efficiency ratio above 1) and they can be proposed as benchmarks for inefficient clubs. We have also found robust efficient teams in all stages in the championship, which opens the possibility of evaluating teams not exclusively from a sport success standpoint, but taking a combination of resources and sport results into account.

We have verified the existence of a temporal trend in efficiency. This means that technical changes occurred between the 2004/2005 and 2012/2013 seasons that led to changes in the efficiency frontier. The best example was Internazionale’s sports performance and efficiency in 2009/2010. Additionally, in our study, teams did not maintain their relative efficiency. This means that, over time, efficient units have changed. This is corroborated by the analysis of efficiency evolution in a window and efficiency in the first period of the window. We have identified that it is hard for clubs to maintain their efficiency during their time in the UCL environment. In fact, we have not found any club capable of maintaining its efficiency.

In all the stages in the competition we have found a positive and significant correlation between sports results and efficiency. However, the most remarkable, significant, positive and strong correlation between estimated efficiency scores and sports results was found in those teams that reached the semifinals of the UCL. This is an important finding for the top best teams in Europe: some simple changes in technology may represent an improvement in efficiency. This finding may assume greater significance if we consider other important findings, namely that an improvement in efficiency could also be an improvement in sports results. Even better is that this improvement could happen without the need for more inputs, which are very scarce nowadays in a football scope.

Although the correlation between efficiency and sport results is weak in teams playing in only the first stages of the championship, the observation between the evolution of efficiency and the evolution of sports results in the majority of the windows indicates that if teams improve their efficiency, they could enhance their sports results. Consequently, even though pursuing efficiency is essential for all kinds of clubs, it seems crucial for small and medium clubs competing in the field with the largest clubs in Europe. The most remarkable benchmarking observation in this regard was attained by Apoel 2011/2012.

Finally, if efficiency is positively correlated with sport results, teams could successfully change their style of play without using more resources. And the lack of continuity in efficient teams adds emotion to the championship.
Acknowledgments
The authors want to thank financial support from Project Grant ECO2016-77-P (AEI/FEDER, UE); Project UZCUD 2015-SOC-01 (Universidad de Zaragoza and CUD); Research group COMPETE S52_17 R (Autonomous Government of Aragon); and ECO2017-86305-C4-4-R (AEI/FEDER, UE). The authors also acknowledge Carlo Morosi and OPTA Sport for the provision of data. Fabiola Zambom-Ferraresi was supported by a predoctoral fellowship from the Public University of Navarre and by Health Economics research group of Public University of Navarre. The authors also thank comments of 8th ESEA European Conference on Sports Economics.

References


Corresponding author
Lucia Isabel García-Cebrián can be contacted at: lgarcia@unizar.es

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com