Predicting consumer behavior using partial least squares structural equation modeling (PLS-SEM)

In recent years, society, technology, politics, economics and many other factors have affected and changed the way consumers behave. Consumer behavior is a major topic of interest in the social sciences, particularly in economics and marketing, which emerged only about 50 years ago. Indeed, the first issue of the *Journal of Consumer Research* was published in 1974, and prior to the late 1970s there was no formal coursework in consumer behavior in colleges of business. With the growing complexity of consumer behavior theories, researchers are increasingly moving away from generalized, universalistic models to more individualized, yet multifaceted models (Codini *et al.*, 2018). For example, there is much more interest in unraveling the contingencies that characterize the differences between subgroups of individuals or environments.

Universalistic models overlook the fact that customers differ in their responses to marketing strategies and tactics, decision-making approaches and purchase motivations, which can easily lead to erroneous conclusions in consumer behavior studies. Understanding such contingencies and their effect on behavior requires a rigorous assessment of both observed and unobserved heterogeneity. Moreover, the practical utility of most marketing studies is limited to hypotheses testing of theoretical relationships embedded in a nomological net (Hair *et al.*, 2018a). The result is most conclusions are skewed toward the believability of the hypotheses being tested in the model (Rigdon *et al.*, 2020). Similarly, researchers are moving beyond the modeling of linear effects and increasingly examine complex nonlinear relationships among constructs of interest (Ahrholdt *et al.*, 2019; Braun and Hadwich, 2017; Palmer, 2010; Pehrsson, 2011; Plotner *et al.*, 2013). Inevitably, the emergence of more advanced analytical techniques and statistical methods facilitates and substantiates investigation of more complex models and analyses that provide the basis for amended findings and discussions.

Partial least squares structural equation modeling (PLS-SEM) has recently gained increasing prominence for analyzing the dynamics and complexities in consumer behavior, especially when prediction is the goal of the analysis (Hair *et al.*, 2017b, 2018b, 2019c). PLS-SEM enables researchers to bridge the concepts of explanation and prediction, because they can expect their model to have high predictive accuracy, while simultaneously being grounded in well-developed causal explanations (Sarstedt *et al.*, 2017). Our objective for this special issue is to introduce and disseminate partial least squares (PLS) path modeling as a prediction-oriented structural equation modeling (SEM) method to a wider audience with the ultimate aim of extending our understanding of consumer behavior. More specifically, we believe the papers begin to address the interrelationships between explanation and prediction in testing theoretical models focusing on consumer behavior.

**Consumer behavior and theoretical model prediction**

Until the last decade, consumer behavior researchers have mostly used covariance-based SEM (CB-SEM; also referred to as factor-based SEM) to pursue their research agenda (Hair *et al.*, 2017a, 2019a). A fundamental concern among many methodologists has been CB-SEM’s sole focus on confirmation, while being limited in terms of prediction and predictive results assessment. Applications of CB-SEM usually involve testing whether a model fits the data well, as expressed by metrics such as chi-squared test of goodness-of-fit, CFI, NFI, RMSEA and SRMR (Bagozzi and Yi, 2012). While testing a model’s fit is an important concern, a well-fitting model provides very little information about the same
model's predictive power (Shmueli, 2010; Shmueli et al., 2016, 2019). Understanding a model’s predictive power is fundamental, however, for any discipline whose goal is to derive actionable recommendations for decision-making.

Confirmation of model relationships is valuable in achieving explanatory goals and can contribute to prediction goals. It helps researchers understand the noisy processes that shape our environment and prediction facilitates better decision-making. But prediction-only research – as implied by machine learning and many business analytics applications (James et al., 2013) – entails the risk of putting too much confidence in patterns that only apply to past situations, and not to the future (Hair and Sarstedt, 2021). Explanation or prediction alone, therefore, has very limited value. Rather, social sciences researchers should focus on what Gregor (2006, p. 626) refers to as explanation and prediction theory, which “implies both [an] understanding of underlying causes and prediction, as well as [a] description of theoretical constructs and the relationships among them.”

Living up to this interplay requires researchers to rethink their methodological choices when estimating latent variable models. CB-SEM is designed solely for confirmation and is generally unsuitable for prediction purposes (Becker et al., 2013). Solely relying on CB-SEM as a method of choice is therefore disconcerting, as ensuring a model’s predictive power is a sine qua non for its practical relevance in decision-making. Moreover, the ability to produce more valid results, a standard argument for preferring CB-SEM over composite-based methods such as PLS-SEM (Hair et al., 2017c), has recently been challenged (Hair and Sarstedt, 2020). Rigdon et al. (2019a) note that while CB-SEM accounts for measurement error in factor models (Hair et al., 2017a), they potentially induce a significant degree of measurement uncertainty. This uncertainty raises doubts, therefore, about the relationship between latent variables in a statistical model and the concepts that they seek to represent, as well as any inferences based on the model estimation (Rigdon et al., 2019a, 2019b). While similar limitations apply to PLS-SEM, acknowledging this limitation makes this method in principle equally suitable for testing relationships among observed and latent variables. This is particularly relevant as PLS-SEM was designed as a “causal-predictive” method (Jöreskog and Wold, 1982, p. 270) that overcomes the apparent dichotomy between explanation and prediction (Chin et al., 2020; Hair et al., 2020).

Beyond the choice of a particular SEM method, researchers need to rethink their use of model evaluation metrics. In the past, the metrics too often focused only on model fit measures, or in the context of PLS-SEM, metrics designed to assess a model’s explanatory power, such as the $R^2$ (Hair et al., 2019b, 2019c). Looking forward, and recognizing the value of both explanation and prediction, it is essential for researchers to begin applying available out-of-sample prediction measures derived from using holdout samples (Sarstedt and Mooi, 2019). Possibilities include k-fold cross validation-type procedures such as PLS$_{\text{predict}}$, which Shmueli et al. (2016) have proposed in the context of PLS-SEM (Shmueli et al., 2019), as well as model selection metrics such as BIC and GM, which excel in striking a balance between model fit and prediction for PLS path models (Danks et al., 2020; Sharma et al., 2019a, 2019b). Finally, Liengaard et al.’s (2020) cross-validated predictive ability test (CVPAT) provides an overall inferential test for predictive model comparison designed to assess whether an alternative competing model achieves better out-of-sample predictive power than another established model.

The previous lack of focus on accurate prediction of theoretical model outcomes by consumer behavior researchers dictates that much greater emphasis must be placed in the future on validating the true predictive accuracy of models. Goodness-of-fit metrics recommended by some methodologists to evaluate PLS-SEM results (Benitez et al., 2020; Schuberth et al., 2018; Tenenhaus et al., 2005) are not a substitute for out-of-sample predictive metrics and, thus, cannot provide insights about the theoretical model’s predictive
power. Stated simply, a well-fitting model is not necessarily an accurate predictive model. The interplay between causal-predictive methods and corresponding metrics is an important step toward further improving the consumer behavior field’s relevance for practice, while at the same time maintaining scientific rigor. A successful balance between explanation and prediction lends authority, therefore, to our understanding of how the study of consumer behavior relates to the fundamental quest of science (Dublin, 1969). Shifting the focus from explanation-only to explanation and prediction will close the gap between theory and practice and contribute to making research more relevant for real-world applications.

While prediction is at the top of priorities for the role of PLS-SEM and consumer behavior research, we would be remiss if we did not also highlight other recent methodological developments in PLS-SEM that some scholars may not yet be aware of – and there are quite a few [e.g. see also Table 1 in Ghasemy et al. (2020)]. The primary PLS-SEM methodological developments we recommend consumer behavior researchers become familiar with include: discrete choice modeling – an option for which PLS-SEM can be used to analyze stated preference data generated through choice experiments (Hair et al., 2018c); endogeneity – a systematic procedure to effectively address endogeneity concerns in PLS-SEM analyses (Hult et al., 2018); multiple mediation and moderation – the preferred approach for these types of analyses is PLS path modeling (Sarstedt et al., 2020); confirmatory composite analysis – a process similar to confirmatory factor analysis that can be followed with PLS-SEM to confirm both reflective and formative measurement models of established measures being updated or adapted to a different research context (Hair et al., 2020; Schuberth et al., 2018); and the CVPAT – it enables researchers to conduct pairwise comparisons of the predictive power of competing theoretical models (Liengaard et al., 2020). In addition, the combination of PLS-SEM with the necessary condition analysis (Richter et al., 2020) and agent-based simulation (Schubring et al., 2016) further expand the portfolio of useful methods consumer behavior researchers can exploit. Familiarity with and application of these methodological tools will ensure consumer behavior researchers fully explore the possibilities of their empirical research.

**Observations on the special issue articles**

Interest in and execution of consumer behavior research has grown substantially in the past 50 years. But much of the research has attempted to explain and confirm relationships between attitudinal and perceptual concepts, and occasionally behavioral opinions, in contrast to prediction, particularly out-of-sample predictions that are useful in generalizing from a sample to the population. We believe these special issue papers are an excellent first effort in demonstrating the importance of adding out-of-sample prediction to the consumer behavior researchers’ toolbox.

The first article, “Psychological ownership in social media influencer marketing,” by Mandy Pick, focuses on the rapidly emerging communication strategy of social media-based influencer marketing. The research examines the impact of consumers’ perceived influencer credibility using the source credibility model, attitudes toward advertising and product, psychological ownership and, ultimately, purchase intentions. The out-of-sample PLS_{predict} methodology demonstrates the value of social media influencers in predicting consumer purchase intentions and ultimately developing more effective marketing strategies.

The second article, “Brand image as the competitive edge for hospitals in medical tourism,” by Tat Huei Cham, Boon Liat Cheng, Mei Peng Low and Jason Boon Chuan Cheok, explores the role brand reputation plays in determining a competitive strategy. Specifically, their research confirms the impact of social (e.g. social media and word-of-mouth communications) and marketing tactics (e.g. hospital advertisement and price perception) on the brand image of medical tourism-based hospitals, and the relationship of these tactics on perceived service quality delivery. Results of a multiple mediation model whose outcomes confirmed high out-of-sample predictive power...
demonstrate the importance of branding for medical tourism. The findings are also relevant for scholars and practitioners involved in developing both regional and global brand strategies.

The third article, “Unbundling subjective career success: A sequential mediation analysis,” by Zubeida Rossenkan, Wei-Chan Au and Pervaiz Khalid Ahmed, examines subjective career success (SCS) using sequential mediation modeling to explore the inter-relationships between types of SCS, including interpersonal, financial, job and hierarchical success. The results show that an individual’s interpersonal success provides a foundation for accomplishing job tasks (job success), which then leads to increased prospects for promotion (hierarchical success) and, ultimately, financial success. The findings provide a nuanced understanding of career behavior among young adults from the perspective of a non-western developing country context, and demonstrate the value of combining the goals of model explanation and accurate out-of-sample prediction metrics in understanding consumer behavior.

The fourth article in this issue, “Exploring consumer–brand engagement: A holistic framework,” by Man Lai Cheung, Guilherme Pires and Philip Rosenberger III, investigates the relationships between consumers’ enduring involvement, ongoing information search behavior, online engagement behavior, consumer–brand engagement (CBE) and, ultimately, brand attitudes in Hong Kong. The findings confirm that the predictive power of these antecedents ultimately enhances customers’ brand attitudes, providing a better understanding of how to strengthen CBE for durable technology products, such as smartphones. The results suggest marketers should seek to heighten customers’ involvement levels by encouraging customer–brand interactions, which are useful not only in encouraging customers’ ongoing search and online engagement behavior, but also critical in strengthening CBE. The results also expand the conceptual underpinnings of CBE and confirm the value of better metrics for evaluating theoretical model predictive power.

The fifth article, “Predicting mobile network operators’ users m-payment intentions,” by Choi-Ming Leong, Kin-Lim Tan, Chin-Hong Puah and Shyh-Ming Chong, empirically tests a theoretical framework that proposes perceived usefulness and ease-of-use mediate the relationship between perceived compatibility and intention to use m-payment systems. The potential to expand m-payment services to other e-wallet platforms is also explored as well as the influence of features such as perceived security and personal innovativeness on the usage behavior of mobile payment services. The findings confirm the value of improving our ability to predict emerging consumer behavior patterns and also extend the limited literature on technology and adoption of services.

The sixth article, “The role of competitive strategy in the performance impact of exploitation and exploration quality management practices,” by Julen Castillo Apraiz, Nicole Franziska Richter, Jesus Matey de Antonio and Siegfried Gudergan, advances our understanding of quality management practices by clarifying how the chosen competitive strategy in the German pharmaceutical market alters the impact of exploitative and explorative quality management practices on performance. Exploitative and explorative quality management practices are clearly related to firm performance and their impact depends on the competitive strategy pursued. Specifically, explorative quality management practices are more relevant for firms following a differentiation strategy, whereas exploitative quality management practices are more relevant for cost leaders. Finally, for strategically ambidextrous firms that simultaneously follow both cost and differentiation approaches, the interplay of the two quality management practices influences the outcome.

The last article, “Assessing formative artscape to predict opera attendees’ loyalty,” by Berta Tubilejas-Andrés, Amparo Cervera-Taulet and Haydee Caleron Garcia, develops a formative higher-order construct (Sarstedt et al., 2019) to measure consumers’ perceptions of a servicescape of performing arts services. Their measurement combines physical (exterior and interior) and
Final thoughts
As evident by this brief overview of the diverse articles in this special issue, explaining and predicting issues within consumer behavior is a thriving and ongoing effort among marketing scholars. As Kaplan (1964, p. 350) notes, “If we can predict successfully on the basis of a certain explanation we have a good reason, and perhaps the best sort of reason, to accept the explanation.” Among our original goals for this special issue was to include papers that illustrate how the proposed advances of the original PLS-SEM method are practically relevant for predicting consumer behavior phenomena. We believe you will agree the articles in this special issue will trigger substantial interest in further advancing the prediction of consumer behavior and inspire exciting follow-up research.

We would like to thank the Editor of European Business Review, Göran Svensson, for giving us the opportunity to prepare this special issue. It was a long and arduous process for everyone, including the authors who stuck with us through numerous rounds of revisions during this pandemic time. Importantly, we would also like to thank the many reviewers, without whom this special issue would not have been possible. Many scholars had to work together to enable what we think is a powerful contribution to better understanding the role of prediction in consumer behavior.

Finally, the accepted manuscripts were submitted and reviewed through the submission system of Emerald Publishing and were selected by a blind peer-review process to ensure their relevance and quality. Moreover, we did our best to not desk reject manuscripts and provided authors an opportunity to resubmit as a new submission whenever realistic. As a result, seven of 27 revised manuscripts were accepted for publication in the special issue, and a number of other papers were selected to appear in later issues of this journal.

Joe F. Hair
Department of Marketing and Quantitative Methods, University of South Alabama, Mobile, Alabama, USA

Jun-Hwa Cheah
Department of Management and Marketing, School of Business and Economics, Universiti Putra Malaysia, Serdang, Malaysia

Christian M. Ringle
Institute of Human Resource Management and Organizations, Hamburg University of Technology, Hamburg, Germany and Waikato Management School, University of Waikato, Hamilton, New Zealand

Marko Sarstedt
Faculty of Economics and Management, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany and Monash University - Malaysia Campus, Bandar Sunway, Malaysia, and

Hiram Ting
Faculty of Hospitality and Tourism Management, UCSI University, Kuala Lumpur, Malaysia and Department of Leisure and Recreation Management, School of Tourism, Ming Chuan University, Taoyuan, Taiwan
References


**Further reading**


