Board diversity and firm innovation: a meta-analysis

Teemu Makkonen
Karelian Institute, University of Eastern Finland, Joensuu, Finland

Abstract
Purpose – It is commonly stated that increased board diversity leads to the heightened financial performance of firms via the impact that it can have on innovation, but the latter association has, thus far, remained empirically controversial. The aim of this paper is to shed light on this unresolved debate and gap in the literature via studying different types of diversity.
Design/methodology/approach – A meta-analysis was conducted on the existing empirical evidence on the topic to show whether such an association exists and compare cognitive (expertise and experience) and demographic diversity (gender, nationality and racial/ethnic).
Findings – The results show that there is indeed a positive and statistically significant association between board diversity and firm innovation. This association is driven more by cognitive diversity of the board members than by demographic diversity.
Research limitations/implications – Potential publication bias, heterogeneity in the quality of the existing studies and the diversity in operationalising innovation and board diversity remain as limitations to this meta-analysis.
Practical implications – Instead of focussing on selecting board members based on demographic (surface-level) diversity, selections should be based on the interplay of the experience, expertise and background demographic characteristics of the potential candidates. Otherwise, the minority members might face a “token” status.
Originality/value – The results of this paper suggest that there is a positive association between board diversity and firm innovation. Future research should examine why this link exists. Therefore, the paper concludes with a research agenda for the benefit of potential further studies.
Keywords Board diversity, Board of directors, Innovation, Meta-analysis
Paper type Research paper

1. Introduction
The discussion on the benefits of board diversity [1] (BD) has revolved to a large extend around the theorem of “the business case for value in diversity”: when it comes to performance, a group that is heterogeneous in their backgrounds should encompass a broader range of network ties (bringing in additional information and outside resources), expertise, knowledge and perspectives compared to homogeneous groups of actors, thus providing the diverse group a competitive advantage (Cox et al., 1991; Robinson and Dechant, 1997). Moreover, if the talent of certain segments of society are systematically excluded from board of directors (BoD) due to, for example, race and gender rather than their actual competence, these BoDs will not function optimally (Torchia et al., 2011). Taken together, these notions imply that BD should have a positive impact on the financial performance of firms. Further, the existing corporate governance literature hypothesises that this is because BD first heightens firm innovation [2], which in turn will result in positive financial outcomes (Miller and Triana, 2009; Carter et al., 2010; Gala and Zenou, 2012). While the latter link is well-established in the literature on corporate governance, less is known about the former.

© Teemu Makkonen. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/legalcode

European Journal of Innovation Management
Emerald Publishing Limited
1460-1060
DOI 10.1108/EJIM-09-2021-0474
In fact, while BD has been discussed widely in the corporate governance literature in relation to the financial performance of firms (Adams et al., 2015), there are still relatively few studies that go beyond these broad performance measures by empirically linking specific aspects of BD and firm innovation together (Torchia et al., 2011; Cook and Glass, 2015).

A recent meta-analysis by Sierra-Morán et al. (2021) points to a positive connection between demographic diversity and firm innovation. However, the notions of the positive impacts of different types (demographic and cognitive) of BD on firm innovation have not always been supported with empirical evidence (Wincent et al., 2010; Zona et al., 2013). As such, there is still no solid consensus on the actual impacts—be it positive or negative—of BD on firm innovation (Ariff et al., 2017) nor which types of BD are expected to be the most beneficial for firm innovation. Therefore, the aim of this paper is to shed light on this topic in the corporate governance literature with the help of meta-analysis. While not without its limitations, meta-analysis is a commonly utilised statistical method for summarising the results of several earlier works. The main research questions of this paper are:

**RQ1.** Is BD associated with heightened firm innovation?

**RQ2.** Which types of BD matter for firm innovation?

The remainder of this article is structured as follows. First, the most commonly utilised theoretical explanations of the benefits and potential drawbacks of BD vis-à-vis firm innovation are discussed followed by a synthesis of the various ways that innovation and different types of BD are operationalised into measurable items in the empirical literature together with an overview on the existing research gaps. Second, the search procedures and methods (meta-analysis) applied in this paper are introduced followed by the results section providing a statistical consensus on the “BD–firm innovation” debate. Lastly, a conclusions chapter sums up the main implications of the results and discusses the limitations of the chosen approach.

### 2. Literature review

#### 2.1 Theoretical issues

Starting from the upper echelon theory (Hambrick and Mason, 1984) indicating that background characteristics of top management teams have a significant impact on corporate decision-making (Liao et al., 2019; Nadeem et al., 2019), several theories have been utilised when motivating the anticipated (positive) impact that BD might have on firm innovation. Of these theories, agency theory (Ross, 1973; Mitnick, 1975; Jensen and Meckling, 1976) is among the most commonly utilised. Research approaching the issue of BD and firm innovation from the aspect of agency theory has focused on the monitoring role of the BoD (Galia and Zenou, 2012). According to agency theory, promoting innovation requires a strong BoD: innovation is always risky, since it includes a degree of uncertainty and investments in innovation require a long-term perspective on potential pay-backs/pay-offs. Therefore, risk-averse executives can purposefully reduce investments in research and development (R&D) and other innovation activities. This might naturally go against the long-term interests of the shareholders. Thus, the BoD functions as an important information system to monitor the behaviour of executives (Ruigrok et al., 2006; Zona, 2014). From the agency theory perspective, whereas homogeneous BoDs are associated with risk-aversion and conservatism (Stiles, 2001; Carter et al., 2010; Margetts and Holland, 2015), diverse BoDs promote firm innovation—via their role as advisors—by bringing in their alternative perspectives (human capital) leading to stronger BoDs that reduce information bias, tolerate risks and potential (short-term) losses and consistently support executives’ innovation initiatives (Ruigrok et al., 2006; Zona et al., 2013). In relation to the latter point, also the stewardship theory (Donaldson and Davis, 1991)—while differing from the agency theory, for example, with regard to the
behaviour of managers—suggests that BD enhances innovativeness through the sharing of different perspectives in board meetings that facilitate the work of the executives (Zona, 2014).

Another widely applied theoretical stance towards BD and firm innovation within the corporate governance literature is resource dependence theory (Pfeffer, 1972; Pfeffer and Salancik, 1978), which moves beyond the owner–manager–board relationship discussed in the agency theory and directs the focus of research to the environment of the firm. Its premise is that firms are dependent on their external environments, which are characterised by a significant degree of uncertainty (Ruigrok et al., 2006). Since a diverse BoD (should) encompass a wide set of skills, know-how, experiences, information, knowledge and contacts, the BoD can be considered a potential provider of useful resources that can reduce this uncertainty and promote firm innovation (Martini et al., 2012). Therefore, the main argument of resource dependence theory vis-à-vis BD is that a diverse BoD helps the firm to acquire critical resources, such as knowledge, legitimacy and networks (Ruigrok et al., 2006). A diverse BoD will provide the firm with resources that promote generating new ideas, reducing narrow-mindedness, understanding markets, finding novel opportunities and improving resource allocation, which can all facilitate firm innovation (Galia et al., 2015; Ariff et al., 2017).

In accordance with resource dependence theory, the behavioural theory of the firm (Cyert and March, 1963; Cox et al., 1991) expects that BD will lead to innovative ideas. Thus, the behavioural theory of the firm suggests that the more diverse the BoD is, the more comprehensive the information available for decision-making, and consequently the more innovative the decisions it makes will be (Miller and Triana, 2009). As such, although the above theories approach the link from BD to firm innovation from a different perspective, they all support the notion of a positive association between the two:

**H1a.** There is a positive association between BD and firm innovation.

The critical mass (or token) theory (Kanter, 1977) states, however, that individual minority board members are unlikely to affect firm innovation alone without the presence of other minority leaders. In other words, individual minority leaders may suffer from a token status. This limits their ability to have an impact on organisational outcomes, since minority board members may face (extensive) scrutiny and negative bias from the other board members, receive less support from their peers and turn to conformity to avoid conflict in boardrooms. Increasing the absolute number (or relative representation) of minority board members over a certain threshold (critical mass) may help to overcome the shortcomings of a negative bias and isolation (Torchia et al., 2011; Cook and Glass, 2015; Rossi et al., 2017). Thus, according to the critical mass theory one would expect that the impact of BD on firm innovation is (statistically) insignificant without a strong presence of minority board members:

**H1b.** The association between BD and firm innovation is (statistically) insignificant.

As put forward by Bhidé (2000), there are both positive and negative effects of heterogeneity (diversity) suggesting that the inter-relationship between BD and firm innovation might be curvilinear rather than linear. This notion is supported by the literature on cognitive and knowledge distance (Nootenboom et al., 2007) stressing that there needs to be an “optimal mix” of diversity in a team for it to function optimally. Optimal cognitive distance is reached when the people involved share enough similarities through which a common understanding can be created, but also differences between each other that enable new combinations of knowledge, ideas and learning. Too much similarity means a narrower base for new ideas, while too much diversity induces problems in mutual understanding (Hautala, 2011).

As such, BD can also lead to negative impacts, such as communication problems, reduced cohesion, increased conflict, delays in reaching consensus and slower reaction times, which can hinder efficient decision-making (Bernile et al., 2018). Research on group dynamics suggests that people might hold prejudices against people who are different from themselves.
Due to these difficulties in integrating all board members into an effective and harmonised group (Erhardt et al., 2003; Huse, 2007; Piekkari et al., 2015), it can take time for board members to get over their interpersonal differences (Milliken and Martins, 1996). From the perspective of cognitive and knowledge distance and group dynamics, it seems that rather than having a positive impact, the effects of BD on firm innovation are negative:

\[ H1c. \] There is a negative association between BD and firm innovation.

### 2.2 Measuring board diversity and innovation

In order to statistically test the association between BD and firm innovation one naturally needs to, first, operationalise firm innovation and BD into quantifiable indicators.

When it comes to indicator choices for depicting innovation, the literature on BD largely follows the general literature on innovation studies: it commonly utilises well-known proxy indicators (with well-known limitations) for innovation, such as R&D expenditure and patents (Li, 2019), which does not necessarily do justice to the complexity of the concept of innovation. Therefore, another commonly utilised operationalisation of innovation is to apply questionnaire items, where firms themselves are asked to evaluate their innovative activities (Galia and Zenou, 2012). This approach also allows the construction of composite innovation indicators containing information from several innovation-related questions (Torchia et al., 2011), such as different innovation types: product, process, service, organisational or marketing innovation (OECD, 2005).

In the literature on BD, there is a common understanding that, rather than being randomly distributed, the cognitive functioning and beliefs of the population tend to vary systematically according to demographic background variables (surface-level diversity) (Robinson and Dechant, 1997). Therefore, race (or ethnicity), nationality (foreign board members) and gender—not least because of the quota systems introduced in several countries to promote gender diversity in BoDs (Kang et al., 2007; Vinnicombe et al., 2009)—are among the most commonly utilised proxies of the different perspectives that individuals can bring to an organisation (Hillman et al., 2002). Alternatively, it has been proposed that, since diverse sets of knowledge and skills (cognitive diversity) are beneficial for innovation, the know-how of the board members should be taken into account, rather than surface-level diversity, when exploring the connection between BD and firm innovation. This aspect of BD is commonly operationalised through variables related to expertise (educational and occupational diversity) and experience (diversity in the age and/or tenure of the board members) (Wincent et al., 2010; Kim and Kim, 2015; Midavaine et al., 2016). Based on this discussion, the following hypothesis is tested here:

\[ H2. \] The association between cognitive diversity and firm innovation is stronger than between demographic (surface-level) diversity and firm innovation.

There are several reasons why certain aspects need to be controlled for in statistical analyses on the influence of BD on firm innovation, of which the most commonly applied measure is board size. Earlier studies have shown that board size is related to innovation (Prencipe, 2016). But here also the evidence is inconclusive whether smaller or larger BoDs might have the advantage in terms of innovation: smaller BoDs might not have sufficient knowledge and expertise to direct firms’ innovative activities, whereas larger BoDs might become dysfunctional and, thus, hamper firm innovation (Zona et al., 2013). Based on this inconsistency in the earlier BD literature, two opposing hypotheses are tested here:

\[ H3a. \] The association between BD and firm innovation is stronger the smaller size of the BoD.

\[ H3b. \] The association between BD and firm innovation is stronger the larger size of the BoD.
2.3 Research gaps
The data and methods applied in the existing literature on BD and firm innovation have resulted in many unaddressed research gaps. First, while it is likely that innovative firms appoint more diverse BoDs than non-innovative firms (Makkonen et al., 2018), the existing studies have largely ignored the issue of reverse causality between firm innovation and BD (Miller and Triana, 2009; Chen et al., 2018). Also, the role of moderating factors that affect the relationship between BD and firm innovation—such as board independence and activities, CEO's characteristics, industry, etc.—and mediating factors—such as R&D productivity, reputation, etc.—that link BD to innovation deserve more attention in the future.

Second, apart from only a few recent exceptions (Cabeza-García et al., 2021; Griffin et al., 2021), the vast majority of the articles included in the meta-analysis had a single-country perspective. However, it is likely that there are country-specific institutional (Terjesen et al., 2015) and cultural differences (Mukarram et al., 2018) and differences in corporate governance systems (John and Senbet, 1998) affecting the link between BD and firm innovation. Therefore, comparative approaches (international comparisons) should be employed to investigate the potential moderating role of these differences.

Third, primary data collection methods are needed for designing more fine-grained measures of BD that would go beyond the existing demographic and cognitive diversity divisions. As criticised by Torchia et al. (2011), particularly the contemporary focus on demographic diversity (cf. Appendix 1) is ill-placed (see also Li and He, 2021). Instead, future studies should be encouraged to focus their attention more on the actual differences (in experience and expertise) between board members. Therefore, a more nuanced interpretation of BD beyond the use of single categories as proxy indicators for underlying differences in board members’ backgrounds would give a better picture of the composition factors that affect the influence of the BoD on firm innovation.

Finally, as argued by Kim and Kim (2015) there seems to be a trade-off point for BD: firms can benefit from BD as long as it does not exceed a certain level, after which the resources that the diverse board members bring with them are counterbalanced by the negative impacts that (excessively) heterogeneous BoDs commonly share in terms of inefficiency, lack of consensus and prolonged decision-making time. This proposed U-shaper relationship between BD and firm innovation has been mainly investigated under the critical mass theory in relation to gender diversity (Ain et al., 2021; Cabeza-García et al., 2021; Saggese et al., 2021). Given the literature on cognitive and knowledge proximity, it would seem more than worthwhile to test whether the critical mass theory applies also to other types of BD.

The above gaps in the empirical literature can be partly explained by an overreliance on single dominant theories. The benefits of adopting theoretical pluralism have been increasingly recognised as there is no one best way to make decisions in a boardroom. Rather, the optimal way of organising a firm is dependent on varying internal and external conditions. In short, BD is expected to exert differing impacts on firm innovation depending on the specific internal organisational contexts and external environments (Zona et al., 2013). Theoretical discussion on BD should explore this pluralism to provide an integrated theoretical framework for studying the impacts of BD on firm innovation.

3. Data and methods
3.1 Meta-analysis
The high volume of papers investigating the link between BD and firms’ (financial) performance has also spurred a number of meta-analyses on the topic—commonly confirming the positive connection between the two (Pletzer et al., 2015; Post and Byron, 2015; Anastasia et al., 2020). Contrarily, the theoretical literature on the link between BD and firm innovation is rather undecided on whether BD has a positive or a negative impact on firm innovation.
Also, the conducted empirical literature repeats this uncertainty: even the most recent studies on the topic still show mixed evidence with either negative, positive or insignificant results (e.g. Hernández-Lara and Gonzales-Bustos, 2019, 2020; Gonzales-Bustos et al., 2020; Liu et al., 2020; Manita et al., 2020; Martinez-Jimenez et al., 2020; Rejeb et al., 2020; Vaivaran and Zhang, 2020; Jhunjhunwala et al., 2021; Luo et al., 2021; Zhong et al., 2021). Therefore, meta-analysis methodology seems particularly suitable for investigating this uncertainty. There are natural limitations to meta-analysis mainly due to the facts that it weighs all studies equally although there is substantial variation in the quality of the included articles (due to, for example, different standards set by journals) and due to the heterogeneity between the included studies in terms of, for example, operationalisation of the variables. The implications of these limitations will be further discussed in Section 5.2. The benefit of meta-analysis is that it can include all the reported effects of previous studies in a single statistical synthesis combining the results of earlier works into a coherent summary figure (Borenstein et al., 2009). This aggregation of information leads to higher statistical power and more robust results than is possible in any individual study (Cohn and Becker, 2003). Therefore, following an approach similar to the one employed by Sierra-Morán et al. (2021) who have investigated the relationship between demographic diversity and firm innovation, a meta-analysis on the impact of BD on firm innovation was performed here to derive a summary correlation measure indicating whether the association is positive or negative (H1), whether cognitive diversity matters more for firm innovation than surface-level diversity (H2) and whether the size of the BoD has an impact on the results (H3).

When undertaking a meta-analysis, the researcher must decide between using fixed or random effects (Borenstein et al., 2009). In the fixed effects model, the studies in the meta-analysis are assumed to constitute the entire universe of studies and, thus, the model only uses within-study variability in the error term. In the random effects model, on the other hand, studies in the meta-analysis are assumed to be only a sample of all possible studies and, thus, the error term includes both within-study variability and variability arising from differences between studies. The random effects model therefore results in much larger standard errors and more conservative significance tests than in the case of the fixed effects model (Field, 2001). As such, the fixed effects model can overestimate the degree of precision in meta-analysis findings (Hunter and Schmidt, 2000), whereas the random effects model facilitates “unconditional” inferences (i.e. inferences that can be generalised beyond the studies included in the meta-analysis). Within the social sciences, researchers typically wish to generalise beyond the sample of included studies, and therefore the random effects model is often more appropriate (Field, 2001) and, thus, also applied in this study (technical details are provided in Appendix 2). The analysis was conducted using Meta-Essentials (see Suurmond et al., 2017) downloadable from: https://www.erim.eur.nl/research-support/meta-essentials/download/.

The calculated overall mean weighted correlations are further subdivided into two categories based on the type of diversity investigated in order to conduct a subgroup analysis: 1) demographic (gender, nationality and racial/ethnic) and 2) cognitive (expertise and experience) diversity. The decision to conduct subgroup analysis is supported by the I² statistic for heterogeneity: 89.3% (see Higgins et al., 2003).

The moderating role of board size is tested with standard regression analysis. The general idea behind examining the role of moderators with meta-analysis tools is, however, to use the results for exploration rather than for statistical testing (ERIM, 2021). Therefore, the result of the moderator analysis can only be interpreted to show tentative support (at best) and would need to be confirmed via more robust analysis with greater statistical power.

Ultimately, it must be noted, that it is quite likely that a number of studies reporting insignificant results have remained unpublished in academic journals. This averseness towards (submitting and) publishing statistically non-significant results is commonly termed as “publication bias” or the “file-drawer problem” (Rosenthal, 1979). A funnel plot is utilised here as a simple-to-use generic mean to estimate the potential existence of publication bias.
While there are naturally a number of alternative reasons for asymmetric funnel plots, if the observed effect sizes are not symmetrically distributed around the combined effect size, this is taken as (likely) evidence of the existence of publication bias (Sterne et al., 2005).

3.2 Sample delineation
The data (covering 15 years: papers published between 2006 and July 2021) were gathered from Elsevier’s Scopus database. The focus was set on scientific journal articles within the subject areas that are most likely to analyse the topic at hand: 1) Business, Management and Accounting; 2) Social sciences; 3) Economics, Econometrics and Finance; 4) Decision sciences and 5) Arts and Humanities (plus multidisciplinary). A word search procedure was used to identify articles with the following terms (that are typically used to describe the topic at hand) appearing in the title, abstract or keywords:

(1) (diversity OR composition OR women OR female) AND board AND innovat*

Initially, 142 scientific journal articles were found (in July 2021). The articles were carefully assessed and those that do not report statistical empirical findings in relation to BD and innovation or do not address firms were excluded. Further, information on the correlation coefficient (r) between BD and firm innovation as well as information on the sample size (n = number of firms included) of the individual studies are needed to run the meta-analysis. Therefore, all the different types of BD–firm innovation linkages analysed in the papers of the sample with sufficient information were included in the meta-analysis (Appendix 1). The search and screening procedures are presented in Figure 1.

The resultant 29 articles (Appendix 1) and the 50 utilisable correlations (effect sizes) in these papers selected for the meta-analysis are, naturally, a sample of English language literature on the topic that can be detected with the chosen keywords from Scopus: there are likely to be articles on the topic on other languages and in journals/books not covered by Scopus.

<table>
<thead>
<tr>
<th>DATABASE: SCOPUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEARCH TERMS:</td>
</tr>
<tr>
<td>(diversity OR composition OR women OR female) AND board AND innovat*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FILTERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title, abstract and keywords</td>
</tr>
<tr>
<td>Subject areas¹</td>
</tr>
<tr>
<td>Years: 2006–2021²</td>
</tr>
<tr>
<td>Language: English</td>
</tr>
<tr>
<td>Document type: Articles</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SCREENING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address firms</td>
</tr>
<tr>
<td>Address innovation</td>
</tr>
<tr>
<td>Address BD</td>
</tr>
<tr>
<td>Statistical evidence</td>
</tr>
<tr>
<td>Sample size reported</td>
</tr>
<tr>
<td>Correlation reported</td>
</tr>
</tbody>
</table>

| Number of articles: 142 |
| Number of articles: 29 |
| Number of effect sizes: 50 |

Note(s): 1) Business, Management and Accounting, Social sciences, Economics, Econometrics and Finance, Decision sciences, Arts and Humanities and multidisciplinary; 2) July 2021

Figure 1.
The utilised search and screening procedures
database. Still, this relatively low number of publications is rather surprising when taking into account how often the claim has been made in the earlier corporate governance literature that BD drives the financial performance of firms exactly through its impact on innovation (Carter et al., 2010). Is the claim then actually supported by the empirical literature? This question is approached here through a meta-analysis.

4. Results
By applying the statistical power calculator (an R script and excel file) for meta-analysis developed by Quintana and Tiebel (2018)–available at: https://osf.io/5c7uz/–and taking into account the presence of high heterogeneity—as shown by the $I^2$ statistic: 89.3%—the calculated statistical power of the conducted meta-analysis is 0.96. This is well above the conventional minimum acceptable statistical power for meta-analyses—which has, following Cohen (1992), commonly been set to 0.80—indicating that the utilised sample is “large enough” for the purposes of meta-analysis.

The results show that BD does indeed have a small but statistically significant ($p < 0.01$) positive association with firm innovation (Figure 2). According to the random effects model, the overall mean weighted correlation between BD and firm innovation is 0.04 (confidence interval: 0.02, 0.07). This result supports H1a postulated based on the commonly utilised agency theory, resource dependence theory and behavioural theory of the firm. In terms of the different BD types, whereas 69% of the included studies considered exclusively demographic diversity, only 28% of them considered cognitive diversity, either exclusively (four studies) or in conjunction with demographic diversity (four studies) (cf. Appendix 1). One study did not differentiate between the two BD types and was used only for testing H1. However, from Figure 2 we can observe that the association between BD and firm innovation is driven more by cognitive rather than demographic diversity: the overall mean weighted correlations are 0.23 and 0.02, respectively (95% confidence intervals: 0.09, 0.37 and −0.01, 0.05 respectively). Thus, it seems that the association between BD and firm innovation is clearer in the case of cognitive diversity than in the case of demographic diversity. The results, thus, support H2.

The result of the moderator analysis points towards a conclusion that board size is negatively associated with the correlation coefficients between BD and firm innovation (Figure 3). The result suggests that the association between BD and firm innovation is stronger in firms with smaller BoDs. However, due to the limitations of the moderator analysis (see Section 3.1) this result can be interpreted to show only tentative support for H3a.

The funnel plot presented in Figure 4 shows that there is evident publication bias also in relation to BD and firm innovation; the observed effect sizes are not symmetrically distributed around the combined effect size (indicated by the imputed data points). Thus, it is very likely (or almost certain as is the case with most meta-analyses) that the applied search procedures have not identified all potential contributions and that the utilisation of a single publication database (Scopus) leaves a number of publications with relevant findings outside the sample presented here. With the help of Google Scholar, it is indeed possible to identify such (working) papers, which generally show insignificant or contradictory impacts of BD on firm innovation.

5. Conclusions

5.1 Contributions and academic and practical implications
The corporate governance literature has generally assumed that BD positively affects the financial performance of firms via its impact on heightening firm innovation. However, the existence of this link between BD and firm innovation has remained without consensus as there are empirical studies supporting as well as studies opposing this hypothesis. The results of the meta-analysis (based on the results of earlier empirical research on the topic)
point to a positive and statistically significant association between BD and firm innovation. Therefore, having diversity in the boardroom seems to positively affect firm innovation. A subgroup analysis further revealed that this association is more likely to be driven by...
cognitive than demographic diversity. Therefore, while the majority of the studies on BD and firm innovation have focused precisely on demographic background characteristics of board members, this type of diversity seems to be much less relevant for firm innovation than cognitive diversity.
However, as stated by Krawiec et al. (2013), while it is easy to accept the statement that diversity is beneficial, it is much harder to pinpoint exactly how and why it is beneficial. Therefore, further studies are called for in order to disentangle, for example 1) the impact of diversity on the mechanisms of BoD selection; 2) the impact of diversity in corporate governance on the actual idea-generating process within firms; and 3) the interaction between BD, CEOs and other members of the top management team.

Finally, an important practical implication of this meta-analysis for policymakers and managers when drafting policies that affect board composition or appointing board members is that also diversity types that go beyond the surface-level should be at the forefront of such decisions: the interplay between demographic background variables and cognitive diversity should be taken into account. Simply appointing minority board members for the sake of increasing (surface-level) demographic diversity might lead them to be viewed as “tokens” and, thus, hamper their potential to make a significant contribution to decision-making in the BoDs.

5.2 Limitations
Caution is always needed when interpreting the results of meta-analyses. As is common when conducting meta-analyses, a number of studies, particularly those reporting statistically insignificant results can remain undetected by standard search procedures in publication databases, because these types of results are usually harder to publish in academic journals (publication bias). Therefore, the reported overall mean weighted correlations are likely to be higher and the direction of the link clearer than would be the case if these undetected studies were included in the sample. As shown by the publication bias analysis and subsequent additional searches, this applies also here—since the funnel plot predicts there are indeed unpublished studies reporting contradictory or insignificant results on the topic. Further, the way that the applied meta-analysis utilises sample sizes to dedicate weights does not take into account the heterogeneity in the quality of the papers: an included “lower quality paper” with a larger sample will have a larger impact on the reported overall mean weighted correlations than an included “high standard” article.

Another point for caution is due to the heterogeneity of the included studies—notably, the varying ways that innovation and BD have been operationalised: the effect sizes in the conducted meta-analysis do not strictly measure the same types of innovation nor use similar indicators for the different types of diversity. Further, moderator analyses can only give tentative evidence for exploratory investigations rather than work as rigorous statistical tools to confirm relationships between variables. These issues naturally limit the generalisability of the results presented here. As such, while this meta-analysis has indicated a positive general picture of the relationship between BD and firm innovation, the results cannot be considered as a rule. The results reflect average effects and, thus, in some contexts the costs can outweigh the benefits: BD can also have negative impacts (reduced cohesion, communication problems, etc.) and the impact of BD is likely to vary on a case-by-case basis depending on the firm, the type of BD, the type of innovation, etc.

Finally, making strong statistical inferences based on small samples can be misleading. It is important to state that the topic under review here has not been studied as extensively as, for example, the link between BD and the financial performance of firms. Thus, while the statistical power of the meta-analysis is methodologically “sufficient”, the size of the sample remained admittedly quite small. This, however, is a result in itself: despite commonly made proposition about the role of BD in driving the financial performance of firms through its impact on innovation, there is a general lack of empirical studies justifying, particularly comparative studies utilising different BD types, and explaining the rationale behind these claims. The identified statistically significant link between the studied phenomena signals
that there is indeed a need for further analyses on the topic to uncover more precisely how do different BD types affect firm innovation.

Notes
1. The definition of diversity by Harrison and Klein (2007) as “the distribution of differences among the members of a unit with respect to a common attribute” has been considered as one of the most academically rigorous treatments of this notion.

2. Commonly defined as the first introduction of an invention to the market (goods and services), new production and delivery methods (processes) or new ways of organising work (OECD, 2005).

References


Board diversity and firm innovation


### Table A1. Overview of the studies included in the meta-analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>N</th>
<th>Diversity types</th>
<th>Operationalisation of innovation (correlations, n)</th>
<th>Average board size reported</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>Miller and Triana (2009)</td>
<td>USA 432</td>
<td>Gender; Race</td>
<td>R&amp;D (1–2)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>2)</td>
<td>Wincent et al. (2010)</td>
<td>Sweden 53</td>
<td>Expertise</td>
<td>Composite variable based on questionnaire items (3)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>3)</td>
<td>Torchia et al. (2011)</td>
<td>Norway 317</td>
<td>Gender</td>
<td>Composite variable based on questionnaire items (4)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>4)</td>
<td>Galiani and Zenou (2012)</td>
<td>France 176</td>
<td>Experience; Gender</td>
<td>Questionnaire items (3–6)</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>5)</td>
<td>Martini et al. (2012)</td>
<td>Italy 69</td>
<td>Gender</td>
<td>R&amp;D (7)</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>6)</td>
<td>Zona et al. (2013)</td>
<td>Italy 301</td>
<td>Expertise</td>
<td>Composite variable based on questionnaire items (8)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>7)</td>
<td>Zona (2014)</td>
<td>Italy 301</td>
<td>Expertise</td>
<td>Composite variable based on questionnaire items (9)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>8)</td>
<td>Cook and Glass (2015)</td>
<td>USA 472</td>
<td>Gender; Race</td>
<td>Product innovations (10–11) Patents (12)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>9)</td>
<td>Kim and Kim (2015)</td>
<td>South Korea 108</td>
<td>Expertise</td>
<td></td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>10)</td>
<td>Midavaine et al. (2016)</td>
<td>USA 25</td>
<td>Experience; Expertise; Gender</td>
<td>R&amp;D (13–15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11)</td>
<td>Ariff et al. (2017)</td>
<td>Malaysia 220</td>
<td>Experience; Expertise; Gender; Race; Gender</td>
<td>R&amp;D (16)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>12)</td>
<td>Mukarram et al. (2018)</td>
<td>India 71</td>
<td>Gender</td>
<td>R&amp;D (17)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>13)</td>
<td>Hernández-Lara and Gonzales-Bustos (2019)</td>
<td>Spain 69</td>
<td>Gender; Nationality</td>
<td>Patents; Two R&amp;D measures (18–23)</td>
<td>Yes</td>
<td>−</td>
</tr>
<tr>
<td>14)</td>
<td>Li (2019)</td>
<td>USA 895</td>
<td>Gender</td>
<td>Patents; R&amp;D (24–25) Patents (26)</td>
<td>No effect</td>
<td></td>
</tr>
<tr>
<td>15)</td>
<td>Liao et al. (2019)</td>
<td>China 688</td>
<td>Gender</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>16)</td>
<td>Nadeem et al. (2019)</td>
<td>UK 424</td>
<td>Gender</td>
<td>R&amp;D (27)</td>
<td>Yes</td>
<td>+</td>
</tr>
</tbody>
</table>

(continued)
Appendix 2

Meta-analysis: Random effects model

For the purposes of the meta-analysis, the correlation measures need to be transformed into effect sizes (Fisher’s z), since the analysis itself is performed using these transformed values. The transformation from sample correlation \( r \) of study \( i \) to Fisher’s z transformed effect size, denoted here as \( Y_i \), is performed as (Borenstein et al., 2009):

\[
Y_i = 0.5 \times \ln \left( \frac{1 + r_i}{1 - r_i} \right)
\]

(Equation 1)

### Table A1.

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>N</th>
<th>Diversity types</th>
<th>Operationalisation of innovation (correlations, ( n ))</th>
<th>Average board size reported</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>17) Gonzales-Bustos et al. (2020)</td>
<td>Spain</td>
<td>86</td>
<td>Gender</td>
<td>R&amp;D (28)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>18) Hernández-Lara and Gonzales-Bustos (2020)</td>
<td>Spain</td>
<td>86</td>
<td>Gender</td>
<td>Patents; R&amp;D (29–30)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>19) Liu et al. (2020)</td>
<td>USA</td>
<td>683</td>
<td>Gender</td>
<td>R&amp;D (31)</td>
<td>Dependent on R&amp;D intensity</td>
<td></td>
</tr>
<tr>
<td>20) Manita et al. (2020)</td>
<td>France</td>
<td>120</td>
<td>Gender</td>
<td>R&amp;D (32)</td>
<td>Yes</td>
<td>–</td>
</tr>
<tr>
<td>21) Martinez-Jimenez et al. (2020)</td>
<td>Spain</td>
<td>100</td>
<td>Gender</td>
<td>Composite variable based on questionnaire items (33)</td>
<td>No effect</td>
<td></td>
</tr>
<tr>
<td>22) Rejeb et al. (2020)</td>
<td>Tunisia</td>
<td>81</td>
<td>Gender</td>
<td>Two composite variables based on questionnaire items (34–35)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>23) Vaivaran and Zhang (2020)</td>
<td>USA</td>
<td>276</td>
<td>Race; Gender</td>
<td>R&amp;D (36–37)</td>
<td>Yes</td>
<td>No effect</td>
</tr>
<tr>
<td>24) Cabeza-García et al. (2021)</td>
<td>Six European countries</td>
<td>231</td>
<td>Gender</td>
<td>R&amp;D (38)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>25) Griffin et al. (2021)</td>
<td>45 countries</td>
<td>12,244</td>
<td>Gender</td>
<td>Three measures based on patents and R&amp;D (39–41) R&amp;D (42)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>26) Jhunjhunwala et al. (2021)</td>
<td>India</td>
<td>947</td>
<td>Gender</td>
<td></td>
<td>Yes</td>
<td>–</td>
</tr>
<tr>
<td>27) Luo et al. (2021)</td>
<td>China</td>
<td>2,393</td>
<td>Nationality; Gender</td>
<td>Patents; R&amp;D (43–46)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>28) Saggese et al. (2021)</td>
<td>Italy</td>
<td>149</td>
<td>Gender; Experience</td>
<td>R&amp;D (47–48)</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>29) Zhong et al. (2021)</td>
<td>China</td>
<td>958</td>
<td>Gender</td>
<td>Patents; R&amp;D (49–50)</td>
<td>No effect</td>
<td></td>
</tr>
</tbody>
</table>
The overall mean weighted effect size $M$ is then calculated as (Hedges and Vevea, 1998; Borenstein et al., 2010):

$$M = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$  \hspace{1cm} (Equation 2)

where $k$ is the number of individual studies included and $W_i$ is the weight of an individual study. The weights are calculated under the inverse variance scheme (Borenstein et al., 2010):

$$W_i = \frac{1}{V_i + T^2}$$  \hspace{1cm} (Equation 3)

The weights are therefore dependent on between-study variance $T^2$ (for details of computing $T^2$, see Borenstein et al., 2010) common to all studies in the sample and on individual studies’ within-study variance $V_i$, which is calculated here ($n$ denotes the sample size) simply, as in the fixed effects model, as (Borenstein et al., 2009):

$$V_i = \frac{1}{n - 3}$$  \hspace{1cm} (Equation 4)

As such, studies with larger samples are assigned more weight (Borenstein et al., 2010). However, the weights assigned under the random effects model are quite balanced: large studies are less likely to dominate, and small studies are less likely to be trivialised in the analysis (Borenstein et al., 2009). That is, since the random effects model under the inverse variance scheme takes into account a between-study variance component (DerSimonian and Laird, 1986), the weights are not solely determined by sample size but are a more nuanced measure (Borenstein et al., 2009).

In calculating the 95% confidence intervals, the so-called “weighted variance method” proposed by Hartung (1999)—which has been shown to outperform the assumption that the effect size would follow a standard normal distribution in the presence of heterogeneity (Sánchez-Meca and Marín-Martínez, 2008), as is the case in this study—was applied assuming a two-tailed ($\alpha/2$) Student’s $t$ distribution with $k - 1$ degrees of freedom for the overall mean weighted effect size (Sánchez-Meca and Marín-Martínez, 2008; Thorlund et al., 2011):

$$CL_{LL} = M - t_{k-1,1-\alpha/2}SE_M$$

$$CL_{UL} = M + t_{k-1,1-\alpha/2}SE_M$$  \hspace{1cm} (Equation 5)

where $LL$ denotes the lower limit and $UL$ the upper limit of the confidence interval $CI$ and $SE_M$ denotes the standard error (Borenstein et al., 2009):

$$SE_M = \sqrt{V_M}$$  \hspace{1cm} (Equation 6)

where $V_M$ is the variance of the overall mean weighted effect size (Thorlund et al., 2011):

$$V_M = \frac{\sum_{i=1}^{k} W_i(Y_i - M)^2}{(k - 1)\sum_{i=1}^{k} W_i}$$  \hspace{1cm} (Equation 7)

For the random effects model, the $p$-value is calculated here by testing the null hypothesis that the mean true effect size is zero. The test statistics $Z$ and $p$ (for two-tailed $\Phi(|Z|)$ standard normal cumulative distribution), thus, are (Borenstein et al., 2010):

$$Z = \frac{M - 0}{SE_M}$$

$$p = 2[1 - (\Phi(|Z|))]$$  \hspace{1cm} (Equation 8)
Finally, the resulting effect sizes were converted back to correlations (for details of converting effect sizes back to correlations, see Borenstein et al., 2009) for presentation. The analysis was conducted using Meta-Essentials, which performs the calculations from correlations to Fisher’s z and back automatically (Suurmond et al., 2017).

Corresponding author
Teemu Makkonen can be contacted at: teemu.makkonen@uef.fi