CHAPTER 6

DOES GENDER MATTER? FEMALE REPRESENTATION ON CORPORATE BOARDS AND FIRM FINANCIAL PERFORMANCE: A META-ANALYSIS


A paper draft was presented at the Academy of Management Conference 2015.
Abstract
In recent years, there has been an ongoing, worldwide debate about the representation of females in companies. Our study aimed to meta-analytically investigate the controversial relationship between female representation on corporate boards and firm financial performance. Following a systematic literature search, data from 20 studies on 3097 companies published in peer-reviewed academic journals were included in the meta-analysis. On average, the boards consisted of eight members and female participation was low (mean 14%) in all studies. Half of the 20 studies were based on data from developing countries and 62% from higher income countries. According to the random-effects model, the overall mean weighted correlation between percentage of females on corporate boards and firm performance was small and non-significant ($r = .01$, 95% confidence interval: -.04, .07). Similar small effect sizes were observed when comparing studies based on developing vs. developed countries and higher vs. lower income countries. The mean board size was not related to the effect sizes in studies. These results indicate that the mere representation of females on corporate boards is not related to firm financial performance if other factors are not considered. We conclude our study with a discussion of its implications and limitations.

*Keywords*: gender diversity, organizational performance, board of directors, meta-analysis
Introduction

Advancing gender equality and female representation in corporate governance has increasingly become the focus of societal and political debates in various countries (Pande & Ford, 2011). Despite extensive efforts to increase women’s presence on corporate boards, men still dominate the corporate world. The financial effects of increased female representation on corporate boards may crucially determine if, and how, regulations to promote females to higher positions are implemented, because pursuing financial success is an innate characteristic of every company. While a number of scientific studies have investigated the relationship between gender diversity and firm financial performance, their conclusions are equivocal (Kochan et al., 2003; Webber & Donahue, 2001). These empirical discrepancies have led to a lack of conclusive evidence about the relationship between increased female representation and firm performance, creating uncertainty for policy makers, CEOs, and investors around the world. Owing to the conflicting evidence from primary studies, systematically summarizing the existing data on the topic in a quantitative meta-analysis has merit. While our general research question is similar to that of Post and Byron (Post & Byron, 2015), the methodological and analytical approach differs substantially between the two analyses. Our study aims to investigate the relationship of interest with a different, more rigorous and controlled methodological approach, and subsequently compares the results of the two meta-analyses. Investigating this relationship in a different sample and with different operationalizations of the variables (compared to Post & Byron, 2015) is especially important, because, in their analysis, the overall mean weighted correlation between female participation on boards and firm performance was very small (only marginally different from zero). Thus, this paper investigates the general relationship between female representation and firm performance using a new and different methodological approach, highlights our additional contribution to the literature, and compares the similarities and differences between the two analyses.

Literature Overview

A board of directors monitors the activities of an organization or company. It sets the corporate strategy, appoints and supervises senior management, and functions as the main corporate governance mechanism. The role of the board in determining the corporate strategy therefore influences firm performance. Since diversity is often considered a double-edged sword (e.g., Milliken & Martins, 1996), meaning that increased diversity can result in advantages and disadvantages regarding desired outcomes, a board composed of diverse directors affects firm performance either positively or negatively. Diversity’s positive and negative effects could also neutralize each other, or could depend on how it is managed.
Along these lines, a meta-analysis by Webber and Donahue (2001) examined the effects of diversity on work group performance in a sample of 45 effect sizes. Low job-related (e.g., age, gender) and highly job-related diversity (e.g., educational background) were measured, but both failed to show a significant relationship with work group performance.

Further, primary studies also do not show a clear consensus on whether gender diversity benefits or disadvantages firm performance (Jackson, Joshi, & Erhardt, 2003; Miller & del Carmen Triana, 2009). At first glance, the relationship between female representation on corporate boards and firm financial performance shows a similar pattern to that of the general diversity-performance relationship, being either positive (Mahadeo, Soobaroyen, & Hanuman, 2011), negative (Pathan & Faff, 2013), or non-significant (Strøm, D’Espallier, & Mersland, 2014). Thus, it remains unclear if increased female representation on corporate boards is associated with firm performance, and, if so, in which direction.

Advocates of greater female representation on corporate boards usually rely on two lines of arguments: the ethical or the business case for diversity (Robinson & Dechant, 1997). The former argues that women should be considered for leadership positions for equality reasons. The aim is therefore not directly to increase performance, but rather that greater female representation is considered a positive and just result in itself (Brammer, Millington, & Pavelin, 2007). Thus, a higher proportion of females on boards might not necessarily be related to better firm performance, but would reflect that boards with more females closely represent the ‘real world’, while other factors than gender alone contribute to better financial outcomes. The business case for diversity holds that if a board comprises heterogeneous directors, diversity leverages financial growth and success (Robinson & Dechant, 1997), indicating that a higher proportion of females could be related to better firm performance. A final outcome, not explained by either of the cases above, would be a negative relationship between a higher proportion of females on boards and lower firm performance. The aim of this article is to summarize the already existing quantitative evidence and attempt to explain the results in light of these cases.

**Positive Effects of Increased Female Representation on Firm Performance**

The *business case* for diversity holds that diverse team members improve corporate governance by introducing broader knowledge bases and experiences (Fondas & Sassalos, 2000; Robinson & Dechant, 1997). Accordingly, the *cognitive resource model* suggests that as (gender) diversity in groups increases, the available cognitive resources increase as well (Jackson, May, & Whitney, 1995; McLeod, Lobel, & Cox, 1996). If used effectively, these
diverse perspectives can contribute to a more thorough search for alternative solutions to problems because they introduce new perspectives to the boardroom (Watson, Kumar, & Michaelsen, 1993). These diverse perspectives also foster a critical analysis of complex problems, prevent premature decision-making (Carter, Souza, Simkins, & Simpson, 2010; Farrell & Hersch, 2005; Milliken & Martins, 1996; van Knippenberg et al., 2004), and develop creative and innovative solutions (Bassett-Jones, 2005). Hence, increased female representation on corporate boards should improve firm financial performance through the diverse perspectives introduced to the boardroom.

Another essential argument in support of the business case for diversity is that women introduce useful female leadership qualities and skills to the boardroom. These include, for example, risk averseness and less radical decision-making (Croson & Gneezy, 2009; Jianakoplos & Bernasek, 1998), as well as more sustainable investment strategies (Charness & Gneezy, 2012). In addition, female leaders fulfil their leadership roles in a more transformational way than their male counterparts, distinguishing themselves especially through their encouraging and supportive treatment of colleagues and subordinates (i.e., individualized consideration; Eagly, Johannesen-Schmidt, & van Engen, 2003). Females are also said to value their responsibilities as directors higher, which is associated with more effective corporate governance (Terjesen, Sealy, & Singh, 2009). Furthermore, diversity on corporate boards generally benefits organizations, by providing wider and better connections and ties to suppliers, organizations, and consumers, which decrease market uncertainties and dependencies (Miller & del Carmen Triana, 2009). In sum, an increased female presence on corporate boards is associated with the introduction of new desirable leadership skills and a variety of strategic advantages for companies. Following this reasoning, we expect a positive relationship between increased female representation and firm financial performance.

Negative Effects of Increased Female Representation on Firm Performance

Individuals are likely to perceive others and themselves in terms of salient social categories, such as gender, thereby creating in- and out-groups (Tajfel, 1978). These categorization tendencies, which might lead to heightened gender salience and a perceived lack of alignment with the group’s stereotypes (Abrams, Thomas, & Hogg, 2011), can compromise functional team processes when demographic subgroups emerge. If the emerging subgroups on corporate boards are based on gender, communication and cooperation might be impaired (van Knippenberg et al., 2004), leading to increased conflicts between board members. The probability of conflict might be further enhanced if the directors identify stronger with the opinions of fellow directors of the same gender (Richard, Barnett, Dwyer, & Chadwick, 2004),
or if the introduction of new perspectives, previously mentioned as one of diverse groups’ advantages, backfires (Jehn, Northcraft, & Neale, 1999). In turn, this potential for interpersonal conflicts might retard the decision-making process and lead to a lack of cohesion between board members and to decreased strategic consensus (Amason, 1996; Knight et al., 1999), hindering corporate boards’ effectiveness. In fast-paced environments, such as on corporate boards where strategic decisions need to be taken quickly, conflict-free communication is crucial to maintain effective performance (Williams & O’Reilly, 1998). And even if these issues can be overcome, the additional time and resources spent on solving them might decrease group and organizational performance (Kyereboah-Coleman, 2006).

These far-reaching potentials for impaired team processes might especially challenge females, who struggle to participate and maintain their standing in the already male-dominated boardroom (Tuggle, Sirmon, & Bierman, 2011) and are at risk of experiencing role ambiguity and role conflict, because they do not conform to typical gender roles in leadership (Koenig, Eagly, Mitchell, & Ristikari, 2011). Such females might be perceived as “tokens” to meet society’s expectations or those of important stakeholders, and could therefore be marginalized and not be taken seriously on the board (Kanter, 1977), which might subsequently hinder their and the entire board’s performance (Jackson & Schuler, 1985; Tubre & Collins, 2000).

All in all, increased female representation could potentially lead to decreased firm financial performance due to a number of strategic disadvantages, increased interpersonal conflicts, and their associated negative consequences. Following this reasoning, we expect a negative relationship between increased female representation and firm financial performance.

In conclusion, it is difficult to determine the relationship between female representation on boards and firm performance a priori. Summarizing all studies measuring the relationship between female representation on corporate boards and firm financial performance could provide substantive evidence to address the question whether increased female representation on corporate boards alone is positively or negatively related to firm financial performance. The meta-analysis by Post and Byron (2015) provided the first systematic summary of this relationship in studies selected from a range of electronic databases and unpublished sources. Using 140 studies (92 published, 48 unpublished), they show that the relationship is very small \( r = .047 \) between female representation and accounting returns, and \( r = .014 \) between female representation and market performance), and that it might depend on moderators, such as shareholder protection or gender parity in a given country. As opposed to Post and Byron (2015), our study follows a more rigorous and controlled methodological approach by investigating the relationship between percentage of females on corporate boards and firm
financial performance, operationalized as return on assets (ROA), return on equity (ROE), and Tobin’s \( Q(\)\), by means of a meta-analysis of articles published in peer-reviewed academic journals. These narrow operationalizations of the variables could increase the certainty with which theoretical and practical implications can be deduced. In light of the study by Post and Byron (2015), we also aim to compare the results of the two meta-analyses. A systematic investigation of the published literature could also reveal potential moderators of such a relationship based on systematic differences between various studies in terms of their source of data and company characteristics, which could only be determined post hoc once we had located and coded all the available data. Based on the studies reviewed above and the meta-analysis of Post and Byron (2015), we hypothesized that female representation on corporate boards is either positively or negatively related to firm financial performance, but that the magnitude of such a relationship is likely to be small.

**Method**

**Systematic Search Strategy**

A systematic literature search was conducted in EBSCO on March 7, 2014, using the search strategy described in Table 1. The databases and the search terms we utilized differ from those used by Post and Byron (2015). Therefore, the two meta-analyses are based on a different sample of articles. In addition, we conducted a hand-search of the reference sections of review articles in this field and of Google Scholar, using the terms listed in Table 1 between March 7 and May 12, 2014. Unlike Post and Byron (2015), we only searched for (and included) studies published in peer-reviewed academic journals.

<table>
<thead>
<tr>
<th>Number of sources ((k))</th>
<th>Search terms and limits</th>
<th>Databases (1986-March 2014)</th>
</tr>
</thead>
</table>
| 325                      | [Subject OR Title (“gender diversity” OR gender OR female OR wom*n OR "board diversity" OR "board of director*" OR "board structure")]
|                          | AND                     | PsycInfo                    |
|                          | [Subject OR Title (“organi*ation* performance” OR “firm performance” OR "financial performance" OR “company performance”)] limits: Academic Articles, English Language | EconLit |
|                          |                         | Business Source Premier     |
|                          |                         | Academic Search Premier     |
Figure 1
*Study Assessment and Exclusion Criteria*

**IDENTIFICATION**

- \( k = 325 \) records published in 1986-2014 from database searching (Table 1)

**SCREENING**

- \( k = 293 \) records after duplicates removed

**ELIGIBILITY**

- \( k = 293 \) records (titles/abstracts) screened by RN

- \( k = 52 \) full-text articles assessed by RN/JP

- \( k = 32/52 \) (62%) articles excluded:
  - \( k = 9 \): Inappropriate performance measures
  - \( k = 9 \): Inappropriate gender diversity measure (Blau’s index, dichotomous coding of gender)
  - \( k = 4 \): Gender diversity was not measured
  - \( k = 3 \): Data for owners or entrepreneurs
  - \( k = 2 \): Does not measure the relationship between gender diversity and performance
  - \( k = 2 \): Not found in the database
  - \( k = 2 \): Does not report sample size
  - \( k = 1 \): Qualitative design

- \( k = 241 \) excluded

**INCLUDED**

- \( k = 20 \) studies included in the quantitative meta-analysis

*Note. \( k \) = number of studies.*

**Study Selection**

The study selection process is summarized on the PRISMA flowchart (Moher, Liberati, Tetzlaff, & Altman, 2009) in Figure 1. Of the 325 sources identified during the electronic search (Table 1) and the 17 sources identified through the hand search, 52 studies met the inclusion criteria and were fully examined (Figure 1). The studies selected for the final meta-analysis had to have a quantitative design and report the Pearson product moment correlation coefficient, \( r \),
between the percentage female representation on boards of directors and firm performance (measured as ROA, ROE, or Q), with the number of observations (number of firms × total length of data collection in years) used as the sample size. Studies not reporting the correlation coefficient, but including the variables required in our analysis, were also included in our sample if the authors (contacted via email) provided these correlations. Studies were excluded if other performance measures were used (e.g., return on investment, firm value). We focused on ROA, ROE, and Q, because they are relatively objective and the most commonly used indicators of organizational performance. Return on investments and firm value were not used because there might be large differences in these measures due to economic differences between countries or due to strategic orientations. Unlike in Post and Byron (2015), studies were excluded from our meta-analysis if female representation was measured as a dichotomous variable (presence vs. absence of females on board), or if they used measures that did not explicitly target the representation of females (e.g., the Blau Index or the Shannon Index), because they might bias the analysis. We also excluded studies that assessed female representation in an inappropriate body (such as the management or owners). This approach allowed us to achieve a reliable measure of female participation on corporate boards.

**Definition of Variables**

**Firm performance measure.** Firm performance was measured with three variables. First, ROA is computed by dividing the “earnings before extraordinary income and preferred dividend in financial year” by the “average of book values of total assets at the beginning and at the end of [a] financial year” (Haslam, Ryan, Kulich, Trojanowski, & Atkins, 2010). The ROA measures the company’s production of “accounting based revenues in excess of actual expenses from a given portfolio of assets measured as amortized historical costs” (Carter et al., 2010). Second, ROE is computed by dividing the company’s “earnings before extraordinary income and preferred dividend in [a] financial year” by the “average of book values of common equity at the beginning and at the end of a financial year” (Haslam et al., 2010). Thus, this measure assesses the returns to the company’s shareholders. ROE and ROA share a characteristic in that they are both “backward looking” accounting-based measures, meaning that they are based on the company’s self-reported financial performance in the recent past (Haslam et al., 2010). Third, Tobin’s Q is a market-based firm performance index calculated by dividing the firm’s “year-end market capitalization and average of book values of total debt at the beginning and at the end of [a] financial year” by the “average of book values of total assets at the beginning and at the end of [a] financial year” (Haslam et al., 2010). This measure is increasingly used in diversity research and scholars argue that it is more reliable than
accounting-based measures (Campbell & Minguez-Vera, 2008), because it represents a “forward looking” measure, meaning that it reflects the future potential of a firm’s performance (Haslam et al., 2010). Beyond this, it is a standardized measure with intuitive interpretation criteria: If the ratio is greater than one, the firm has a higher ability to create value by effectively allocating its resources, indicating a high competitive advantage for that firm (Rose, 2007). In contrast to ROE and ROA, this measure lacks an accounting convention bias, because it is considered objective by not relying on self-reported data (Campbell & Minguez-Vera, 2008).

**Female representation.** Female representation was measured as the percentage of females on corporate boards.

**Moderator variables.** In addition to performance measures and female representation on boards, the following variables were included in the current analysis due to the relevant data in all, or most, of the studies:

1. *Country development and income.* Using the countries in which data were collected, we classified studies into dichotomous groups based on their economic development status (developed: \( k = 9 \); developing: \( k = 10 \)) and their Gross National Income per capita (GNI; high income: \( k = 13 \); low income: \( k = 6 \); United Nations, 2012).

2. *Mean board size.* This variable was a scale measure of the mean number of directors on boards.

**Data Extraction**

Data were extracted independently by three authors of this study from \( k = 20 \) studies (Post and Byron, 2015, included 13 of the 20 studies from our sample in their analysis), resulting in 34 coded effect sizes, and any inconsistencies were resolved during discussions. The authors of 16 studies were contacted via e-mail by the third author and seven provided additional data (either the sample size used for correlations or the correlation coefficients if they were not reported). Table 2 lists the study characteristics and effect size data.
## Table 2

**Study Characteristics and Effect Size Data in 20 Studies included in the Meta-Analysis**

<table>
<thead>
<tr>
<th>Study</th>
<th>Period</th>
<th>CO</th>
<th>DEV/DC</th>
<th>GNI</th>
<th>PM</th>
<th>%F</th>
<th>r</th>
<th>N</th>
<th>No. Firms</th>
<th>Board</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahern &amp; Dittmar (2012)</td>
<td>2001 - 2009</td>
<td>NO</td>
<td>DEV</td>
<td>HI</td>
<td>Q</td>
<td>24.6</td>
<td>-.067</td>
<td>1074</td>
<td>248</td>
<td>5.5</td>
</tr>
<tr>
<td>Behren &amp; Staubo (2013)</td>
<td>2000 - 2009</td>
<td>NO</td>
<td>DEV</td>
<td>HI</td>
<td>ROA</td>
<td>17.0</td>
<td>.029</td>
<td>1560</td>
<td>-274</td>
<td>5.6</td>
</tr>
<tr>
<td>Darmadi (2011)</td>
<td>2007</td>
<td>IN</td>
<td>DC</td>
<td>LI</td>
<td>ROA</td>
<td>12.0</td>
<td>-.070</td>
<td>169</td>
<td>169</td>
<td>8.4</td>
</tr>
<tr>
<td>Darmadi (2011)</td>
<td>2007</td>
<td>IN</td>
<td>DC</td>
<td>LI</td>
<td>Q</td>
<td>12.0</td>
<td>-.160</td>
<td>169</td>
<td>169</td>
<td>8.4</td>
</tr>
<tr>
<td>Galbreath (2011)</td>
<td>2005 - 2007</td>
<td>AU</td>
<td>DEV</td>
<td>HI</td>
<td>ROA</td>
<td>9.0</td>
<td>-.060</td>
<td>151</td>
<td>151</td>
<td>7.6</td>
</tr>
<tr>
<td>Galbreath (2011)</td>
<td>2005 - 2007</td>
<td>AU</td>
<td>DEV</td>
<td>HI</td>
<td>ROE</td>
<td>9.0</td>
<td>-.220</td>
<td>151</td>
<td>151</td>
<td>7.6</td>
</tr>
<tr>
<td>Garba &amp; Abubakar (2014)</td>
<td>2004 - 2009</td>
<td>NI</td>
<td>DC</td>
<td>LI</td>
<td>ROA</td>
<td>9.5</td>
<td>.139</td>
<td>72</td>
<td>12</td>
<td>9.1</td>
</tr>
<tr>
<td>Garba &amp; Abubakar (2014)</td>
<td>2004 - 2009</td>
<td>NI</td>
<td>DC</td>
<td>LI</td>
<td>Q</td>
<td>9.5</td>
<td>.046</td>
<td>72</td>
<td>12</td>
<td>9.1</td>
</tr>
<tr>
<td>Haslam et al. (2010)</td>
<td>2001 - 2005</td>
<td>UK</td>
<td>DEV</td>
<td>HI</td>
<td>ROA</td>
<td>8.4</td>
<td>.040</td>
<td>486</td>
<td>97</td>
<td>11.3</td>
</tr>
<tr>
<td>Haslam et al. (2010)</td>
<td>2001 - 2005</td>
<td>UK</td>
<td>DEV</td>
<td>HI</td>
<td>ROE</td>
<td>8.4</td>
<td>.030</td>
<td>486</td>
<td>97</td>
<td>11.3</td>
</tr>
<tr>
<td>Haslam et al. (2010)</td>
<td>2001 - 2005</td>
<td>UK</td>
<td>DEV</td>
<td>HI</td>
<td>Q</td>
<td>8.4</td>
<td>-.110</td>
<td>486</td>
<td>97</td>
<td>11.3</td>
</tr>
<tr>
<td>Julizaerna &amp; Sori (2012)</td>
<td>2008 - 2009</td>
<td>ML</td>
<td>DC</td>
<td>HI</td>
<td>ROA</td>
<td>10.6</td>
<td>-.015</td>
<td>280</td>
<td>280</td>
<td>7.6</td>
</tr>
<tr>
<td>Lümkerath-Rovers (2011)</td>
<td>2005 - 2007</td>
<td>NL</td>
<td>DEV</td>
<td>HI</td>
<td>ROE</td>
<td>4.0</td>
<td>.328</td>
<td>297</td>
<td>99</td>
<td>7.8</td>
</tr>
<tr>
<td>Mahadeo et al. (2011)</td>
<td>2007</td>
<td>MA</td>
<td>DC</td>
<td>HI</td>
<td>ROA</td>
<td>3.1</td>
<td>.337</td>
<td>42</td>
<td>42</td>
<td>9.6</td>
</tr>
<tr>
<td>Pathan &amp; Faff (2013)</td>
<td>1997 - 2011</td>
<td>US</td>
<td>DEV</td>
<td>HI</td>
<td>ROE</td>
<td>7.9</td>
<td>-.100</td>
<td>2640</td>
<td>212</td>
<td>12.7</td>
</tr>
<tr>
<td>Rodriguez-Dominguez et al. (2010)</td>
<td>2004 - 2006</td>
<td>ES</td>
<td>DEV</td>
<td>HI</td>
<td>Q</td>
<td>---</td>
<td>-.061</td>
<td>288</td>
<td>96</td>
<td>---</td>
</tr>
<tr>
<td>Shafique et al. (2014)</td>
<td>2008 - 2012</td>
<td>PK</td>
<td>DC</td>
<td>LI</td>
<td>ROA</td>
<td>33.3</td>
<td>-.049</td>
<td>30</td>
<td>6</td>
<td>---</td>
</tr>
<tr>
<td>Shukeri et al. (2012)</td>
<td>2011</td>
<td>ML</td>
<td>DC</td>
<td>LI</td>
<td>ROE</td>
<td>9.8</td>
<td>.094</td>
<td>300</td>
<td>300</td>
<td>7.4</td>
</tr>
<tr>
<td>Strom et al. (2014)</td>
<td>1998 - 2008</td>
<td>WW</td>
<td>---</td>
<td>---</td>
<td>ROA</td>
<td>---</td>
<td>-.004</td>
<td>497</td>
<td>329</td>
<td>---</td>
</tr>
<tr>
<td>Strom et al. (2014)</td>
<td>1998 - 2008</td>
<td>WW</td>
<td>---</td>
<td>---</td>
<td>ROE</td>
<td>---</td>
<td>-.003</td>
<td>462</td>
<td>329</td>
<td>---</td>
</tr>
<tr>
<td>Wellalage &amp; Locke (2013)</td>
<td>2006 - 2010</td>
<td>SL</td>
<td>DC</td>
<td>LI</td>
<td>ROA</td>
<td>7.4</td>
<td>-.011</td>
<td>440</td>
<td>88</td>
<td>7.3</td>
</tr>
<tr>
<td>Wellalage &amp; Locke (2013)</td>
<td>2006 - 2010</td>
<td>SL</td>
<td>DC</td>
<td>LI</td>
<td>Q</td>
<td>7.4</td>
<td>-.430</td>
<td>440</td>
<td>88</td>
<td>7.3</td>
</tr>
<tr>
<td>Zhou et al. (2012)</td>
<td>2000 - 2009</td>
<td>CH</td>
<td>DC</td>
<td>HI</td>
<td>ROA</td>
<td>12.3</td>
<td>-.012</td>
<td>3197</td>
<td>-320</td>
<td>2.2</td>
</tr>
<tr>
<td>Zhou et al. (2012)</td>
<td>2000 - 2009</td>
<td>CH</td>
<td>DC</td>
<td>HI</td>
<td>Q</td>
<td>12.3</td>
<td>.087</td>
<td>3197</td>
<td>-320</td>
<td>2.2</td>
</tr>
</tbody>
</table>

**Note.** Abbreviations: *AU* = Australia; *Board* = mean size of the board; *CH* = China; *CO* = country of data collection; *DC* = developing country; *DEV* = developed country; *ES* = Spain; *GH* = Ghana; *GNI* = Gross National Income Classification; *HI* = high-income; *IN* = Indonesia; *KN* = Kenya; *LI* = low-income; *N* = number of observations (number of firms × total length of data collection in years); *MA* = Mauritius; *ML* = Malaysia; *NI* = Nigeria; *NL* = Netherlands; *NO* = Norway; *No. Firms* = Number of firms in sample; *Period* = time frame in which data were collected; *PK* = Pakistan; *PM* = Performance measure; *SL* = Sri Lanka; *US* = United States; *WW* = Worldwide; *%F* = percentage of female board members.
Data Analysis (Meta-Analysis)

The effect size used in the current meta-analysis was the Pearson product moment correlation coefficient, \( r \). The interpretation criteria for the absolute magnitude of \( r \) are: .1 small, .3 medium, and .5 large (Cohen, 1988). The meta-analysis was computed using Comprehensive Meta-Analysis 2.0 (CMA; Biostat, USA) according to the random-effects model with inverse-variance weights (Borenstein et al., 2009). The analysis was conducted in the following steps (Borenstein et al., 2009):

1. The effect size data (\( r \) and \( N \)) were reported for each study. Since the variance of \( r \) is biased based on the magnitude of \( r \), CMA converts \( r \) to Fisher’s \( z \), computes all subsequent analyses on Fisher’s \( z \), and converts the final results back to \( r \).

2. Each effect size (Fisher’s \( z \)) was weighted according to the inverse-variance method (inverse of the sum of within- and between-study variance; (DerSimonian & Laird, 1986).

3. The overall mean weighted effect size of all studies was computed according to the random-effects model, where overall \( r \) is the sum of the product of all \( r \) (expressed as Fisher’s \( z \)) and weights divided by the sum of all weights. This model was used, because we assumed that the effect sizes would differ between studies in the analysis due to differences in study characteristics and because we only identified a random sample of all studies on this topic in our literature search.

The heterogeneity between study effect sizes was computed using a \( Q \) statistic and an \( I^2 \) index, where \( I^2 = 100\% \times (Q-df)/Q \) with \( df = k-1 \) and \( k = \) number of studies (Borenstein et al., 2009). The \( I^2 \) index quantifies the variability in the effect sizes due to real (rather than chance) differences between studies. This variability can be interpreted as low (25%), moderate (50%), or high (75%) heterogeneity due to real differences between studies (Higgins, Thompson, Deeks, & Altman, 2003).

Publication Bias Analyses

Studies with statistically significant and high effect sizes are more likely to be published in academic journals (Borenstein et al., 2009). Such a publication bias could have inflated the result of our meta-analysis, which focused specifically on findings published in academic journals. Thus, we controlled for publication bias using methods available in CMA. Specifically, publication bias could be present if:

1. Rosenthal’s Fail Safe-\( N \) is low, meaning that it takes only a few theoretically missing studies with low effect sizes to nullify the result of a meta-analysis (Rosenthal, 1979),
2. A funnel plot of standard error by Fisher’s z in each study is not symmetrical (Egger et al., 1997) and mathematically correcting for symmetry (using the trim-and-fill analysis) changes the interpretation of the overall analysis (Duval & Tweedie, 2000).

3. The study effect sizes and precision differ systematically according to the statistically significant Begg and Mazumdar correlation (Begg & Mazumdar, 1994) and Egger’s regression (Egger et al., 1997).

**Sensitivity and Moderator Analyses**

The stability of the overall mean weighted \( r \) over time was investigated by adding one study at a time to all previous studies (cumulative analysis). The influence of individual studies on the overall mean weighted \( r \) was investigated by removing one study at a time from the overall analysis (one-study removed analysis). The moderator analyses (subgroup analyses and univariate meta-regressions) were used to test the influence of systematic differences between studies on the overall mean weighted \( r \).

**Results**

**Study Characteristics**

A total of 20 studies with 34 effect sizes were included in the current meta-analysis. Firm performance was measured according to ROA in 85% of effect sizes \((k = 17)\), ROE in 45% \((k = 9)\), and Q in 40% \((k = 8\); Table 3). All 34 effect sizes were based on an average of 734 observations from 146 firms collected in slightly more than five years (Table 3). On average, the boards in those firms consisted of almost eight members with a low female representation (14%; Table 3). Half of all effect sizes were based on data from developing countries and data in 62% of the studies came from high income countries (Table 3).

**Relationship between Female Representation and Firm Performance**

There was a small positive, but not statistically significant, relationship between the percentage female representation on corporate boards and the combined mean of the three firm performance measures; overall mean weighted \( r = .01 \), 95% confidence interval, \( CI [-.04, .07] \), \( p = .598 \), \( k = 20 \) (Figure 2). High variability in effect sizes among the 20 studies existed \((I^2 = 87\%, Q = 142.84, df = 19, p < .001)\).
Table 3
Descriptive Statistics for \( k=20 \) Studies included in the Current Meta-Analysis

<table>
<thead>
<tr>
<th></th>
<th>( k ) studies (%) of 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>WESP</td>
<td></td>
</tr>
<tr>
<td># developing countries (% within outcome)</td>
<td>9 (45%)</td>
</tr>
<tr>
<td># developed countries (% within outcome)</td>
<td>10 (50%)</td>
</tr>
<tr>
<td># high-income countries (% within outcome)</td>
<td>13 (62%)</td>
</tr>
<tr>
<td>GNI</td>
<td></td>
</tr>
<tr>
<td># low-income countries (% within outcome)</td>
<td>6 (32%)</td>
</tr>
<tr>
<td>Mean number of observations (SD)</td>
<td>734 (875)</td>
</tr>
<tr>
<td>Mean number of firms (SD)</td>
<td>146 (104)</td>
</tr>
<tr>
<td>Mean board size (SD)</td>
<td>7.89 (2.49)</td>
</tr>
<tr>
<td>Mean % female (SD)</td>
<td>13.82 (9.52)</td>
</tr>
<tr>
<td>Mean data collection period (SD)</td>
<td>5.32 (3.99)</td>
</tr>
</tbody>
</table>

Note. Abbreviations: GNI = Gross National Income per capita; Q = Tobin’s Q; ROA = Return on Assets; ROE = Return on Equity; SD = standard deviation; WESP = World Economic Situations and Prospects classification.

Figure 2
Forest Plot of the Association between Percentage Female Representation on Corporate Boards and Firm Performance

Note. ‘Combined’ = mean effect size in studies using multiple measures of firm performance; ‘Total’ = total number of observations per study (number of firms \( \times \) number of years); The diamond depicts the overall mean weighted effect size \( r \) of all \( k=20 \) studies; overall mean weighted \( r = .01 \), 95% confidence interval, 95%CI: -.04, .07).
Publication Bias Analyses

Since the overall result of our analysis was not statistically significant, Rosenthal’s Fail-Safe $N$ was not applicable. Nevertheless, there was little evidence of publication bias in the current analysis, because the funnel plot (Figure 3) was mathematically symmetrical around the central vertical line (corresponding to the overall mean weighted effect size) according to the trim-and-fill analysis. Thus, the overall mean weighted effect size in the analysis (unfilled diamond) and the overall effect corrected for potential missing studies (filled diamond) are aligned and no filled circles (theoretically missing studies) are shown on the plot (Figure 3). The symmetry is particularly evident in the area towards the top of Figure 3, showing the studies (unfilled circles) with the lowest estimated variability in effect sizes, and, thus, the highest precision. These studies contributed the most weights to the calculation of the overall mean weighted effect size in the current meta-analysis. Therefore, confirming the results on the forest plot (Figure 2), the studies with the most weights had either positive, negative, or close to null effect sizes, indicating that there was no preference for high positive, or high negative, effects in the current meta-analysis. Although the studies towards the bottom of Figure 3 appear less symmetrically distributed, these had high estimated variability of effect sizes, low weights, and, thus, little influence on the overall mean weighted effect size in the current analysis. Finally, study effect sizes and precision did not differ systematically according to the non-significant Begg and Mazumdar correlation coefficient ($p = .770$) and Egger’s regression intercept ($p = .374$).

Sensitivity Analyses

The cumulative analysis showed that the overall mean weighted $r$ remained consistently small and non-significant as studies were added one at a time (based on publication date) from 2006 to 2014 (Figure 4). Similarly, the overall mean weighted $r$ was not influenced by any one study, because it remained small and non-significant as one study at a time was removed from the overall analysis.
Figure 3
*Funnel Plot of the Estimated Variability (Standard Error of the Mean, SEM) and Effect Size r (expressed as Fisher’s z) in each Study*

Note. This plot shows that the effect sizes in the individual studies (circles) were symmetrically distributed around the overall mean weighted effect size shown on the vertical line.

Figure 4
*Forest Plot of the Cumulative Analysis*

Note. ‘Combined’ refers to the mean effect size of studies that have used multiple measures of firm performance. ‘Total’ refers to the total number of observations (number of firms × number of years) as one study is added to all previous studies in each row. The plot shows how the overall mean weighted effect size r (referred to as ‘Point’) changes as each study is added over time to all previous studies. The diamond depicts the overall mean weighted effect size r of all $k = 20$ studies.
Subgroup Analyses and Meta-Regressions

According to a subgroup analysis, the overall outcome of the current meta-analysis was not dependent on the performance measure. Specifically, the overall mean weighted $r$ remained consistently small and non-significant when studies were grouped according to individual performance indicators (Table 4).

### Table 4

**Results of the Moderator Analyses and the Meta-Regression**

<table>
<thead>
<tr>
<th>Subgroups</th>
<th>$k$ studies (% of 20)</th>
<th>Mean weighted $r$ (95%CI)</th>
<th>$p$ two-tailed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance Measure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROA</td>
<td>17 (85%)</td>
<td>.00 (-.05, .05)</td>
<td>.861</td>
</tr>
<tr>
<td>ROE</td>
<td>9 (45%)</td>
<td>.08 (-.02, .19)</td>
<td>.125</td>
</tr>
<tr>
<td>Q</td>
<td>8 (40%)</td>
<td>-.10 (-.21, .02)</td>
<td>.107</td>
</tr>
<tr>
<td><strong>WESP Development</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Developed countries</td>
<td>9 (47%)</td>
<td>-.02 (-.06, .10)</td>
<td>.661</td>
</tr>
<tr>
<td>Developing countries</td>
<td>10 (53%)</td>
<td>.01 (-.07, .10)</td>
<td>.753</td>
</tr>
<tr>
<td><strong>GNI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower income countries</td>
<td>6 (32%)</td>
<td>-.02 (-.17, .12)</td>
<td>.750</td>
</tr>
<tr>
<td>Higher income countries</td>
<td>13 (68%)</td>
<td>.03 (-.03, .10)</td>
<td>.310</td>
</tr>
<tr>
<td><strong>Meta-regression predictor</strong></td>
<td>$k$ studies</td>
<td>Slope</td>
<td>Slope $p$ two-tailed</td>
</tr>
<tr>
<td>Mean Board Size</td>
<td>16</td>
<td>-.01</td>
<td>.400</td>
</tr>
</tbody>
</table>

Similarly, the overall mean weighted $r$ remained consistently small and non-significant when studies were grouped according to either country development (developed vs. developing) or country income (lower vs. higher income); Table 4. Finally, as shown in Table 4, the univariate meta-regression showed that the weighted effect sizes in individual studies (outcome) were not dependent on the mean board size (predictor).

**Discussion**

The main finding of the current study, based on data from 20 studies (34 effect sizes) published only in peer-reviewed academic journals, is that the relationship between the percentage of female directors on corporate boards and firm financial performance is consistently small and non-significant. The general magnitude of this result is in line with findings of Post and Byron (2015), who, based on data from 140 published and unpublished studies, also found a small correlation between gender diversity on corporate boards and firm financial performance. This is especially noteworthy, because both meta-analyses are based on different study samples and different operationalizations of the main measures (female representation and firm performance), providing further evidence to conclude that female representation and firm performance are not strongly associated. Interestingly, the two meta-analyses differ in that Post and Byron (2015) find a statistically significant correlation between
increased gender diversity on corporate boards and higher accounting returns. But concluding
that a higher female representation on corporate boards has practical implications for the
generation of profits from assets and investments seems debatable due to the overall small effect
size. By testing the relationship of interest with our more rigorous and controlled
methodological approach (for example, by including only peer-reviewed and published
studies), we provide further evidence that female representation on corporate boards is not
associated, positively or negatively, with firm performance. Although both meta-analyses
indicate only a very small correlation, primary research should further investigate the
relationship between boards’ gender composition and firm performance. This is because, as
argued below, the relationship might be too complex to be investigated on a univariate level.

In our international sample, female representation on corporate boards was not
significantly related to firm financial performance, as measured by the “backward-looking”
measures ROA and ROE and the forward-looking measure Tobin’s Q, if we did not control for
any other factors. This result is in line with other primary studies and meta-analyses (Ali, Ng,
& Kulik, 2013; Webber & Donahue, 2001; Wellalage & Locke, 2013), indicating only a small
association between (gender) diversity and firm financial performance, while contrasting
individual studies that find either a positive (Mahadeo et al., 2011) or negative relationship
(Pathan & Faff, 2013) between gender diversity and firm performance.

Results of both meta-analyses provide little evidence to support the business case for
gender diversity. However, more importantly, a higher representation of females on corporate
boards is also not associated with a detrimental effect on firm financial performance, which
supports the ethical case for diversity. If increased female representation on corporate boards is
not positively or negatively associated with firm performance, it seems reasonable to promote
gender equality in board representation. Given the current underrepresentation of females on
corporate boards in all studies included in our sample, and possibly in all countries worldwide
(the largest average percentage of female directors included in our sample was 37%
(Kyereboah-Coleman, 2006) and the overall average was only 14%), women should be
prioritized for promotions if they are equally qualified. By bringing the performance-based and
the ethical view together, fostering gender diversity in boardrooms seems justified and
desirable.

We do, however, acknowledge that our univariate approach to this intricate research
question is rather simplistic and might not do justice to the vast econometric complexity present
when studying the relationship between gender diversity and firm performance. Numerous
other variables not investigated in this study might influence the relationship between female
representation on corporate boards and firm financial performance. Future meta-analyses should investigate the relationship between various other diversity variables, such as age, tenure or education, on corporate boards and firm financial performance and subsume them in one analysis. Such an inclusive approach would yield benefits for practitioners and scientists.

The board’s limited influence on the firm financial performance might be a reason for the small overall effect size in the current meta-analysis. According to Bertrand and Schoar (2003), the CEO and CFO only explain about 5 to 6% of variance in corporate performance measures. Hence, the boards’ effect on actual firm performance may be limited in general. Various other factors which the board cannot alter or influence, such as the current economic and political situation, influence companies’ performance. Although such factors might be more important for firm performance than gender alone, they are difficult to quantify numerically for meta-analytic purposes.

When the Business Case Might Still Matter

The business case for diversity should not be abolished altogether. Whether an increased representation of females and the concomitant increase in gender diversity on corporate boards lead to performance benefits for the firm might depend on contextual factors, or on how diversity is managed. In our analysis, we aimed to find moderators of this relationship based on systematic differences between studies. Not surprisingly, the process of finding specific moderators was difficult, because there was high heterogeneity among studies in terms of reported study characteristics which could be used as potential moderators.

The relationship between female representation and firm performance remained independent of how firm financial performance was measured. This supports our initial assumption that these outcome measures are relatively objective and measure firm financial performance similarly. In addition, this finding increases the certainty with which our results can be interpreted, because the non-significant relationship seems to be independent of the outcome measure. On a descriptive level, the results of this subgroup analysis are also in line with Post and Byron (2015), who find that accounting returns, such as ROA and ROE, increase and market performance (Tobin’s Q) remains unrelated to whether there are more female directors on corporate boards. While the significant positive relationship between accounting returns and female representation on corporate boards, which Post and Byron (2015) find, deviates from our nonsignificant finding regarding these performance measures, their large sample size, which increased the power of their analysis and the likelihood of finding a significant result, is likely the reason for this deviation. Regardless of statistical significance,
the overall magnitude of their effect size for ROA and ROE was similar to our effect sizes, and
deducing practical implications from such small effect sizes might be debatable and misleading.

The characteristics of the country in which data were collected in the individual studies
had little influence on the effect sizes in the current analysis. Neither a country’s development
status, nor its GNI per capita influenced the relationship between female representation and
firm financial performance. Thus, whether a country is considered developed/developing or
“poor”/“rich” does not influence the effect that female board members have on firm
performance. While previous studies have shown that a country’s characteristics, such as a long
history of female participation in politics (Terjesen & Singh, 2008), influence the presence of
women on corporate boards, our results suggest that a country’s characteristics related to
economic performance do not influence the relationship between female representation and
firm financial performance. This indicates that aiming for equality should guide decisions
regarding future promotion to corporate boards, irrespectively of a country’s economic status.

In addition to our more rigorous meta-analytical approach to the research question, we
also tested what Post and Byron (2015) called for: the moderating effect of board size on the
relationship between female representation and firm performance. Larger boards might make it
more difficult for directors to influence decisions and might limit the influence of directors on
important decisions overall (Carpenter & Westphal, 2001; Westphal & Bednar, 2005), and the
percentage of female directors on larger boards would therefore also matter less. However, the
number of directors on corporate boards did not significantly influence the effect size
distribution in our analysis, suggesting that the non-significant relationship between female
representation and firm performance remains similar, regardless of how many directors are on
corporate boards, at least on the meta-analysis level. This is somewhat surprising because larger
boards usually experience a higher complexity in all decision-making processes (Sanders &
Carpenter, 1998) and have been shown to be associated with decreased financial performance
(ROE; Conyon & Peck, 1998). The increased complexity on larger boards might make it even
more difficult for females to have an impact, given their apparently wide underrepresentation.
In conclusion, although a higher representation of females on boards does not appear to be
directly associated with financial performance, more females on corporate boards might
indirectly influence firms’ financial performance. For example, females might provide a
protective effect against larger boards’ apparently increased interpersonal conflicts. However,
this requires further research.
Limitations and Future Research

There were a number of limitations in the current study. First, the main limitation is the small sample size of only 20 studies in our analysis, making it difficult to accept the null hypothesis due to low statistical power. However, overall, Post and Byron (2015) also report a similar result based on the higher number of studies in their analysis. Thus, it is unlikely that the apparent lack of a meaningful relationship between female representation on corporate boards and firm performance just resulted from low statistical power in both meta-analyses. Interestingly, regardless of the studies’ high heterogeneity, the variability of the overall mean weighted effect sizes is reasonably consistent in our analysis. Specifically, as shown on Figure 2, the effect sizes in all 20 studies can be classified as either positively or negatively small (most weighted correlation coefficients between ± .3). Thus, although it cannot be ruled out, it is unlikely that statistical power alone, or one or more hidden factors, consistently suppressed a relationship between female representation on boards and firm financial performance based on data from different countries and various industries. Future primary studies should attempt to focus on other diversity factors, such as educational level or the seniority of female board members of successful companies to determine how these characteristics affect firm performance.

Second, the quality of the included studies was not assessed by means of standardized scales. Instead, we only selected studies published in academic journals, assuming that such studies are of a higher quality than unpublished sources, because experts had reviewed them. Furthermore, we also assumed that academic research on the topic might be more value-free than those in unpublished sources. Our assumption might not have been entirely correct, because the quality of the reported statistical data was generally poor in many of the examined studies. For example, although most studies conducted complex multivariate statistical analyses, it was often unclear if the reported regression coefficients were unstandardized or standardized (meaning that high-quality multivariate data corrected for various factors could not be used in our analysis). Therefore, the authors of future primary studies should use standardized guidelines to report quantitative results, which can be subsequently utilized in meta-analyses.

Third, including only published sources could potentially lead to a publication bias (inflation of effect sizes) in a meta-analysis. However, based on the outcome of the various tests we conducted, there was little evidence of publication bias in our analysis. Interestingly, in the current analysis, the strongest evidence against publication bias is a simple visual inspection of the data in Figure 2. All the effect sizes in this figure are small, suggesting that the opposite of
publication bias might have occurred in this analysis. That is, we might have failed to find studies with high magnitude effect sizes in the positive or negative direction.

Fourth, the current study relies on a linear model to determine the relationship between female representation and firm financial performance. It is possible that such a relationship depends on multiple factors in a non-linear fashion and that this changes over time. Thus, the linear assessment of data collected over a number of years might have contributed to the low effect sizes in the current meta-analysis. In general, it is difficult to determine the correct analytical approach to such a complex topic. Specifically, the relationship between female representation and firm financial performance was low, not only on a univariate study level, but also after the application of multivariate linear and non-linear approaches in the primary studies included in our analysis. Thus, the relationship between gender and financial performance might be truly negligible compared to other factors that might affect financial performance. Thus, future studies should focus on devising novel analytical approaches to study this topic.

Fifth, the data in our analysis come from countries with differing legal and board systems. Most pronounced is the difference between the one-tier and the two-tier board system, with various smaller differences between countries. In a one-tier board system, which is prevalent in the United States, the board is solely responsible for all corporate decisions. Inside directors, who are directly employed in the company, represent the interests of the company’s stakeholders, while outside directors, who are usually employed in other companies, bring a different perspective and objectivity to the boardroom. A two-tier board system, which, for example, is common in Germany, consists of an executive board and a supervisory board. The executive board manages the day-to-day business and the supervisory board supervises the executive board’s decisions. We did not investigate the influence of these factors, because there are too many minor differences between countries and they are too widespread to classify. While we do not expect these country differences to have a large influence on the current results, they might have contributed to the high heterogeneity in the studies’ effect sizes. Meta-analytically investigating the differing role that the diversity of internal and external directors might play in the boardroom with regards to firm performance could provide a future contribution to the scientific literature.

Sixth, two studies included in our sample were classified as statistical outliers due to their extreme effect sizes in opposite directions (Abdillahi & Manini, 2013; Wellalage & Locke, 2013). The removal of these studies did not change the outcome of our analysis. These studies possibly were outliers, because both included a relatively small number of firms, nine and 88,
respectively (compared to the mean number of firms of 146 in all the studies; see Table 3). This might have increased the influence of individual firms and thereby skewed those studies’ results.

Seventh, it is possible that not all 20 studies were based on independent data, meaning that some firms might have contributed more data to the overall effect size than others. There was a possible overlap in the included firms from the same countries, because the primary studies did not rely on single firm data, but rather on data from business databases from the same country. For example, data from Norway were included in three studies in the current analysis (Ahern & Dittmar, 2012; Bøhren & Staubo, 2014; Dale-Olsen et al., 2013). To reduce this potential overlap in the data, we only once included data from the same time period, utilizing the same outcome measure in the same country.

Eighth, an overall higher representation of females might be needed in order to identify a relationship between diversity and performance. A shortcoming of the included data is that they are restricted in range, because the largest percentage of female directors included in our sample was 37% and the average was about 14%. This limits the meaningfulness of our findings, because few female directors are present in all the studies in general. This result supports the notion that more females should be promoted to director positions to meaningfully investigate the effects of gender diversity on performance. In accordance with this proposition, Joecks and colleagues (2012) suggest that a relatively low representation of females on boards first has a negative effect on firm performance (contradicting our findings), which only becomes positive after a critical mass of 30% female directors is reached. Thus, the current representation of 14% females on corporate boards in our sample might not be sufficient to show either the positive or negative effect that increased gender diversity might have on firm performance.

Lastly, women generally experience higher levels of chronic stress than men (Matud, 2004), especially when working in male-dominated industries, where they suffer from increased stress and worse mental health than men (Gardiner & Tiggemann, 1999) and might be perceived as tokens (Kanter, 1977). These differences in experienced stress might influence the relationship between gender diversity and firm performance, depending on the industry in which it is measured. For example, it would be interesting to examine if women’s influence is more pronounced in traditionally female-dominated industries as opposed to male-dominated industries. Similarly, Post and Byron (2015) also suggest that female directors might influence firm performance stronger in customer-proximal industries.
Conclusion

In recent years, interest in gender diversity and the representation of females in leadership positions has evidenced a steep increase. Many scientific studies have investigated the relationship between female representation on corporate boards and firm financial performance, but, so far, the results are contradictory. The results of the current meta-analysis show that a higher representation of females on corporate boards is neither related to a decrease, nor to an increase in firm financial performance, confirming findings from a similar meta-analysis on this topic (Post & Byron, 2015). These results do not support the business case for diversity, which suggests that diversity is associated with an increase in performance. However, they allow the conclusion that gender diversity should be promoted for ethical reasons to promote fairness. If a larger representation of female directors does not matter with regard to firm performance, females, if equally qualified, should be given priority when promotion decisions are made.