Gender Diversity in Boards: an empirical analysis of performance differences

The effect of gender diversity in boards on firm performance is a trending area of research. In this paper, data from 67 Dutch firms listed on the Amsterdam Euronext Stock Exchange are utilized to test this effect. Based on the token status, agency, resource dependence and critical mass theories, a positive effect of gender diversity in boards on firm performance is hypothesized. Using a firm and time fixed effects methodology, the firms are compared in the period 2010-2018. Gender diversity is measured in three different manners, including a new measure unseen in previous literature named the equality index. The results of the models show no statistically significant effect of gender diversity on firm performance for all three measures of gender diversity. Therefore, it cannot be concluded that gender diversity has a positive effect on firm performance.

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis Economics and Business Economics

Final draft

Name student: Laurence Nowak

Student ID number: 476267

Supervisor: dr. A.A.F. Gerritsen

Second assessor: dr. A.L. Boring

Date final draft version: 03/07/2020

Contents

1. Introduction

Since the first of January 2013, the Dutch law contains the Act on Management and Supervision, which requires large Dutch firms to strive for a balanced distribution of board positions between men and women (Leenders, Pouwels, & Van den Brink, 2019). In practice, a balanced distribution means that the seats in management and supervisory boards of large Dutch firms should be filled with at least 30 percent women and at least 30 percent men. On a yearly basis, firms are categorized as large depending on their asset value, net turnover and average number of employees. If a large firm does not reach the 30 percent target of both genders, it is obliged by law to explain in its annual report why it did not reach the target, which attempts were made to reach the target and how it plans to achieve the target in the near future.

The Act on Management and Supervision was voted for by the Dutch parliament to encourage firms to move towards more equal opportunity within the Netherlands. In 2007, the share of female board members was just 4.5 percent on average in the top 5000 firms (Dirven & Merens, 2018). There are various explanations for this inequality, such as the part time working culture, risk aversion and competition aversion, thinking in stereotypes and the existing organisational culture (Ebregt, Jongen, Lanser & Merens, 2019). The relative importance of these explanations is difficult to measure. What seems even more difficult, is for firms to reach the target share of women on their boards. Four years after the law was introduced, the share was still far below the target; 15.4 percent on average (Dirven & Merens, 2018). Large profit-minded firms might not be so interested in values of equality and attach more value to their profits. Hence, will meeting the gender target, potentially resulting in a higher firm performance, convince firms to employ more women in top positions? The research question of this paper is as follows:

What is the effect of gender diversity in boards on firm performance?

This question will be researched using data of 67 Dutch firms which are listed on the Amsterdam Euronext Stock Exchange during the period 2010-2018. This paper contributes to the academic literature by looking at a more recent period in the Netherlands compared to the existing literature. The share of women on boards has increased by 10.9 percent between 2007 and 2017 in the Netherlands (Dirven & Merens, 2018). In addition, a new measure for gender diversity is introduced, called the equality index, which is believed to be more representative compared to measures used in previous literature. In previous studies, gender diversity is often measured as the percentage of female directors. By using this as a measure, a board with only females would have the highest value of gender diversity. In this paper, the highest value of gender diversity is achieved when the percentage of female directors equals the percentage of male directors.

First, in the related literature section, the positive link between board performance and firm performance is established. Then, three underlying theories of the effect of gender diversity in boards on

firm performance theories are discussed which are the token status theory, agency theory and resource dependence theory. Through the token status theory, it is argued that the stereotype of women having a lower quality of leadership for board positions leads to fewer female directors being employed because of their gender (Lee & James, 2007). Because this discrimination towards women in top functions is based on their gender and not their ability, a higher share of women can lead to better firm performance (Neumark, 2018). This theory is extended to the critical mass theory which explains that three or more women in a board has the possibility to create a critical mass which can positively influence firm performance (Kramer, Konrad, Erkut & Hooper, 2006). The critical mass theory applied to gender diversity is tested for the first time in the Netherlands. Via agency theory it is argued that the interests of female directors are more likely to be in line with the interest of the firm and shareholders compared to male directors (Liu, Wie & Xie, 2014). Resource dependence theory is used to argue that female directors have access to different resources which can benefit firm performance. In the same section, three empirical papers by Smith, Smith and Verner (2006), Carter, D'Souza, Simkins and Simpson (2010) and Lückerath-Rovers (2013) are highlighted, an overview of the relevant empirical literature is presented and the hypotheses are formulated. Based on the underlying theories, a positive relationship between gender diversity and firm performance is hypothesized using three different measures of gender diversity. These hypotheses test the effect of the equality index, a female dummy and a critical mass dummy on firm performance.

Second, in the data and methodology section, the data is introduced and the variable measures are discussed. In addition, the methodology and regression equations are explained and the descriptive statistics are presented. Tobin's Q, return on assets (ROA) and return on equity (ROE) are utilized to measure firm performance. Firm size, firm age and board size are added as control variables. A firm and time fixed effects methodology is used to answer the research question. Third, in the results section, the relevant results are discussed and the estimated coefficients are interpreted. The results of the models show no statistically significant positive effect of gender diversity on firm performance in all three cases. Therefore, it cannot be concluded that gender diversity has a positive effect on firm performance.

Next, in the robustness section, analyses to identify problems with omitted variables, reverse causality and outliers are performed and discussed. Afterwards, in the discussion section, the main outcomes and weaknesses of the analyses are debated. The sample, time period of the data, external validity and issues with causality are critically analyzed. Finally, in the conclusion section, the results are summed up and suggestions for future research are made.

2. Related literature

2.1 Theories

First of all, the connection between board functioning and firm performance must be established. Finkelstein, Hambrick and Cannella (1996) argue boards have two main functions which have a large connection to firm performance. First, in general, boards determine the strategy direction of a firm and have a large influence on decision-making within the firm. In addition, boards function as a monitoring body for areas such as monitoring the use of the firm's wealth, representing the firm's shareholders, responding to take-over threats and hiring and monitoring the top managers who are not part of the board. They thus make important decisions for the firm, but also monitor the progress of those decisions and have the power to intervene in processes of the firm when needed. After linking board functioning and firm performance positively, comes the link between board gender diversity and firm performance. Many theories have been developed and applied to the area of research of women on corporate boards (WOCBs). Terjesen, Sealy and Singh (2009) divide the theories applied to WOCBs into four levels of perspective: individual, board, firm and industry/environment. The individual level refers to the director, or aspirant director. The board level represents the processes and interactions that occur within the board. The industry/environment level contains local and national industrial and external environments. The following section will focus on the firm level as firm financial performance is part of this level and thus is the most relevant level to the research question. This level captures the strategy and structure of the firm and the events that happen outside of the board but within the firm. At the firm level, the most convincing theories are the token status theory, agency theory and resource dependence theory (Liu et al., 2014).

First, Kanter (2008) explains the token status theory as the situation where the dominant group has a distorted image of the minority. He refers to the minority as tokens or solos, which are defined as individuals who are the only representative of a demographic group. The theory suggests that the dominant group distorts the image of female token directors in ways which are closer linked to femininity and further away from the qualities of leadership. This leads to sex-role stereotypes which are inconsistent with the dominant group's perceptions of leaders. The consequence is that male job applicants are more likely to be associated with qualities of leadership compared to female job applicants, because men hold most director positions (Powell & Butterfield, 2002). If there however is a case of discrimination towards women in top functions, a higher share of women can lead to better results (Neumark, 2018). Tokenism of WOCBs also reinforces the stereotype of women having a lower quality of leadership for board positions (Lee & James, 2007). Due to tokenism and the sex-role stereotypes of female directors, the impact of a lone female director on corporate decisions is likely to be restricted (Liu et al., 2014).

Extending this theory, applying the critical mass theory to board gender diversity postulates that three or more women on a board has the possibility to create a critical mass, where women have a substantial influence on the content and process of board discussions (Kramer et al., 2006). When the critical mass is reached, the impact female directors might have on firm performance should be more noticeable. Joecks, Pull and Vetter (2013) test the critical mass theory empirically using 151 German firms over a period of 6 years. They find evidence that gender diversity at first negatively affects firm performance with use of a random effects methodology. However, once a female percentage of around 30 is reached, these firms perform better than firms with full-male boards. Strengthening the theory of Kramer et al. (2006), 30 percent translates into around 3 women on a board based on their sample of firms.

Second, agency theory can be described by the relationship between the principal and the agent of the principle (Terjesen et al., 2009). Regarding firms, agency problems occur when the best interest of the director (agent) is not in line with the best interest of the shareholder (principal) (Liu et al., 2014). According to Fama and Jensen (1983) efficient board guidance and monitoring are of great importance to mitigate the conflict of interests. The theory assumes that outside directors will act more independently compared to inside directors, as they tend to be more active in monitoring activities and therefore be better monitors for the interests of the shareholders (Liu et al., 2014). Outside directors will thus cause board decisions to be pointed more towards the interest of the shareholders instead of individual interest of the directors (Bianco, Ciavarella & Signoretti, 2015). Directors are classified as outsiders if they do not have a business relation with the firm, are not current or former employees of the firm and are not related or interlocked with management (Adams & Ferreira, 2009; Fields & Keys, 2003). Inside directors do have connections to the firm via at least one of those channels.

Empirical research by Adams and Ferreira (2009) and Farrell and Hersch (2005) shows that female directors overwhelmingly serve on boards as outside directors. Female directors will consequently cause decisions of the board to be pointed more towards the interest of the shareholders. Adams and Ferreira (2009) find a positive relationship between gender-diverse boards and the demand for more audit efforts and managerial accountability. They also find that the quality of a firm's governance has a negative relationship with the effect of board gender diversity on corporate decisions. On the other hand, board gender diversity can hurt the value of a well-governed firm through over-monitoring which was unnecessary in the first place.

Third, resource dependence theory postulates that a firm needs resources from their external environment to survive (Pfeffer & Salancik, 2003). Being dependent on these resources forms a risk for the firm. By forming connections between the external environment and the firm, the firm can reduce dependency and the risks that come with it. This dependency on the external environment requires a board

with access to a broad selection of resources, such as financing, industrial/functional/geographic knowledge, prestige, legitimacy, and diversity (Terjesen et al., 2009).

Pfeffer and Salancik (2003) claim there are three areas of benefit that arise from forming connections between the external environment and the firm: advice and counsel, legitimacy and communication channels. Regarding advice and counsel, Kravitz (2003) and Huse and Solberg (2006) suggest that gender-diverse boards have a positive impact on the quality of discussions on complex topics. In addition, Smith (2018) finds that new perspectives within boards can be a consequence of a higher share of women. Hillman, Cannella Jr. and Harris (2002) found that Fortune 1000 female directors are more likely to hold advanced degrees and have non-business backgrounds compared to men. Women also represent diversity, soft values and women's issues (Huse & Solberg, 2006). Furthermore, women have stronger connections to other boards with female directors, that can offer advice (Hillman, Shropshire and Cannella Jr., 2007). On the other side, more diversity can lead to more disagreement, which can hinder the decision-making of the board (Adams, De Haan, Terjesen & Van Ees, 2015). Regarding legitimacy, firms can legitimize their practices by accepting societal norms and values (Liu et al., 2014). Directors of large firms are highly visible to societal actors, such as institutional investors and political parties, who can grant legitimacy and are influenced by prestige (Hillman et al., 2007). Directors can thus legitimize firms. In terms of communication channels, female leaders are more capable of connecting their firms to female customers, workers and women in society at large due to their different perspectives and life experiences (Liu et al., 2014). Hillman et al. (2007) apply the resource dependence theory and find that gender-diverse boards in the United States can accumulate the benefits of advice and counsel, legitimacy and communication channels. This theory thus suggests beneficial effects of gender-diverse boards.

2.2 Empirical literature

In addition to the theoretical research, many empirical studies have been conducted to examine the effect of gender diversity in boards on firm performance. Table 2.1 contains a chronological overview of the relevant empirical literature on this topic. The results of these studies differ vastly, which can partially be explained by researchers using different datasets, methods and measures of variables. In general, methods using panel data and listed firms are most common. Although Tobin's Q and ROA are the most popular firm performance measures, many other measures are also used such as return on shares, ROE and return on investment. There are large differences in the amount and selection of control variables which are used to avoid omitted variable bias. For example, Liu et al. (2014) use eleven control variables that represent board characteristics, ownership characteristics and firm characteristics. Lückerath-Rovers (2013) includes just three control variables, board size, firm size and a financial sector dummy. To convey the differences

between the literature, three empirical studies will be discussed in detail. The studies by Smith et al. (2006) and Carter et al. (2010) were selected based on their high number of citations, high number of observations and relatively high number of firm years in the dataset. The study by Lückerath-Rovers (2013) was selected for its representativeness towards this study, with it being the most highly regarded study in this area using Dutch listed firms. In addition, all three studies differ in their choice of performance measures.

Table 2.1 Chronological overview of the empirical literature

Author(s) (year)	**Gender diversity measure**	**Performance measure**	**Database (*n*, country, years)**	**Method**	**Main result**
Marinova et al. (2016)	Women's ratio, dummy (women on the board: yes/no)	Tobin's Q	102 Dutch and 84 Danish listed firms (2007)	2SLS	No link
Liu et al. (2014)	Women's ratio, number of women	ROA, ROS	2000+ Chinese listed firms (1999-2011)	Panel regression with firm fixed effects	+
Joecks et al. (2013)	Blau index	ROE	151 German listed firms (2000-2005)	OLS, random effects	+ critical mass, - no critical mass
Lückerath-Rovers (2013)	Dummy	ROE, ROS	99 Dutch listed firms (2005-2007)	Correlation, panel regression	+ ROE, - ROS
Mahadeo et al. (2012)	Women's ratio	ROA	39 Mauritian listed firms (2007)	OLS	+
Bøhren & Strøm (2010)	Women's ratio	ROA, ROS, Tobin's Q	203 Norwegian listed firms (1989-2002)	Fixed effects	-
Carter et al. (2010)	Number of women	ROA, Tobin's Q	641 US listed firms (1998-2002)	OLS, 3SLS	No link
Dobbin & Jung (2010)	Number of women	ROA, Tobin's Q	432 US listed firms (1997-2006)	Firm and year fixed effects	No link profits, - stock value
Haslam et al. (2010)	Dummy, women's ratio	ROA, ROE, Tobin's Q	126 British listed firms (2001-2005)	ANOVA	No link ROA, ROE, - Tobin's Q

Table 2.1 Continued

Author(s) (year)	Diversity measure	Performance measure	Database (*n*, country, years)	Method	Main result
Adams & Ferreira (2009)	Dummy women's ratio	ROA, Tobin's Q	1939 US firms (1996-2003)	IV panel regression	-
Miller & del Carmen Triana (2009)	Blau index	ROI, ROS	326 US listed firms (2003)	OLS	No link
Campbell & Mínguez-Vera (2008)	Dummy, women's ratio, Blau index, Shannon index	Tobin's Q	68 Spanish firms (1995-2000)	2SLS panel regression	+ women's ratio, Blau and Shannon index, no link dummy
Rose (2007)	Women's ratio	Tobin's Q	100+ Danish listed firms (1998-2001)	Cross-sectional OLS	No link
Smith et al. (2006)	Women's ratios	CM/NS, GP/NS, NIAT/NA, ONI/NA	2500 largest Danish firms (1993-2001)	Fixed effects panel (IV) regression	+ subject to performance measure
Carter et al. (2003)	Dummy, women's ratio	ROI, Tobin's Q	638 US listed firms (1997)	2SLS	+ (Tobin's Q)
Erhardt et al. (2003)	Minorities' and women's ratio	ROA, ROI	127 large US firms (1993 and 1998)	Correlation and hierarchical regression	+ (including demographic diversity)
Shrader et al. (1997)	Women's ratio	ROA, ROE, ROI, ROS	200 US firms (1992)	Hierarchical regression	

Note. + = positive link, - = negative link, 2SLS = two stage least squares, ROA = return on assets, ROS = return on shares, ROE = return on equity, OLS = ordinary least squares, 3SLS = three stage least squares, IV = instrumental variable, ROI = return on investment, CM = contribution margin, NS = net sales, GP = gross profit, NIAT = net income after tax, NA = net assets, ONI = operating net income.

Smith et al. (2006) estimate the relationship between management diversity and firm performance. The 2500 largest Danish firms are studied between the period of 1993 to 2001. With Denmark being a relatively small country, such a large dataset also includes smaller firms. Therefore, the external validity is larger compared to previous empirical literature, where often only large (listed) firms are studied. Four alternative measures for firm performance are analyzed to test the robustness of the results: Gross profit/net sales, contribution margin/net sales, operating net income/net assets, net income after tax/net assets. Management diversity is measured as both the proportion of women among CEOs and the proportion of

women on boards of directors. In addition, control variables such as firm size, firm age and export orientation are added. Smith et al. (2006) apply the Hausman test and determine that a fixed effects model is more suitable than a random effects model.

To deal with potential problems of endogeneity, an instrumental variable (IV) is used: the average length of education of the spouses of other CEOs of the firm. The instrumental variable must have a high correlation with the independent variable, in this case the proportion of women in management, but no correlation with the dependent variable, in this case firm performance. If this is not the case, the IV will not project the full effect of the proportion of women in management on firm performance. Smith et al. (2006) hypothesize that CEOs who are married to spouses with a higher length of education are more positive and have a less traditional view on the abilities of female CEOs. Therefore, they are more willing to hire and accept a female CEO compared to CEOs who are married to spouses with a shorter length of education. Furthermore, the IV is assumed not to affect firm performance. This reasoning is in line with the token status theory, as the stereotype for women having a lower quality of leadership for board positions is more likely present for a CEO with a spouse who has a shorter length of education. Therefore, the choice of the instrumental variable is justified. However, due to lack in data this instrumental variable can only be applied to the models analyzing the effects of only female CEOs and not the effects of all female directors.

The results of Smith et al. (2006) are mixed and depend on both the management diversity and firm performance measure. Gross profit/net sales and contribution margin/net sales are more positively and significantly affected by the proportion of female CEOs than operating net income/net assets and net income after tax/net assets. These effects on performance are related to female CEOs with a university degree. Female CEOs without a university degree have a much smaller or insignificant effect. For female directors on boards, Smith et al. (2006) find a difference between female directors which were elected by employees of the firm and other female directors. The female directors which were elected by employees, had a significant positive effect on all four performance measures. The other female directors had a significant negative effect on operating net income/net assets and net income after tax/net assets and an insignificant effect on gross profit/net sales and contribution margin/net sales. An explanation for this is that the other directors are family members of members of the board and are therefore not employed for their abilities but their relationships with individual members of the board. However, this cannot be tested as data on family ties of board members was not available.

Carter et al. (2010) examine the relationship between the number of women and ethnic minority directors and firm financial performance. A sample of 641 US firms which are listed in the S&P 500 index is utilized over the period of 1998 to 2002. Tobin's Q as well as ROA is used to measure of financial performance. Carter et al. (2010) estimate their model by using both an ordinary least squares (OLS)

regression and a three stage least squares (3SLS) regression with firm and time fixed effects. To address potential problems of endogeneity, lagged dependent variables are included. The 3SLS estimation avoids problems with reverse causality. 3SLS estimation is chosen above two stage least squares (2SLS) estimation because it addresses both potential endogeneity and cross-correlation between equations. The results of the OLS regression give a significant positive relationship between ROA and the number of women on the board and no significant relationship between Tobin's Q and the number of women on the board. However, the results of Hausman Tests suggest the need for the 3SLS regression results to answer the hypotheses (Carter et al., 2010). In this regression the coefficient of number of women on the board is insignificant for both Tobin's Q and ROA. Therefore, the 3SLS results provide no support for a link between gender diversity of the board and financial performance.

In the Netherlands, Lückerath-Rovers (2013) studied 99 Dutch firms which are listed on the Amsterdam Euronext Stock Exchange for the period of 2005 to 2007. Three factors are determined which can explain the varying results of previous empirical studies: time, causality and critical mass. Regarding time, Lückerath-Rovers (2013) identifies differences in the measure of gender diversity. A static measure of number of women in the board can lead to other results than the change of number of women in the board. Ryan and Haslam (2005) also find a lower firm performance during the 5 months prior to the firms that appointed a female board member compared to firms that appointed a male board member. This suggests women are more likely to be appointed to poor performing firms, which can bias the results of a study negatively which does not consider firm performance in previous periods. Regarding causality, in times of bad firm performance, shareholders are more likely to influence the decisions of directors (Lückerath-Rovers, 2013). Therefore, they might increase the pressure to have a more diverse board. This pressure is even higher if the shareholders believe a less gender-diverse board is less critical. Critical mass theory can also influence results, but in the case of Lückerath-Rovers (2013) it is difficult to be tested due to the limited number of boards in the sample with more than one female director.

In the data, it is common that many firms have either 0 or 1 female(s) on their board. Consequently, a dummy variable is used to measure board gender diversity. For firm performance multiple measures are utilized such as ROE, return on shares (ROS), return on invested capital, operating result, stock price growth and total stock revenue. In this research two different correlation analyses which build on the methods performed by Joy, Wagner and Narayanan (2007) and Desvaux, Devillard-Hoellinger and Baumgarten (2007) are used. In addition, Lückerath-Rovers (2013) performs a regression analysis with control variables. For the regression analysis the control variables include the logarithm of total assets, board size and a financial sector dummy. The dummy is included because the firms in the financial sector in the dataset are on average the largest firms but also have the highest number of female directors.

Using the method of Joy et al. (2007), Lückerath-Rovers (2013) finds a statistically positive relationship between return on shares and ROE and having a female in the board by comparing the mean of firms with and without female directors. The results for return on invested capital, total stock revenue and stock price growth are insignificant. Using the method of Desvaux et al. (2007), where the mean of all firms is compared with firms with at least one female directors, the results only show a significant positive correlation between ROE and the firms with at least one female director. Total stock revenue and stock price growth show an insignificant correlation. The regression analysis is used to further explore the relationship between ROE and the presence of female directors (Lückerath-Rovers, 2013). The analysis finds a significant positive relationship between ROE and the female directors dummy as well as the share of female directors. It is, however, made clear that these relationships cannot be treated as causal with certainty, as there might be other factors that influence firm performance.

2.3 Hypotheses

As previously discussed, this paper will attempt to contribute to the empirical literature by looking at a more recent period in the Netherlands, where firms have employed more female directors than before (Dirven & Merens, 2018). In addition, a new measure for gender diversity is introduced which was not found in any of the empirical studies in table 2.1. This measure, called the equality index, is believed to be more representative of gender diversity in boards compared to the percentage of female directors in a board, which was a common measure in previous literature. It is utilized in the first hypothesis. To allow for a broader understanding of gender-diverse boards and firm performance, three hypotheses have been formulated.

Hypothesis 1: Firm performance is higher when a board has a higher equality index.

The accompanying null hypothesis for hypothesis 1 is that firm performance is equal or lower when there is a higher equality index in boards. The equality index is calculated via the following steps:

$$If\ the\ number\ of\ females\ in\ the\ board\ \leq \frac{Board\ size}{2}$$

$$Then\ equality\ index = \frac{Number\ of\ females\ in\ the\ board}{Board\ size/2} \times 100\%$$

$$Otherwise\ equality\ index = \frac{Board\ size - Number\ of\ female\ board\ members}{Board\ size/2} \times 100\%$$

The equality index is a way to measure the share of female directors in a board. The reasoning for it being believed to be more representative of gender diversity in boards is explained in section 3.2.2.

Hypothesis 2: Firm performance is higher when there is at least one female in a board.

The accompanying null hypothesis for hypothesis 2 is that firm performance is equal or lower when there is at least one female in a board. This hypothesis is helpful to answer the research question, because a large share of firms in the dataset that have at least one female, have just one female. Out of all firms with females, 53.09 percent of their firm years have one female in the board. Even with one female director, there is more gender diversity than with zero, and therefore the positive effect is hypothesized. The difference between in testing between hypothesis 1 and 2, is that hypothesis 1 compares the firms in their share of female directors, whereas hypothesis 2 compares firms with and without female directors.

Hypothesis 3: Firm performance is higher when there is 30 percent or more females in a board.

The accompanying null hypothesis for hypothesis 3 is that firm performance is equal or lower when there is 30 percent or more females in a board. By testing this hypothesis, firms that have reached the critical mass of 30 percent will be compared to firms with female directors that have not reached the critical mass and firms with no female directors.

3. Data and methodology

In this section, the sample of Dutch firms will be described. Next, the different measures of dependent, independent and control variables are discussed. Then, the regression equation using a firm and time fixed effects methodology is formed. Afterwards, the effect of sectors is discussed and issues of reverse causality and outliers are considered. In this section the descriptive statistics are presented during the introductions of variables. To understand these statistics better, the control variables of firm size, board size and firm age were not yet transformed into natural logarithms.

3.1 Sample description

In this paper data of 100 Dutch firms during the period 2010-2018 will be utilized, which are listed on the Amsterdam Euronext Stock Exchange on the first of May 2020. After narrowing down the dataset, as described next, 67 of the 100 firms are used for the main regression. The data was gathered via the database Orbis by Bureau van Dijk Electronic Publishing. Orbis consists of firm data collected on a worldwide scale with information on more than 365 million firms (Bureau van Dijk, 2020). Firms which are located in differing countries have to oblige to other tax law systems which can affect after-tax performance measures,

such as ROA (Lückerath-Rovers, 2013). There are also many countries with binding gender quotas on boards of directors, such as Belgium, Finland, France, Germany, Iceland, Italy and Norway (Smith, 2018). Firms will have a higher share of women in their board if the law obliges them to have a certain share of women. A binding quota might also affect firm performance, as firms will have to invest more time and money into making sure they reach the quota. If a firm does not reach the quota, it will also have to pay a fine. So, due to differences between countries in tax law systems and gender diversity in boards, only firms which were legally located in The Netherlands were included. In addition, private firms were excluded as they are not obliged by law to release reports to the public and therefore merely public listed firms are included (García Álvarez, 2016).

Once the dataset had been narrowed down, the values of Tobin's Q were available for 71 firms for six or more years of the 9 years for the period 2010-2018. The data before 2010 had significantly more missing values and many of the firms had not released their results for years after 2018 on the first of May 2020. To measure gender equality, the number of women on boards is of great importance. Orbis includes access to annual reports for most of the 71 firms, where firms report the number and gender of board members. If the gender of a board member was unclearly stated, the gender pronouns were looked up in the annual report. If it was still unclear, the director was looked up on LinkedIn. The missing annual reports were found on the firm's websites. There were however a few firms with annual reports that could not be found via Orbis or their websites, so the final number of firms was narrowed down to 67 firms.

3.2 Variable measurement

3.2.1 Firm performance

The dependent variable of the research question is firm performance. Haslam, Ryan, Kulich, Trojanowski and Atkins (2010) identify two categories of measures which emphasize different outcomes of firms: stock-based and accounting-based measures. A common stock-based measure is Tobin's Q and two common accounting-based measures are ROA and ROE for firm performance in the related literature, as can be found in Table 2.1. These measures are commonly used in governance investigations (Carter et al., 2010). First, *Tobin's Q* is calculated in Orbis by dividing market capitalization by total assets. Market capitalization equals the total number of shares outstanding multiplied by the current market price of each share. To an extent, Tobin's Q forecasts future cash flows of a firm. If the value of Tobin's Q is higher than 1, investors expect firms to create more value by utilizing the assets it has available effectively (Campbell & Mínguez-Vera, 2008). Thus, a value of Tobin's Q below 1 is associated with poor usage of available assets. Wernerfelt and Montgomery (1988) claim Tobin's Q is a good measure for firm performance as it reflects what the market expects from earnings in the future. Additionally, it accounts for the risk investors take by

investing in the firm, as this risk is part of the market price of the shares, and cannot be subject to reporting inconsistencies between firms because of differences in accounting practices or tax laws (Lindenberg & Ross, 1981). Demsetz and Villalonga (2001) and Dezsö and Ross (2012) argue for Tobin's Q above accounting-based measures, as Tobin's Q focusses on expectations of future performance whereas accounting-based measures are based on previous events and thus give insight into past performance.

This paper will however also consider accounting-based measures to conduct robustness analyses. In addition, using these measures reduces the possibility of the capital structure of the firm simply reflecting the observed values of performance (Haslam et al., 2010). Second, *ROA* indicates the operating ability of a firm to generate accounting-based income using its assets (Carter et al., 2010). It gives shareholders a sign of what the firm might be able to generate with an investment. ROA is calculated by adding up operating profit and depreciation and then dividing by total assets. Orbis defines operating profit as all operating revenues minus all operating expenses and depreciation as the total amount of depreciation and amortization of assets. It therefore does not include interest expenses. Third, *ROE* indicates the financial or capital ability of a firm to generate profits using the investments of shareholders. ROE is calculated in Orbis by dividing net income by shareholder funds. Net income is defined as operating profit plus financial profit minus all taxes related to the accounting period. The difference between ROA and ROE is that ROE does not contain debt, whereas ROA does. In our calculation, interest expenses are included in ROE but not in ROA, because ROE uses shareholder funds in the denominator and from the perspective of a shareholder interest expenses are real expenses. ROA and ROE are commonly used to indicate the earnings of a firm and returns to shareholders (Shrader, Blackburn & Iles, 1997). They also provide a basic measure for the overall profitability of firms.

Table 3.2 shows significant negative relationships between Tobin's Q and ROA (-0.44) and between Tobin's Q and ROE (-0.24) at the 5% level. It is however important to keep in mind that Tobin's Q focusses on expectations of future performance, whereas accounting-based measures give insight into past performance (Demsetz & Villalonga, 2001; Dezsö & Ross, 2012). Therefore, even though all three measures represent performance, they portray different moments in time. So, a positive relationship is not necessarily to be expected. There is a significant positive relationship at the 5% level between ROA and ROE (0.68). These measures do measure the same period. Figure 3.1 shows how Tobin's Q, ROA and ROE developed on average over the studied period. Tobin's Q was at its lowest point in 2011, meaning that shareholders had low expectations of future performance on average. ROA was below zero in 2012 and ROE was below zero in 2013, which gives reason to believe the predictions of the shareholders were possibly correct. The positive correlation between ROA and ROE in table 3.2 does not seem to hold in each

year. For example, between 2011 and 2012 ROA decreases whereas ROE increases and between 2012 and 2013 vice versa.

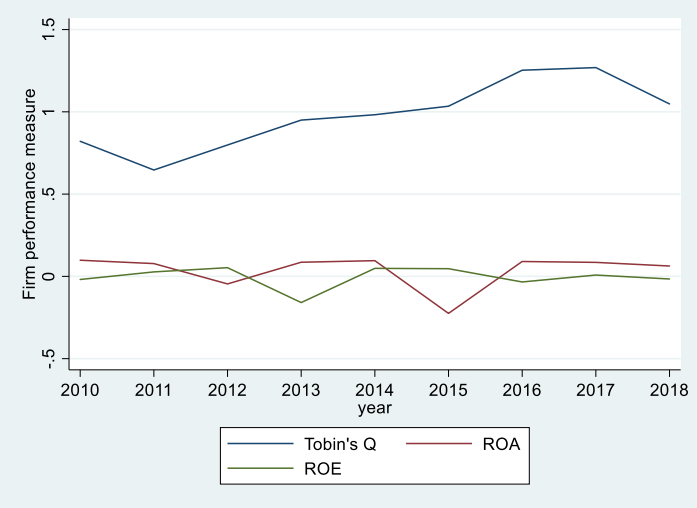

Figure 3.1 Development of means of firm performance measures, 2010-2018

3.2.2 Gender diversity

Following from the resource dependence theory, agency theory and token status theory, more diversity in boards might lead to better firm performance. For hypothesis 1, it is important to calculate a share of female directors in a board. In the related literature, this share is often calculated as the total number of women divided by the total number of board members. The flaw in using this formula is that a board with solely females will have the highest amount of gender diversity. In this paper the share will be calculated differently and has been named the *equality index*. This is more representative of gender diversity, as board is seen as fully gender-diverse if the number of male directors equals the number of female directors. The steps to calculate the equality index are described in section 2.3. The equality index will equal 100 if the number of women on the board is equal to half of the total amount of board members and will equal 0 if the number of women on the board is equal to 0. For hypothesis 2, a dummy variable will be used which equals 1 if a firm has at least one female director in the board and 0 if a firm has no females in the board. This dummy is called *female dummy*. For hypothesis 3, firms will be split into three groups represented by dummy variables. One dummy variable will equal 1 if the firm has reached the critical mass of at least 30 percent female directors in the board, and 0 otherwise. This dummy variable is called *critical mass*. The other dummy variable will equal 1 if the firm has more than 0 percent and less than 30 percent female directors in the board, and 0 otherwise. This dummy variable is called *small share*.

Table 3.2 shows a positive correlation between the equality index and all firm performance measures, but this is only significant at the 5% level for ROE (0.17). In fact, ROE is also the only firm performance measure with a significant correlation with the female dummy (0.17), critical mass dummy (0.091) and small share (0.11) at the 5% level. Table 3.3 shows that the average board had an equality index of 25.06 and thus 12.53% females in the board. Out of 599 observations, 59% of the boards had at least one female director. This 59% consists of 11% boards with 30% or more females, and 48% boards with between 0 and 30% females. Figure 3.2 shows the development of the equality index over time. The average equality index has increased vastly over time, with an increase of more than 20%. Between 2013 and 2014, there is a large increase of around 5%. This is the year when the Act on Management and Supervision was introduced, which might have provided firms an incentive to employ more female directors. Figure 3.3 shows an increase in the number of firms with at least one female in their board. The average of female dummy increased from just below 0.4 in 2010 to just below 0.8 in 2018. In addition, the amount of boards that reached the critical mass increased greatly. In 2010, less than 5% of the firms reached the critical mass, whereas this was just over 20% in 2018. There is no clear pattern between Tobin's Q and the equality index, as can be seen in figure 3.4. This is in line with the aforementioned insignificant correlation between the two variables. The scatterplots of ROA and ROE and the equality index are included in Appendix A in figure 3.6 and figure 3.7. Figure 3.7 shows the positive significant correlation between ROE and equality index might be caused by a few outliers with no females in the board. This will be discussed in section 3.5.

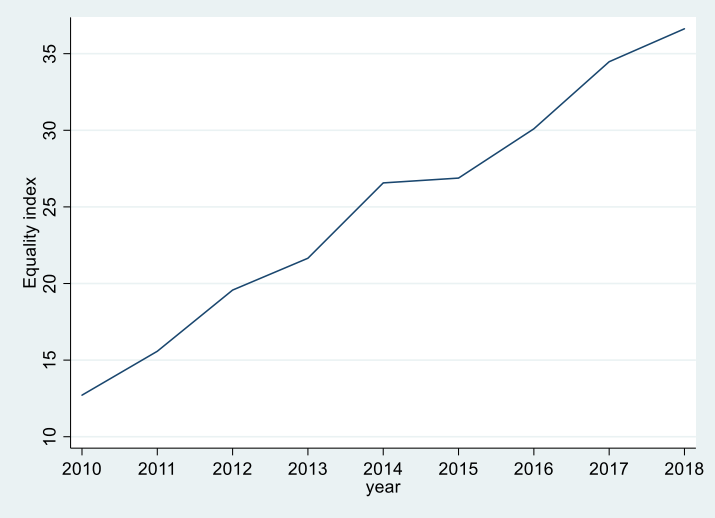

Figure 3.2 Development of mean of equality index, 2010-2018

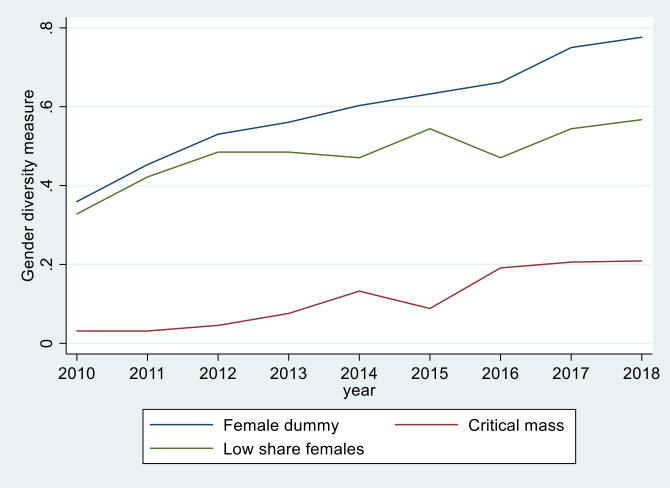

Figure 3.3 Development of means of gender diversity dummies, 2010-2018

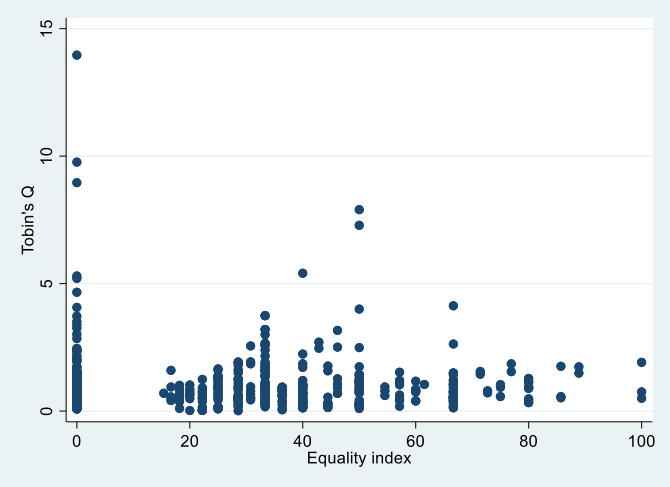

Figure 3.4 Scatterplot of Tobin's Q and equality index

3.2.3 Control variables

A large concern of the fixed effects methodology is the problem of omitted variable bias. Omitted variable bias is present when there are time-varying variables that are not included in the regression analysis which influence both gender diversity and firm performance. To avoid this problem, control variables can be included in the regression analysis. By including control variables, the coefficient of the measure of gender diversity will be less biased, which will be explained next. Firm size, firm age and board size will be included as control variables. First, Firm size, measured as the natural logarithm of the number of employees, is expected to have a positive relationship with firm performance. A firm with more employees will likely have a larger market share, and therefore be more able to capitalize from the market, leading to a better firm performance (Smith et al., 2006). Following from the resource dependence theory, a larger

firm is expected to have a positive relationship with gender diversity in boards. With women's equal rights becoming more mainstream in society, a larger firm will be more exposed to this and will feel more pressure to have a gender-diverse board. By not including firm size, the coefficient of gender diversity will thus be positively biased. Firm size is frequently included in previous studies, but is also often measured as total sales or total assets next to the number of employees (Lückerath-Rovers, 2013; Adams & Ferreira, 2009; Bøhren & Strøm, 2010; Campbell & Mínguez-Vera, 2008; Carter, Simkins & Simpson, 2003; Carter et al., 2010; Lee & James, 2007; Liu et al., 2014; Marinova, Plantenga & Remery, 2016; Rose, 2007; Smith et al., 2006).

Second, firm age is measured as the natural logarithm of the number of years since the firm was incorporated. The effect of firm age on firm performance is argued to have an inverse U-shape (Smith et al., 2006). Younger firms are less experienced, need to build up their market position and market share and often have higher capital costs compared to older firms, which leads to a lower firm performance. On the other hand, old firms might become laxer and their product life cycle might be coming to an end which decreases earnings. Ujunwa (2012) suggests that complexity of capital structures increases with firm age. Also, the relationship between firm age and gender diversity is unclear. Older firms might have more traditional views and therefore the sex-role stereotypes might be more present, leading to lower gender diversity. On the other hand, for a younger firm gender diversity might be less of a concern as it is more focused on gaining its market share. Although the direction of the bias by not including firm age will cause is unclear, it is still important to include it in the regression. It is used repeatedly in previous studies and is usually defined as the number of years since the firm was incorporated (Hillman et al., 2007; Smith et al., 2006; Miller & del Carmen Triana, 2009; Marinova et al., 2016; Liu et al., 2014; Dezsö and Ross, 2012).

Third, board size is measured as the natural logarithm of the total number of board members and is expected to have a positive relationship with firm performance too. In line with the resource dependence theory, larger boards will have access to more resources in firm decision-making because of greater knowledge from more directors. Better decisions will consequently lead to better firm performance. Adams and Ferreira (2009) explain the relationship between board size and gender diversity is largely mechanical and has no interesting economic implications. For example, since the Act on Management and Supervision, Dutch firms might have simply added a board position to their current board and assigned it to a female. It is however important to control for board size to ensure the effects that are found are due to gender diversity and not board size (Adams & Ferreira, 2009). Carter et al. (2003), Hyland and Marcellino (2002) and Brammer, Millington and Pavelin (2007) all find a positive relationship between board size and gender diversity. By not including board size, the coefficient of gender diversity will thus be positively biased. It is therefore often used as a control variable in previous studies, most commonly measured as the total

number of board members (Hillman et al., 2007; Joecks et al, 2013; Liu et al., 2014; Lückerath-Rovers, 2013; Marinova et al., 2016; Mahadeo, Soobaroyen & Hanuman, 2012; Adams & Ferreira, 2009; Bøhren & Strøm, 2010; Bianco et al., 2015; Carter et al., 2010; Erhardt, Werbel, & Shrader, 2003; Haslam et al., 2010; Carter et al., 2003). It is however not possible to account for all time-varying variables that influence gender diversity and firm performance, which will be explained in section 6.4.

Interestingly, table 3.2 shows board size is the only control variable with a significant relationship with the firm performance measures. All control variables are significantly correlated with all the gender diversity measures, except for firm size and the critical mass dummy. The minima and maxima in table 3.3 of firm size and firm age might seem extreme. These will be discussed in section 3.5.

Table 3.1 Variable definitions

Variable	Definition	Formula
Firm performance		
Tobin's Q	The market value of firm divided by its assets' replacement cost	$\frac{Market\ Capitalization}{Total\ Assets}$
ROA	Operating ability of firm to generate operating profit using its assets	$\frac{Operating\ profit + Depreciation}{Total\ Assets}$
ROE	Financial ability of firm to generate income using the investments of shareholders	$\frac{Net\ Income}{Shareholder\ Funds}$
Gender diversity		
Equality index	The ratio of males and females on boards	$See\ Section$ 2.3
Female dummy	Equals 1 if at least one female on board, 0 otherwise	
Critical mass	Equals 1 if at least 30 percent females on board, 0 otherwise	
Small share	Equals 1 if between 0 and 30 percent females on board, 0 otherwise	
Control variable		
Firm size	The natural logarithm of the number of employees	$\log Employees$
Firm age	The natural logarithm of number of years the firm has existed	$\log (Current\ Year - Founding\ Year)$
Board size	The natural logarithm of total number of board members	$logNumber\ of\ Board\ Members$

Table 3.2 Correlation matrix

		1	2	3	4	5	6	7	8	9	10
1	**Tobin's Q**	1.00									
2	**ROA**	-0.44*	1.00								
3	**ROE**	-0.24*	0.68*	1.00							
4	**Equality index**	0.0013	0.059	0.17*	1.00						
5	**Female dummy**	-0.026	0.066	0.17*	0.84*	1.00					
6	**Critical mass**	-0.016	0.049	0.091*	0.68*	0.30*	1.00				
7	**Small share**	-0.015	0.049	0.11*	0.39*	0.80*	-0.34*	1.00			
8	**Firm size**	-0.013	0.021	0.057	0.18*	0.20*	0.019	0.19*	1.00		
9	**Firm age**	-0.0048	-0.027	0.092*	0.22*	0.23*	0.091*	0.17*	0.075	1.00	
10	**Board size**	-0.097*	0.11*	0.16*	0.37*	0.47*	0.13*	0.38*	0.38*	0.15*	1.00

* Significance at the 5% level.

Table 3.3 Summary statistics

Variable	**Obs.**	**Mean**	**Std.**	**Min.**	**Max.**
Firm performance					
Tobin's Q	587	.98	1.13	.014	13.96
ROA	523	0.036	0.88	-18.68	0.88
ROE	563	-0.0053	0.68	-8.33	4.34
Gender diversity					
Equality index	599	25.06	24.76	0	100
Female dummy	599	0.59	0.49	0	1
Critical mass	599	0.11	0.32	0	1
Small share	599	0.48	0.50	0	1
Control variable					
Firm size	533	25277.89	89031.6	1	709720
Firm age	599	60.31	55.87	0	335
Board size	599	7.26	2.83	1	16

3.3 Methodology

To empirically test the effect of gender diversity in boards on firm performance, a time and fixed effects regression methodology will be used. The large advantage of this type of regression is that it can account for time-invariant differences between individuals, or in this case firms. Therefore, firms from different sectors can be used for the same analysis, for example. Hereby an important assumption is made that the potential effect of gender diversity on firm performance is equal for all firms. This assumption will be further discussed in section 3.4. The following regression equation will be used:

$$Firm\ Performance_{i,t} = \alpha_i + \rho Gender\ Diversity_{i,t} + \sum_{j=1}^{N} \beta_j X_{j,i,t} + \gamma_t + \varepsilon_{i,t}$$

This equation describes the effect of the independent variable gender diversity in boards on the dependent variable firm performance. Firm performance is measured as Tobin's Q, ROA and ROE. Gender diversity is measured differently per hypothesis, as described in table 3.1. $\sum_{j=1}^{N} \beta_j X_{j,i,t}$ describes the effect of the control variables on firm performance. The control variables which will be included are firm size, firm age and board size. These control variables could influence the share of women in boards and the firm performance. α_i describes the firm fixed effects, γ_t describes the time fixed effects and $\varepsilon_{i,t}$ is the error term. In all regressions, robust standard errors will be included.

3.4 Sector effect

Using the methodology described in section 3.3, the assumption is made that ρ is equal for each firm. Naturally, there are time-invariant differences between firms that could influence the gender diversity in boards and firm performance. In this paper, a robustness test will be conducted to test the effect of the time-invariant factor of which sector the firm operates in, as this is included in many previous studies. Campbell and Mínguez-Vera (2008) argue firms match their gender diversity in boards to the gender diversity of potential customers and employees to increase their ability to penetrate the market. As the gender composition of customers and employees varies between sectors, gender diversity in boards is expected to vary systematically between sectors. Adams and Ferreira (2009) find a significant variation in the share of female directors across industries. The average development of firm performance might also be influenced by sector. Sectors differ in which stage of the sector life cycle they are in. Firms in a sector which is in the growth stage will likely differ in their average performance development compared to firms in a sector which is in the decline stage. Both Joecks et al. (2013) and Adams and Ferreira (2009) find significant differences between sectors of the effect of gender diversity on firm performance.

Orbis splits the firms into 19 major Bureau van Dijk sectors. Due to the small number of firms in the dataset, some of those sectors only consist of one or two firms. The sectors have therefore been subcategorized into 6 sectors defined as: 1. Communications: 2. Food, beverages, tobacco and retail; 3. Machinery, equipment, furniture and recycling; 4. Other services; 5. Banking, insurance and financial services; 6. Primary sector. Sector 5 was also added because Lückerath-Rovers (2013) found that Dutch listed firms that are in the financial sector are on average the largest firms but also have the highest share of female directors. Because sector is a time-invariant variable, it will be added as an interaction term between sector and the gender diversity measure. Figure 3.5 shows that the average equality index increased in all sectors between 2010 and 2018. As expected, the increases in the equality index vary per sector. For example, the average equality index in the food, beverages, tobacco and retail sector moved from just above 20 in 2010 to close to 50 percent in 2018. The communications sector average equality index increased from just above 25 percent in 2010 to just above 35 percent in 2018.

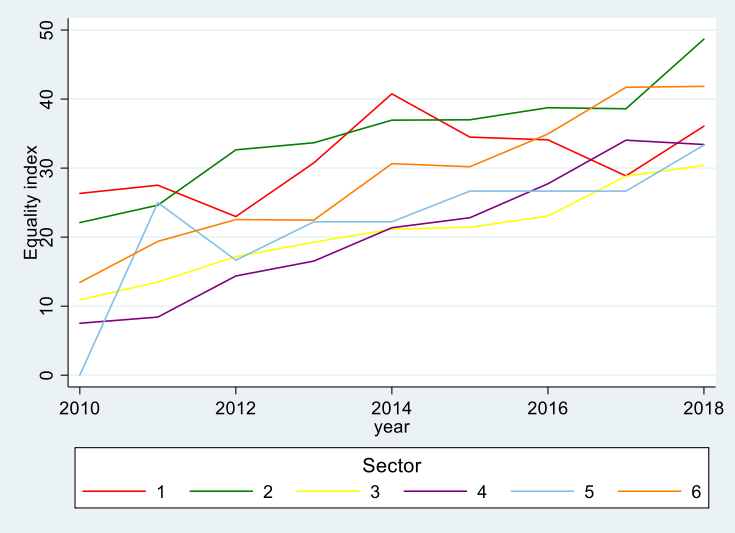

Figure 3.5 Development of mean of equality index per sector, 2010-2018

3.5 Reverse causality

Next to omitted variable bias, another concern for a fixed effects methodology is reverse causality. Reverse causality would exist if the dependent variable would (also) affect the independent variable. According to Smith et al. (2006), well-performing firms might employ more women, because they decide to follow a riskier strategy of recruiting board members. In addition, Dezsö and Ross (2012) argue women may be able to self-select themselves into firms because of the scarcity of experienced female directors. Well-performing firms might also be more likely to respond to pressure to conform to the desired norm of gender diversity, because the need for legitimacy is larger or because they have more resources to do so (Meyer &

Rowan, 1977; Pfeffer & Salancik, 2003). If this reasoning is true, positive causality runs from firm performance to gender diversity, and not the reverse.

In previous literature, this problem is tackled via three different methods. First, Smith et al. (2006) include an instrumental variable which has no connection to firm performance but is related to gender diversity in boards. How this works, is explained in section 2.2. Second, Dezsö and Ross (2012) conduct their main regression with one-year lagged values of the dependent variables as a control variable. Third, Joecks et al. (2013) and Liu et al. (2014) add one-year lagged variables of their gender diversity and board characteristics variables. In this paper reversed causality will be controlled for following the second and third method in the robustness section. The main regression will first be conducted with one-year lagged values of Tobin's as a control variable. Then, the main regression using Tobin's Q will be conducted with one-year lagged values of gender diversity and board size.

3.6 Outliers

When studying the data, a few noteworthy data points were found regarding the minima and maxima of the firm performance measures. These data points deviated largely from the other data points, making them outliers. It is likely these outliers exist because of measurement errors. For example, the maximum of Tobin's Q is 13.96, which would mean investors expected a firm to create more than 13 times the value of the assets it currently has available in the near future. This number seems very unrealistic. Outliers can have a large influence on the coefficient of the independent variable. Therefore, a robustness test will be conducted where the outliers are removed to see if the coefficients change significantly. If the data points do not fall between the following range, they will be removed for the robustness test:

$$[Q_1 - 1.5(Q_3 - Q_1), Q_3 + 1.5(Q_3 - Q_1)]$$

Q_1 is the median of the lower half of the data and Q_3 is the median of the upper half of the data.

Firm size and firm age also had noteworthy minima and maxima. However, these deviations can be explained. The minimum of 1 for firm size, meaning the firm had 1 employee, can be explained by the fact that holding firms are part of the dataset. A holding firm owns shares of other firms and can therefore be large and thus listed on the Amsterdam Euronext Stock Exchange. It does not necessarily produce goods or services itself, and consequently does not need many employees. This can also explain why the minimum of firm age is zero because some of these holding firms are created by existing large firms. The maximum of firm size is 709720 of the firm Randstad N.V, which seems a very large number for employees. It is however a realistic number, because this firm is an enormous employment agency that operates on an international scale.

4. Results

In this section, the results for the different regressions are presented and the outcomes of the hypotheses are discussed.

4.1 Hypothesis 1

Table 4.1 presents the results of the fixed effects model to test hypothesis 1. The coefficient of the equality index in regression 1, 2 and 3 is insignificant at the 10% level. This would mean an increase in the equality index would have no positive effect on firm performance. Regression 1 has the highest within R-squared value at 0.14. This means it explains the variation in firm performance within firms the most compared to regression 2 and 3. The only significant coefficient at the 10% level is of firm size in regression 3. Here a one percent increase in the number of employees will lead to an increase in ROE of 0.0015. This positive relationship is in line with the expected relationship explained in section 3.2.3. Hypothesis 1 states the following:

Hypothesis 1: Firm performance is higher when a board has a higher equality index.

From the provided analysis, hypothesis 1 cannot be accepted as the coefficient of the equality index in regression 1, 2 and 3 is not statistically different from zero. Therefore, the null hypothesis, which states that firm performance is equal or lower when a board has a higher equality index, cannot be rejected. Using the percentage of females in a board, previous studies did find a positive relationship with firm performance. Liu et al. (2014) find a significant positive relationship between the percentage of female directors and ROA while also using a firm fixed effects methodology. Lückerath-Rovers (2013) finds a significant positive relationship between the percentage of female directors and ROE while also studying Dutch listed firms, but with an OLS regression methodology.

4.2 Hypothesis 2

The results for the fixed effects model using the female dummy variable can be found in table 4.2. In regression 4, 5 and 6 the coefficients of the female dummy are all insignificant at the 10% level. This means that having at least one female in a board did not have a significant effect on firm performance, compared to having no female directors. The coefficients of the control variables do not differ a lot compared to regression 1, 2 and 3, as the same firms are studied. Once again, firm size is the only significant coefficient at the 10% level and has the same effect on firm performance as in regression 3. Hypothesis 2 is formulated as follows:

Hypothesis 2: Firm performance is higher when there is at least one female in a board.

From the provided analysis, hypothesis 2 cannot be accepted as the coefficient of the female dummy in regression 4, 5 and 6 is not statistically different from zero. Therefore, the null hypothesis, which states that firm performance is equal or lower when there is at least one female in a board, cannot be rejected. Campbell and Mínguez-Vera (2008) also find an insignificant effect of having at least one female in the board on Tobin's Q, using a fixed effects methodology. On the other hand, Lückerath-Rovers (2013) finds a significant positive relationship between the female dummy and ROE in Dutch listed firms, using an OLS regression.

Table 4.1 Effect of equality index on firm performance

	Tobin's Q	**ROA**	**ROE**
Variable	**(1)**	**(2)**	**(3)**
Equality index	-0.0010	0.0018	0.0034
	(0.0020)	(0.0021)	(0.0029)
Firm size	-0.10	-0.10	0.15*
	(0.086)	(0.12)	(0.079)
Firm age	0.25	0.047	-0.32
	(0.38)	(0.21)	(0.33)
Board size	-0.80	1.43	-0.11
	(0.66)	(1.28)	(0.23)
Obs	526	494	516
Within R^2	0.14	0.085	0.046

* Significance at the 10% level, ** Significance at the 5% level.

Table 4.2 Effect of female dummy on firm performance

	Tobin's Q	**ROA**	**ROE**
Variable	**(4)**	**(5)**	**(6)**
Female dummy	0.027	-0.0025	0.137
	(0.086)	(0.054)	(0.11)
Firm size	-0.010	-0.10	0.15*
	(0.085)	(0.11)	(0.081)
Firm age	0.24	0.039	-0.41
	(0.39)	(0.20)	(0.39)
Board size	-0.80	1.42	-0.036
	(0.64)	(1.27)	(0.21)
Obs	526	494	516
Within R^2	0.14	0.084	0.044

* Significance at the 10% level, ** Significance at the 5% level.

4.3 Hypothesis 3

Table 4.3 presents the results of the model to test the critical mass theory. The coefficients for the critical mass dummy in regression 7, 8 and 9 are all insignificant at the 10% level. The same goes for the coefficients of the small share dummy. This would mean having more than 30 percent females in the board would not have a positive significant effect on firm performance compared to firms with 0 to 30 percent female directors and firms with no female directors. Once again, the coefficients of the control variables do not differ largely from regression 1, 2 and 3, as the same firms are studied over the same period in both models. Hypothesis 3 states the following:

Hypothesis 3: Firm performance is higher when there is 30 percent or more females in a board.

From the provided analysis, hypothesis 3 cannot be accepted as the coefficient of the critical mass dummy in regression 7, 8 and 9 is not statistically different from zero. Therefore, the null hypothesis, which states that firm performance is equal or lower when there is 30 percent or more females in a board, cannot be rejected. Joecks et al. (2013) do find a positive significant relationship between firms that reach a critical mass of 30 percent and ROE while using an OLS and random effects methodology.

Table 4.3 Effect of critical mass on firm performance

	Tobin's Q	**ROA**	**ROE**
Variable	**(7)**	**(8)**	**(9)**
Critical mass	-0.084	0.010	0.22
	(0.13)	(0.12)	(0.20)
Small share	0.029	-0.0052	0.14
	(0.085)	(0.055)	(0.11)
Firm size	-0.010	-0.097	0.15*
	(0.086)	(0.11)	(0.080)
Firm age	0.21	0.066	-0.39
	(0.38)	(0.22)	(0.37)
Board size	-0.82	1.44	-0.021
	(0.65)	(1.29)	(0.23)
Obs	526	494	516
Within R^2	0.14	0.085	0.045

* Significance at the 10% level, ** Significance at the 5% level.

5. Robustness

In this section four different robustness tests will be performed and analyzed. The first test is for an understanding of reverse causality in the data. The second test will analyze the data if outliers are removed and the third test will consider the effect of which sector the firm is in on gender diversity in a board. Tobin's Q is the only firm performance measure that will be utilized in the first and second robustness test, as ROA and ROE were used in the results section as robustness tests for firm performance. In addition, the regression models using Tobin's Q in the results section had the highest within R-squared values, meaning these models explained the variation in firm performance within firms the most compared to ROA and ROE. In the third robustness test ROA and ROE will be included because those variables had suspicious outliers.

5.1 Sector robustness test

Table 5.1 shows the results for the interaction terms between sector and the equality index. In the main regression, the assumption is made that ρ is equal for each firm. Sector is included to test whether the effect of gender diversity on firm performance differs per sector. The interaction term between the equality index and sector 5 is omitted because there was insufficient variation of the equality index for the firms in sector 5 between 2010 and 2018. The coefficient of the interaction term between the equality index and sector 4 is the only significant coefficient at the 5% level. The coefficients of the other interaction terms are insignificant, meaning their effect of gender diversity on firm performance do not statistically differ from sector 1. Sector 4 is the other services sector. This means the effect of gender diversity on firm performance differs in the other services sector compared to the other sectors. In other words, the ρ value for sector 4 is 0.0084 lower compared to the other sectors. The other service sector consists of business, construction, transport, freight & storage, property, travel, personal & leisure and printing & publishing services. Because of this broad definition, it is difficult to conclude which type(s) of services cause(s) the differing effect. The significant result does however show it is important to control for sector.

Table 5.1 Effect of equality index on Tobin's Q with sector interaction terms

	Tobin's Q
Variable	**(10)**
Equality index	0.0035
	(0.0031)
Equality index*sector 2	-0.00044
	(0.0048)
Equality index*sector 3	-0.0031
	(0.0051)
Equality index*sector 4	-0.0084**
	(0.0035)
Equality index*sector 5	Omitted
Equality index*sector 6	-0.0038
	(0.0037)
Firm size	-0.098
	(0.087)
Firm age	0.30
	(0.39)
Board size	-0.80
	(0.65)
Obs	526
Within R^2	0.15

* Significance at the 10% level, ** Significance at the 5% level.

5.2 Reverse causality robustness tests

To control for reverse causality, the main regression will first be conducted with one-year lagged values of Tobin's Q as a control variable. Then, the main regression using Tobin's Q will be conducted with one-year lagged values of gender diversity and board size. Table 5.2 shows the regression with one-year lagged values of Tobin's Q as a control variable. The coefficients of the one-year lagged Tobin's Q variable and the equality index are insignificant in regression 11, 12 and 13 at the 10% level. There would be a case of reverse causality if the hypothesized positive relationship between gender diversity and firm performance disappears when prior firm performance is added to the regression. Because this positive relationship was not found for any of the gender diversity measures in regressions 1, 4 and 7 in the results section, it cannot be concluded whether or not reverse causality is present.

Table 5.2 Effect of gender diversity measures on Tobin's Q with one-year lagged Tobin's Q variable

	Tobin's Q		
Variable	**(11)**	**(12)**	**(13)**
Tobin's $Q_{(t-1)}$	0.0054	0.0039	0.0039
	(0.052)	(0.050)	(0.050)
Equality index	-0.00067		
	(0.0024)		
Female dummy		0.069	
		(0.12)	
Critical mass			-0.065
			(0.050)
Small share			0.071
			(0.11)
Firm size	-0.051	-0.051	-0.048
	(0.092)	(0.091)	(0.091)
Firm age	0.21	0.18	0.13
	(0.34)	(0.36)	(0.34)
Board size	-1.09	-1.10	-1.12
	(0.74)	(0.75)	(0.75)
Obs	466	466	466
Within R^2	0.16	0.16	0.17

* Significance at the 10% level, ** Significance at the 5% level.

Table 5.3 presents the results of the regressions with one-year lagged values of gender diversity and board size. By using these lagged values, reverse causality can be avoided as current firm performance should not be able to influence board characteristics of the past. The coefficients of all gender diversity measures are insignificant at the 10% level. A positive significant effect of previous board characteristics on firm performance is thus not found.

Table 5.3 Effect of one-year lagged gender diversity measures on Tobin's Q

	Tobin's Q		
Variable	**(14)**	**(15)**	**(16)**
Equality index$_{(t-1)}$	0.00031		
	(0.0024)		
Female dummy$_{(t-1)}$		-0.011	
		(0.096)	
Critical mass$_{(t-1)}$			0.053
			(0.17)
Small share$_{(t-1)}$			-0.011
			(0.096)
Firm size	-0.10	-0.10	-0.10
	(0.089)	(0.089)	(0.089)
Firm age	0.091	0.095	0.11
	(0.34)	(0.33)	(0.34)
Board size$_{(t-1)}$	0.21	0.21	0.23
	(0.21)	(0.22)	(0.21)
Obs	471	471	471
Within R^2	0.092	0.092	0.093

* Significance at the 10% level, ** Significance at the 5% level.

5.3 Outliers robustness test

Following the method described in section 3.6, outliers were removed from the data to perform a robustness test. For Tobin's Q 28 observations were dropped. For ROA 61 observations and for ROE 76 observations were removed. The coefficients of the gender diversity measures remain insignificant, as in regressions 1 to 9. The only significant change of removing the outliers is visible in regressions 17, 18 and 19 in table 5.4 using Tobin's Q. Here, firm size has a significant negative effect on Tobin's Q in all three regressions. For the regressions using Tobin's Q we also find a higher R-squared value after removing outliers, meaning these regressions explained the variation in firm performance within firms more compared to before removing outliers. However, the regressions using ROA and ROE have a lower R-squared value after removing outliers. Removing outliers of firm performance thus does not have a significant effect on the effect of gender diversity on firm performance.

Table 5.4 Effect of gender diversity measures on firm performance with no outliers

	Tobin's Q			**ROA**			**ROE**		
Variable	**(17)**	**(18)**	**(19)**	**(20)**	**(21)**	**(22)**	**(23)**	**(24)**	**(25)**
Equality index	0.0011			0.00027			0.00034		
	(0.0012)			(0.00025)			(0.00038)		
Female dummy		0.056			0.019			0.0030	
		(0.070)			(0.014)			(0.024)	
Critical mass			0.040			0.023			0.018
			(0.090)			(0.016)			(0.028)
Small share			0.056			0.019			0.0031
			(0.070)			(0.014)			(0.024)
Firm size	-0.059*	-0.057*	-0.057*	0.0050	0.0051	0.0050	0.0093	0.0096	0.0094
	(0.035)	(0.034)	(0.034)	(0.0046)	(0.0045)	(0.0045)	(0.013)	(0.013)	(0.013)
Firm age	0.090	0.052	0.047	-0.0072	-0.019	-0.018	-0.037	-0.039	-0.034
	(0.27)	(0.26)	(0.26)	(0.025)	(0.027)	(0.027)	(0.090)	(0.090)	(0.092)
Board size	0.080	0.070	0.067	-0.0084	-0.012	-0.011	-0.010	-0.012	-0.010
	(0.084)	(0.085)	(0.085)	(0.025)	(0.026)	(0.026)	(0.029)	(0.028)	(0.028)
Obs	498	498	498	465	465	465	440	440	440
Within R^2	0.24	0.24	0.24	0.020	0.028	0.028	0.039	0.036	0.038

* Significance at the 10% level, ** Significance at the 5% level.

6. Discussion

In this section, the main outcomes and weaknesses of the analyses are discussed. The sample, time period of the data, external validity and issues with causality are critically analyzed.

6.1 Sample

In this paper 100 Dutch firms listed on the Amsterdam Euronext Stock Exchange were used. However, the database Orbis did not include data for all 100 firms. Based on whether a firm had a Tobin's Q value for six or more years out of nine, sufficient data was available for just 67 firms. It is not clear why the data of some firms was available, while the data of other firms was not. The data before 2010 had significantly more missing values and many of the firms had not released their results for years after 2018 on the first of May 2020. A higher number of firms would lead to more observations. The significance of the coefficients in the regressions are based on the t-test. By adding more observations, for the same t-value, the P-value will be lower. A coefficient is considered significant, for example, at the 10% level if the P-value is below 0.1. By increasing the number of observations, it is thus possible to increase the significance of the coefficients of variables. In addition, for the robustness test for the effect of which sector a firm is in, the firms were subcategorized from 19 major sectors into six different sectors due to the fact that otherwise some of the sectors would only include one or two firms. By including more firms, it would be possible to divide the firms into more categories and therefore create a clearer picture of the effects per specific sector.

6.2 Time

The studied period was 2010-2018. Although Dutch firms have had an increasingly higher share of females in their boards in this period, 9 years is not very long (Dirven & Merens, 2018). As the Act on Management and Supervision is a part of Dutch law since 2013, it is likely gender diversity will increase in the coming years (Leenders et al., 2019). As for hypothesis 3, in just 71 firm years the critical mass was reached. Coming back to this hypothesis in a few years will make it more relevant in the Netherlands as the number of firms that have reached the critical mass is likely to have increased. The number of observations of firms that have reached the critical mass will therefore increase, which can lead to more significant coefficients of variables as explained in section 6.1.

6.3 External validity

It is important to question whether the sample of Dutch listed firms is representative for all other firms. A method to increase the external validity would be to include listed firms from other countries, European countries for example. If firms from other countries are added, it would be important to account differences

between countries. Firms from differing countries have to oblige to other tax law systems which can affect after-tax performance measures (Lückerath-Rovers, 2013). Many European countries also have binding gender quotas for boards of directors, such as Belgium, Finland, France, Germany, Iceland, Italy and Norway (Smith, 2018). A binding quota might also affect firm performance, as firms will have to invest more time and money into making sure they reach the quota. If a firm does not reach the quota, it will also have to pay a fine. Another strategy to increase the external validity of would be to send a survey to firms. Although firm size is accounted for, this dataset consists of only listed firms which are typically large. By collecting data via a survey, it would be possible to include firms which are not listed and public. By doing so, private firms but also smaller firms could be included which increases the external validity of the results.

6.4 Causality

The fixed effects methodology comes with two major challenges. First, the issue of omitted variable bias. Omitted variable bias is present when there are time-varying variables that are not included in the regression analysis which influence both gender diversity and firm performance. If these variables are omitted, the coefficient of the dependent variable can be biased, because it includes the effect of the omitted variable into its own coefficient. The bias of a coefficient is the difference between the expected value and the true value of the coefficient. Control variables are helpful to decrease the potential bias in the coefficient of the independent variable, which was explained in section 3.2.3. In the main regression, firm size, firm age and board size are included as control variables. For example, firm size was expected to have a positive relationship with firm performance and gender diversity. By not including firm size, the coefficient of the gender diversity measure will be positively biased. The hypothesized positive effect of firm size would be included in the coefficient of the gender diversity measure, which would thus lead to the positive bias.

The challenge of omitted variable bias is to identify all relevant omitted variables. For example, if a variable is included which has a positive significant effect on firm performance and is positively affected by gender diversity, which is defined as a mechanism, it can remove the causal effect of gender diversity. The positive effect of gender diversity will then be captured by the coefficient of the mechanism. The coefficient of the gender diversity measure will then be negatively biased.

Next to including the wrong type of control variables, not including relevant control variables also needs to be avoided. The main regression includes just three control variables. There might however be more variables which influence gender diversity and firm performance. Lui et al. (2014) and Carter et al. (2010) include the share of outside directors as a control variable. As argued through the agency theory in section 2.1, the interests of outside director are more in line with the firm than inside directors which can positively influence firm performance. Adams and Ferreira (2009) and Farrell and Hersch (2005) find that

female directors overwhelmingly serve on boards as outside directors. A higher share of outside directors would then positively influence firm performance and gender diversity. Not including the share of outside directors would thus lead to a positive bias. Smith et al. (2006) include export orientation as a control variable because firms that are active in the export market potentially operate in larger markets, which is expected to affect firm performance positively.

The second large challenge of the fixed effects methodology is reverse causality. Reverse causality would exist if the dependent variable would (also) affect the independent variable. There are multiple arguments for positive causality running from firm performance to gender diversity, and not the reverse, which were explained in section 3.4. If there is a case of reverse causality which is not accounted for, the coefficient of the gender diversity measure will be positively biased. The coefficient of the gender diversity measure will portray a positive effect on firm performance, while in reality firm performance would be positively influencing gender diversity. Three different methods to tackle this problem were identified in previous literature. First, Smith et al. (2006) include an instrumental variable which has no connection to firm performance but is related to gender diversity in boards. How an instrumental variable works, is explained in section 2.2. Second, Dezsö and Ross (2012) conduct their main regression with one-year lagged values of the dependent variables. Third, Joecks et al. (2013) and Liu et al. (2014) add one-year lagged variables of their gender diversity and board characteristics variables. In this paper reversed causality was tested following the second and third method in section 5.2. No evidence was found for prior performance affecting the relationship between gender diversity and firm performance, but this is due to the fact a relationship between gender diversity and firm performance was not found beforehand.

7. Conclusion

The Act on Management and Supervision was voted for by the Dutch parliament to encourage firms to move towards more equal opportunity within the Netherlands. Equal opportunity may not provide a sufficient business driver for large profit-minded firms. Based on the token status, agency, resource dependence and critical mass theories, this paper is set up to investigate the effect of gender diversity in boards on firm performance. These theories were used to formulate the main research question. The research question of this paper is as follows:

What is the effect of gender diversity in boards on firm performance?

To answer this question, three hypotheses using different measures of gender diversity were formulated. Hypothesis 1 states firm performance is higher when a board has a higher equality index. Hypothesis 2 states firm performance is higher when there is at least one female in a board. Hypothesis 3 states firm

performance is higher when there is 30 percent or more females in a board. To test these hypotheses, a firm and time fixed effects model was used. The results of the models show no statistically significant effect of gender diversity on firm performance for all three measures of gender diversity. The three hypotheses cannot be accepted. Therefore, it cannot be concluded that gender diversity has a positive effect on firm performance. Four robustness tests are performed to test for a sector effect, reverse causality and outliers. A significant effect of the other services sector is found, showing that the sector a firm is in can influence the effect of gender diversity on firm performance. A larger dataset would be useful to find out more specifically in which sector(s) firm performance is affected differently by gender diversity. For reverse causality, no significant effect is found of the one-year lagged variables of Tobin's Q and the board characteristics. Removing outliers also does not yield a significant effect of gender diversity on firm performance. Our results do not provide an incentive for firms to strive for a gender-diverse board based on their potential firm performance. Therefore, if the Dutch parliament wants to move towards more equal opportunity within the Netherlands and specifically within board rooms of large firms, a binding quota might be necessary.

For future research, the research question will stay relevant. This topic is of great importance to firms, as in the end it can be the key to performing well or poorly. Besides, more gender diversity increases gender equality, which is something countries or even the European Union should be working towards. It would be interesting to study firms in different countries, but also firms of different sizes. For smaller startups it would also be beneficial to have the knowledge whether it is important to have a gender-diverse board. In line with this, researching whether the timing of achieving gender diversity is also of great interest. Should firms strive for gender diversity from the start, or does it have a greater impact on firm performance once the firm is up and running?

Gender diversity is not the only measurable form of diversity. Researching the effect of age, ethnicity and board experience of directors on firm performance will add an understanding of the potential importance of diverse boards. In addition, in many countries in Europe, large firms are now expected to have a minimum share of females in their boards (Smith, 2018). In some countries, this is even required by law. It would be valuable to research if the effect of gender diversity is different when it is introduced artificially, compared to the situation where firms make their own choices. If all firms suddenly have to find many female directors, it could be the case that many female directors are appointed with little prior experience. Boards with less experience might have a lower board performance, which can lead to a lower firm performance. An artificial introduction of required gender diversity could thus have a negative effect on firm performance. Finally, boards are not the only place where greater gender diversity can be achieved.

A further recommendation is to study gender diversity within the whole firm, including all types of employees.

Bibliography

Adams, R. B., De Haan, J., Terjesen, S., & Van Ees, H. (2015). Board diversity: Moving the field forward. *Corporate Governance: An International Review*, *23*(2), 77-82.

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of financial economics*, *94*(2), 291-309.

Bianco, M., Ciavarella, A., & Signoretti, R. (2015). Women on corporate boards in Italy: The role of family connections. *Corporate Governance: An International Review*, *23*(2), 129-144.

Bøhren, Ø., & Strøm, R. Ø. (2010). Governance and politics: Regulating independence and diversity in the board room. *Journal of Business Finance & Accounting*, *37*(9-10), 1281-1308.

Brammer, S., Millington, A., & Pavelin, S. (2007). Gender and ethnic diversity among UK corporate boards. *Corporate Governance: An International Review*, *15*(2), 393-403.

Bureau van Dijk. (2020). *Worldwide Company Data.* Retrieved from https://www.bvdinfo.com/en-gb/our-products/data/international/orbis?gclid=CjwKCAjw2a32BRBXEiwAUcugiKVwQUrmn5QnbUYVIwFHQmw4cVX-I-COzxN_2FdBS-78bxe-_IDGJxoC46UQAvD_BwE

Campbell, K., & Mínguez-Vera, A. (2008). Gender diversity in the boardroom and firm financial performance. *Journal of business ethics*, *83*(3), 435-451.

Carter, D. A., D'Souza, F., Simkins, B. J., & Simpson, W. G. (2010). The gender and ethnic diversity of US boards and board committees and firm financial performance. *Corporate Governance: An International Review*, *18*(5), 396-414.

Carter, D. A., Simkins, B. J., & Simpson, W. G. (2003). Corporate governance, board diversity, and firm value. *Financial review*, *38*(1), 33-53.

Demsetz, H., & Villalonga, B. (2001). Ownership structure and corporate performance. *Journal of corporate finance*, *7*(3), 209-233.

Desvaux, G., Devillard-Hoellinger, S., & Baumgarten, P. (2007). *Women matter: Gender diversity, a corporate performance driver*. McKinsey.

Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic management journal*, *33*(9), 1072-1089.

Dirven, H.-J. & Merens, A. (2018). *Emancipatiemonitor 2018.* Retrieved from https://digitaal.scp.nl/emancipatiemonitor2018/komen-er-meer-vrouwen-in-topfuncties/

Dobbin, F., & Jung, J. (2010). Corporate board gender diversity and stock performance: The competence gap or institutional investor bias. *NCL Rev.*, *89*, 809.

Ebregt, J., Jongen, E., Lanser, D. & Merens, A. (2019). Vrouwen aan de top. Retrieved from https://www.cpb.nl/sites/default/files/omnidownload/cpb-notitie-vrouwen-aan-de-top.pdf

Erhardt, N. L., Werbel, J. D., & Shrader, C. B. (2003). Board of director diversity and firm financial performance. *Corporate governance: An international review*, *11*(2), 102-111.

Fama, E. F., & Jensen, M. C. (1983). Separation of ownership and control. *The journal of law and Economics*, *26*(2), 301-325.

Farrell, K. A., & Hersch, P. L. (2005). Additions to corporate boards: The effect of gender. *Journal of Corporate finance*, *11*(1-2), 85-106.

Fields, M. A., & Keys, P. Y. (2003). The emergence of corporate governance from Wall St. to Main St.: Outside directors, board diversity, earnings management, and managerial incentives to bear risk. *Financial Review*, *38*(1), 1-24.

Finkelstein, S., Hambrick, D., & Cannella, A. A. (1996). Strategic leadership. *St. Paul: West Educational Publishing*.

García Álvarez, A. (2016). *Firm Performance and Gender Quotas in Supervisory Boards.* Retrieved from http://arno.uvt.nl/show.cgi?fid=142286

Haslam, S. A., Ryan, M. K., Kulich, C., Trojanowski, G., & Atkins, C. (2010). Investing with prejudice: The relationship between women's presence on company boards and objective and subjective measures of company performance. *British Journal of Management*, *21*(2), 484-497.

Hillman, A. J., Cannella Jr., A. A., & Harris, I. C. (2002). Women and racial minorities in the boardroom: How do directors differ?. *Journal of management*, *28*(6), 747-763.

Hillman, A. J., Shropshire, C., & Cannella Jr., A. A. (2007). Organizational predictors of women on corporate boards. *Academy of Management Journal*, *50*(4), 941-952.

Huse, M., & Solberg, A. G. (2006). Gender-related boardroom dynamics: How Scandinavian women make and can make contributions on corporate boards. *Women in management review*, *21*(2), 113-130.

Hyland, M. M., & Marcellino, P. A. (2002). Examining gender on corporate boards: A regional study. *Corporate Governance: The international journal of business in society*.

Joecks, J., Pull, K., & Vetter, K. (2013). Gender diversity in the boardroom and firm performance: What exactly constitutes a "critical mass?". *Journal of business ethics*, *118*(1), 61-72.

Joy, L., Wagner, H. M. & Narayanan, S. (2007). *The Bottom Line: Corporate Performance and Women's representation on boards*. Catalyst.

Kanter, R. M. (2008). *Men and women of the corporation: New edition*. Basic books.

Kramer, V. W., Konrad, A. M., Erkut, S., & Hooper, M. J. (2006). *Critical mass on corporate boards: Why three or more women enhance governance* (pp. 2-4). Wellesley, MA: Wellesley Centers for Women.

Kravitz, D. A. (2003). More women in the workplace: is there a payoff in firm performance?. *Academy of Management Perspectives*, *17*(3), 148-149.

Lee, P. M., & James, E. H. (2007). She'-e-os: gender effects and investor reactions to the announcements of top executive appointments. *Strategic Management Journal*, *28*(3), 227-241.

Leenders, J., Pouwels, B., Van den Brink, M. (2019). *Bedrijvenmonitor Topvrouwen 2019.* Retrieved from https://www.rijksoverheid.nl/documenten/rapporten/2019/09/20/bedrijvenmonitor-topvrouwen-2019-samenvatting

Lindenberg, E. B., & Ross, S. A. (1981). Tobin's q ratio and industrial organization. *Journal of business*, 1-32.

Liu, Y., Wei, Z., & Xie, F. (2014). Do women directors improve firm performance in China?. *Journal of corporate finance*, *28*, 169-184.

Lückerath-Rovers, M. (2013). Women on boards and firm performance. *Journal of Management & Governance*, *17*(2), 491-509.

Mahadeo, J. D., Soobaroyen, T., & Hanuman, V. O. (2012). Board composition and financial performance: Uncovering the effects of diversity in an emerging economy. *Journal of business ethics*, *105*(3), 375-388.

Marinova, J., Plantenga, J., & Remery, C. (2016). Gender diversity and firm performance: Evidence from Dutch and Danish boardrooms. *The International Journal of Human Resource Management*, *27*(15), 1777-1790.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American journal of sociology*, *83*(2), 340-363.

Miller, T., & del Carmen Triana, M. (2009). Demographic diversity in the boardroom: Mediators of the board diversity–firm performance relationship. *Journal of Management studies*, *46*(5), 755-786.

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, *56*(3), 799-866.

Pfeffer, J., & Salancik, G. R. (2003). *The external control of organizations: A resource dependence perspective*. Stanford University Press.

Powell, G. N., & Butterfield, D. A. (2002). EXPLORING THE INFLUENCE OF DECISION MAKERS' RACE AND GENDER ON ACTUAL PROMOTIONS TO TOP MANAGEMENT. *Personnel psychology*, *55*(2), 397-428.

Rose, C. (2007). Does female board representation influence firm performance? The Danish evidence. *Corporate Governance: An International Review*, *15*(2), 404-413.

Ryan, M. K., & Haslam, S. A. (2005). The glass cliff: Evidence that women are over-represented in precarious leadership positions. *British Journal of management*, *16*(2), 81-90.

Shrader, C. B., Blackburn, V. B., & Iles, P. (1997). Women in management and firm financial performance: An exploratory study. *Journal of managerial issues*, 355-372.

Smith, N. (2018). Gender quotas on boards of directors. *IZA World of Labor*.

Smith, N., Smith, V., & Verner, M. (2006). Do women in top management affect firm performance? A panel study of 2,500 Danish firms. *International Journal of productivity and Performance management*.

Terjesen, S., Sealy, R., & Singh, V. (2009). Women directors on corporate boards: A review and research agenda. *Corporate governance: an international review*, *17*(3), 320-337.

Ujunwa, A. (2012). Board characteristics and the financial performance of Nigerian quoted firms. *Corporate Governance: The international journal of business in society*.

Wernerfelt, B., & Montgomery, C. A. (1988). Tobin's q and the importance of focus in firm performance. *The American Economic Review*, 246-250.

Appendix A: Scatterplots of ROA & ROE and the equality index

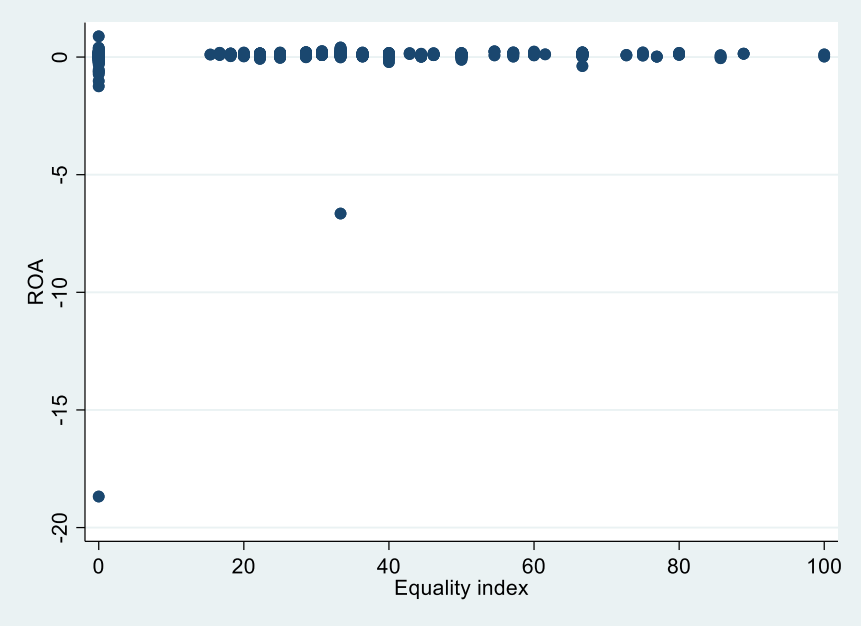

Figure 3.6 Scatterplot of ROA and equality index

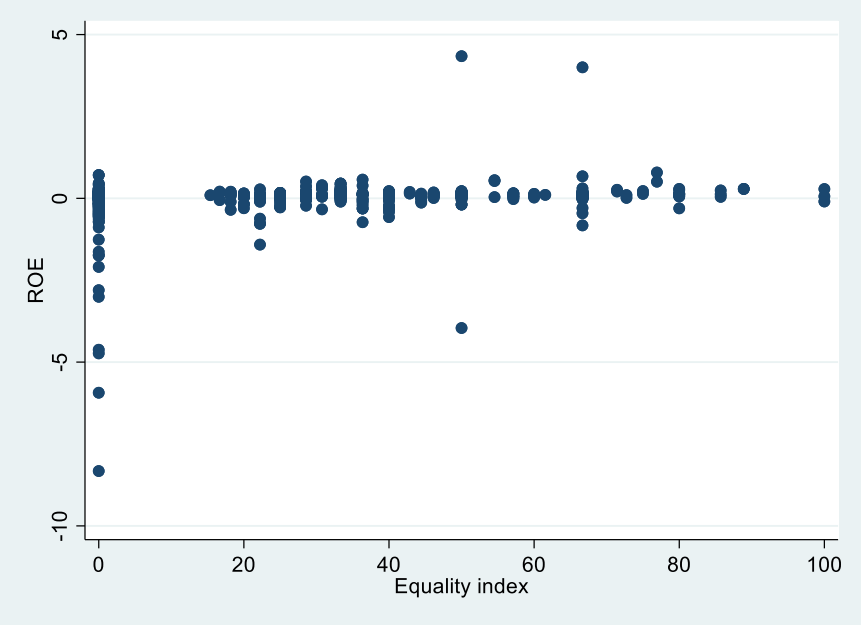

Figure 3.7 Scatterplot of ROE and equality index