University of Zurich UZH

Department of Banking and Finance

Centre of Competence for Sustainable Finance

Gender Diversity on the Board of Directors and ESG Performance in Banking – An Analysis of European Banks

Bachelor Thesis in Business and Finance

Naomi Gerber

Full Text Version

CCSF Thesis Series no.36 (2022)

August 2022

CCSF Thesis Series (formerly CSP/CMF Thesis Series)

Gender Diversity on the Board of Directors and ESG Performance in Banking – An Analysis of European Banks
Bachelor Thesis in Banking and Finance

Naomi Gerber
Advisor: Dr. Annette Krauss
Professor: Professor Dr. Marc Chesney

Full Text Version
CCSF Thesis Series no. 36 (2022)

Zurich: University of Zurich, Department for Banking and Finance / Center of Competence for Sustainable Finance, Plattenstrasse 14, 8032 Zurich, Switzerland

Abstract

This thesis examines the effect of board gender diversity on European banks' ESG performances while considering endogeneity. The findings of the GMM on a significant positive effect highlight the importance of increasing the share of female directors in creating a sustainable banking sector. However, the positive connection is neither due to stereotypical characteristics of women nor a significant gap in values between genders of directors. Rather differences in skills and backgrounds caused by the double glass ceiling are a valid explanation. We conclude that banks should focus on overall board diversity, not just on the gender of their board members.

Executive Summary

Researchers are going as far as blaming the male dominance in the financial sector for the financial crisis back in 2008. After all, not only has the male dominance in finance not changed in the last century, but the next big crisis caused by a collapse of our ecosystem might not be as far away as one wishes. If the financial sector does not act now, we will be locked on a fatal collision course. Among the various approaches to initiate this change, reinforcing gender diversity to shift the mindset of the boards toward a more sustainable corporate strategy might be one of the most promising concepts since it focuses on corporate governance as well as internal processes.

This thesis aims to shed empirical light on the connection between gender diversity and environmental performance by analyzing panel data of 93 European banks over a 10-year period. Moreover, we intend to raise awareness for the importance of considering endogeneity in the analysis by providing an extensive overview of the most frequent sources for the bias and comparing different models when dealing with endogeneity. By means of the General Method of Moments, this paper attempts to control for endogeneity, a common source of discrepant research results.

Based on a literature review on gender differences, we formulate five hypotheses that cover the following questions:

1. What is the general impact of a higher proportion of female directors on the ESG performance?
2. Is Kanter's critical mass theory applicable to the banking sector? The theory states that female directors can only exert an influence if the share of women exceeds a critical mass of 30%.
3. Does the gender of a bank's CEO affect the shape of the effect of gender diversity?
4. What is the state of gender homophily in the cases we consider? Based on this question, the critical mass theory is connected to the gender of the CEO and it is investigated whether this effect changes after a critical mass of women is reached.
5. Does the relationship of the variables change based on the country where the bank is located due to differences in existing gender quotas?

Four modified regression models are developed to investigate the hypotheses by means of the difference and system General Method of Moments. Nonetheless, for this thesis, the usage of the difference General Method of Moments is more efficient. The system General Method of Moments produces alarmingly high Sargan/Hansen test statistics, which indicates that the strict assumptions used for this method might not hold. Nevertheless, all methods show a positive influence of an increased percentage of female directors on the environmental performance.

Existing literature states that a potential connection of the variables is due to differences in characteristics and values between genders. Women are more likely to pursue long-term strategies, are more innovative, less committed to the status quo, and more likely to invest sustainably. Yet, this paper displays that the findings of the general population of women cannot be applied to the group of female directors due to the existing double glass ceiling in the banking sector. Women must often demonstrate supplementary characteristics to shatter the imaginary barrier and reach the board of directors. Since most boards still consist predominantly of male directors, the additional attributes are likely characteristics that society connects to men. Consequently, we assume that the gap in values between genders is smaller for directors. If we further connect this to our findings of a significant positive influence of the variable skills and background, we conclude that it is essential for an increased ESG score that the additional director creates new value. Anyhow, this is not automatically correlated to one's gender. Although this thesis finds a significantly positive effect of gender diversity on environmental performance, we should also consider different approaches.

The results do not support Kanter's critical mass theory, as the general effect of gender diversity on environmental performance represents an inverted U-shape. A positive effect of an increased share of women on the boards is visible, even before a critical mass of women is reached. Consequently, the results claim that gender balanced boards are most efficient for enhancing the environmental performance.

When considering the gender of a bank's CEO, the findings of this thesis strengthen that the previous observation of an inverted U-shaped effect is mainly valid for men-led banks. In contrast, for women CEOs, the effect changes concavity and has the form of a U-shape.

Therefore, the critical mass theory is supported for women-led European banks due to the possible existence of gender homophily. This suggests that one should not only help women break the first glass ceiling to reach the boards but, most importantly, shatter the second glass ceiling to increase the number of female CEOs in the banking sector.

Ultimately, this thesis attempted to investigate the connection between the geographical location of the banks and differences in the effect of gender diversity due to existing gender quotas. However, due to the following two limitations, it was not yet possible to include a dummy variable for gender quotas. First, most countries just recently established gender quotas; accordingly, it might be too early to observe the effect of this regulation. Second, the data available per country is relatively little in the compiled dataset. Hence, this paper simply compares the mean ESG scores and the mean shares of female directors between countries and notices widely spread records. By connecting this to the assumed differences between quota women and female directors in literature, we state that static gender quotas, which are mandatory for all banks, might not be efficient. We alternatively propose to target a specific ESG score for all the banks and an overall percentage of female directors for the sector rather than implementing static gender quotas whose impact is unclear for individual banks.

In summary, our finding of a positive association between a higher proportion of female directors and a bank's environmental performance sheds empirical light on the importance of gender diversified boards. Our results are particularly relevant since only few previous papers have examined this effect while simultaneously controlling for endogeneity. In contrast to extensive literature, which argues that the positive effect is due to stereotypical characteristics of women, we state that differences in skills and backgrounds caused by the double glass ceiling present a more valid explanation. Therefore, banks should focus on diversity as a whole, not just on the gender of their board members.

Table of contents

List of figures

List of tables

List of Abbreviations

CSR	Corporate Social Responsibility
ESG	Environmental, Social and Governmental
GICS	Global Industry Classification Standard
GMM	General Method of Moments
OLS	Ordinary Least Squares
ROA	Return on Assets
ROE	Return on Equity

1 Introduction

Researchers are going as far as blaming the male dominance in the financial sector for the financial crisis back in 2008. Excessive risk taking is associated with male dominated boards by several studies and already over 10 years ago, researchers suggested that gender balanced boards could potentially help overcome this "old boys club" (Buallay et al. (2020), Adams, Gupta, and Leeth (2009)). After all, not only has the male dominance in finance not changed since, but the next big crisis caused by a collapse of our ecosystem might not be as far away as one wishes. If we do not manage now to escape the fatal collision course we are currently on, it will be too late (Townsend (2016)).

Banks play an essential role in climate change policies by ending their relationships with industries linked to high pollution and contributing to a more sustainable allocation of capital (HLEG - High-Level Group on Sustainable Finance, European Commission (2018)). Moreover, experts expect banks to be increasingly involved in environmental issues (Alberici and Querci (2016)). There is an urgent need for some significant changes in the financial industry. Reinforcing gender diversity to change the mindset of the boards toward a more sustainable company strategy is one of the most promising approaches since this concept focuses on corporate governance as well as internal processes. Women are more likely to pursue long-term strategies, are more innovative, and less committed to the status quo (Eagly, Johannesen-Schmidt, and van Engen (2003), Dezsö and Ross (2012), Torchia, Calabrò, and Huse (2011), Adams and Funk (2012)). Especially the willingness to take a new approach is essential for implementing Environmental, Social, and Governmental (ESG) practices in the banking sector. Just because things worked in the past, it does not imply they will still work in the new century of the green revolution.

Given the glass ceiling phenomenon, this paper discusses the extent to which the findings on the general population of women apply to female directors and CEOs. Based on the assumption that female directors have to break the glass ceiling, they also need to demonstrate additional characteristics. Since these attributes are mostly connected to men by society, this leads to the question of whether differences in characteristics and values can even be found if a woman shatters the imaginary barrier. Therefore, gender balanced boards would not increase the ESG score of the banks. In the context of the banking sector, women even face a double glass ceiling. The overall percentage of women in the banking industry amounts to 52%, whereas middle managers make up 38% (first glass ceiling), and the share of women in executive positions is approximately 16% (second glass ceiling) (SKEMA Observatory on the Feminization of Companies (2017)). This further suggests that even the characteristics and values between female directors and CEOs significantly differ.

There does not exist extensive literature on the connection between gender diversity and ESG performance in the banking sector since most papers focus on firms in general and further are not considering the endogeneity bias. For example, Husted and Sousa-Filho (2019), Cucari, Esposito De Falco, and Orlando (2018) find a negative connection between the two parameters for firms in Italy or Latin American countries. In contrast, the few existing papers on the banking sector of Shakil, Tasnia, and Mostafiz (2020), Birindelli et al. (2018), Birindelli, Iannuzzi, and Savioli (2019), Buallay et al. (2020) state a positive effect. Consequently, to the contradicting findings, this paper contributes to shrinking the gap in research by analyzing and comparing the results of this paper to previous studies on general firms and connecting it to the little existing literature on banks. Moreover, this thesis attempts to control for endogeneity by means of the General Method of Moments (GMM). Compared to other methods, the GMM produces consistent results in the presence of various sources of endogeneity. The data panel used for the regression covers data from 93 listed European banks over a 10-year period.

Following this chapter, a literature overview on differences in values and characteristics between genders is presented and subsequently, hypotheses are developed. In the 3. chapter, the selection of the panel data is explained, and the variables used in the regression models are classified. Chapter 4 displays the most frequent sources for the endogeneity bias. This is the fundament for chapter 5, the introduction of the GMM, which also includes a comparison of different methods when dealing with the endogeneity bias. In addition, the assumptions taken in Stata are specified and the general regression model is introduced in this section. Finally, chapter 6 presents the results and either rejects or supports the hypotheses. Lastly, limitations are pointed out, the findings are discussed, and the thesis makes propositions for future research.

2 Literature review and hypotheses development

This chapter introduces social research findings on differences in values and characteristics and links them to the economic research of the banking sector. Additionally, the existing literature is reviewed and potential factors influencing the effect of female directors on a bank's ESG performance are discussed. Subsequently, the resulting hypotheses are developed in every subchapter. At the end of this chapter, it is argued to what extent the findings of previous literature apply to this work.

2.1 Differences in values and characteristics between genders

The percentage of women on the boards is constantly increasing; it is time to consider what women can bring to the table. Women will have more and more influence on banks' decision-making and therefore affect the capital allocation and ESG performances.

The social role theory can explain different values and characteristics between genders. Differences are based on two factors: One's position within the organization and the existing gender roles and expectations (Eagly, Johannesen-Schmidt, and van Engen (2003), Eagly and Karau (2002)). Women are encouraged to pursue cooperative and relationship-building behaviors. In comparison, society trains men to be autonomous, individual-oriented, and competitive (Chodorow (1974), Gilligan (1982)). As a result of the social role theory, women are expected to adopt different leadership styles and focus on other organizational priorities than men (Eagly, Johannesen-Schmidt, and van Engen (2003), Dezsö and Ross (2012), Torchia, Calabrò, and Huse (2011), Adams and Funk (2012)).

While investigating the relationship between women on the board of directors and a bank's ESG performance, it is necessary to understand the various features of women themselves (Birindelli et al. (2018)). One component may be the educational and professional background of women. Hence, they tend to be more sensitive toward ESG related topics than men (Williams (2003), Bear, Rahman, and Post (2010)). Therefore, based on the fluency theory, it is probable that women hesitate less to adopt ESG in their decision-making. Women, in addition, are associated with the relatively more socially oriented psychological traits: Helpfulness and sensitivity to the welfare of others (Zhang, Zhu, and Ding (2013), Eagly, Johannesen-Schmidt, and van Engen (2003)). Furthermore, women are more likely to pursue long-term strategies, are more innovative, and are less committed to the status quo (Glass and Cook (2018)). Additionally, previous research showed that women are more likely to invest sustainably (Apesteguia, Azmat, and Iriberri (2012), Charness and Gneezy (2012)). In conclusion, these factors may lead to a positive influence of women on the ESG performance of a bank. Moreover, a higher share of women might not only increase the ESG score but assures that the bank performs legitimate ESG activities (Kyaw, Olugbode, and Petracci (2017), Aouadi and Marsat (2018). In a nutshell: Gender diversified boards are less likely to engage in controversies and greenwashing.

Nevertheless, Adams and Funk (2012) mention that one must critically question the general findings when applying them to women on the board of directors. In reality, given the glass ceiling phenomenon, women must demonstrate additional characteristics to break the glass ceiling and reach the board of directors. These supplementary characteristics are probably associated with men since most boards still consist predominantly of male directors. Hence, one can question whether a difference in characteristics between genders stands out when a woman is able to reach the board of directors.

Generally, women tend to be more risk-averse than men (Croson and Gneezy (2009)). However, beyond the glass ceiling, research finds that female directors are more risk-loving (Niederle, Segal, and Vesterlund (2013), Adams and Funk (2012)). A plausible explanation is that the existence of the glass ceiling may lead to female directors being more willing to accept challenges. Yet, consequently to the elevated risk, female executives also have a higher chance of failing, a potential reason for the reluctance to appoint women to leadership positions (Adams, Gupta, and Leeth (2009)). Moreover, female board members must be competitive, not only to prove themselves against male dominance but also over other female candidates.

Even though the gap in characteristics between genders seems to be smaller after breaking the glass ceiling, there exists legal evidence that women often do not get promoted by men because they are seen as too feminine (Branson (2006)). On the other hand, Berger et al. (2013) find that female CEOs often tend to promote other women into executive positions in the banking industry. Yet, since women still occupy such few seats, it is likely that women themselves sometimes hesitate to promote other female colleagues if they could potentially risk their position. This argumentation is based on the anecdote that it is still not well seen for women to promote other female colleagues. On these terms, everyone could automatically assume that the decision is based on solidarity, even though the woman has all the skills needed for a promotion. Thus, not only men should be blamed for the low percentage of women on the board of directors; women themselves are also jointly responsible.

Adams and Funk (2012) examined over 600 Swedish CEOs' and directors' responses on whether a difference in values exists between genders beyond the glass ceiling. Based on the survey responses, a significant difference in values is identified. For example, men focus more on achievements and power, whereas women emphasize universalism and benevolence. Yet, in contrast to extensive literature, they have found women CEOs and directors to be more risk-loving than men. Thus, the argument that women often do not get appointed to leadership positions because they are more risk-averse than men seems invalid.

2.1.1 Effect of female directors on the ESG performance

Several papers state that a higher proportion of women on the board of directors has a positive effect on a firm's ESG performance (Birindelli et al. (2018), Buallay et al. (2020), Shakil, Tasnia, and Mostafiz (2020)). On the contrary, some researchers find a negative connection between the two parameters for firms in Italy or Latin American countries (Husted and Sousa-Filho (2019), Cucari, Esposito De Falco, and Orlando (2018)).

The negative findings indicate that the point of view of female directors does not vary from that of male directors (Cucari, Esposito De Falco, and Orlando (2018)). This further suggests that the effect changes for different countries. In contrast to most European countries, Cucari, Esposito De Falco, and Orlando (2018) state a negative influence of female board members on the ESG score for Latin American countries. One explanation the authors of the paper mention is based on the critical mass theory. With only 53 out of 176 firms having at least one woman on their board, the percentage of women is so low that the female directors do not influence the board's decision-making. Therefore, estimating the potential effect of an increased percentage of women is difficult. Moreover, based on Hofstede (1984), all Latin American countries are characterized by high collectivism rather than individualism. This implies that one individuum has little influence on the decision-making. Consequently, a single female director probably will not increase the ESG performance of a firm. This behavior gives room to preferences of subgroups within the board and preferences of external stakeholders (Cucari, Esposito De Falco, and Orlando (2018)). Lastly, it is relevant to mention that the study was based on general firms, not specific on the banking sector.

In the context of banks, Shakil, Tasnia, and Mostafiz (2020) find a significant positive relationship between gender diversity and ESG performance. Else, Birindelli et al. (2018) showed a limited positive effect of gender diversity on a bank's ESG performance. Buallay et al. (2020) present similar findings: If the percentage of female directors lies between 22-50%, the effect on the ESG disclosure is significantly positive. But if the percentage exceeds 50%, the effect turns negative.

The literature supports the existence of differences in values and characteristics between female and male directors, even beyond the glass ceiling. Moreover, since most negative findings were based on firms in general or countries outside of Europe, the first hypothesis is the following:

H_1: A higher percentage of women on the board of directors has a positive effect on the ESG performance of a bank

2.1.2 Critical mass theory

Kanter's critical mass theory expresses that women can only exert an influence if the share of women is 30% or more. Consequently, the effects of having women on the boards are only visible if a certain critical mass of women is reached (Kanter (1977)). Namely, the critical mass can be quantified as an absolute number of at least three women (Kramer et al. (2006)). If a critical mass of women is not reached, female directors cannot provide new perspectives and bring in their skills. Kanter's theory implies a U-shaped effect of gender on the ESG score.

Joecks, Pull, and Vetter (2013) give a detailed overview of Kanter's critical mass theory. They display that the U-shaped curve can be explained with Kanter's classification of four groups according to the share of women. The first group is called the uniform group and consists exclusively of members with the same characteristics, in the context of this thesis, male directors. Next is the skewed group, composed predominantly of men and only a few women. The third classification, the titled group, is made up of a critical mass of women. Lastly, there exists a balanced group where the genders are equally represented. According to Kanter's research, the skewed group is the most problematic. Two scenarios are possible for this group: Either the minority of women gets overlooked, or it is in focus. Yet, in both cases, the women get stereotyped. The women can react to the situation in two ways: They potentially hide their unique characteristics or pretend that no differences in values and characteristics between gender exist. Moreover, the male directors will behave differently in this group. These factors lead to the finding of the relative outperformance of the skewed group by the uniform group. Thus, this is a possible explanation for the U-shaped curve of the critical mass theory.

Kanter's theory has widely been tested in legislative and political research but rarely in the business sector (Birindelli, Iannuzzi, and Savioli (2019)). One of the few examples is Liu (2018), who found a positive effect on the experienced environmental lawsuits if a firm has three or more female directors. Further, companies with a critical mass of women on their board of directors are more likely to follow an innovative management strategy, which is an essential requirement for implementing successful ESG practices (Miller and Triana (2009), Torchia, Calabrò, and Huse (2011)).

In the banking sector, the critical mass theory was tested by a few papers that primarily focused on different outcomes, such as vulnerability to the financial crisis (Farag and Mallin (2017)). Birindelli, Iannuzzi, and Savioli (2019) are some of the first to test the critical mass theory for banks, especially regarding the effect on the environmental performance. However, the hypothesis that a critical mass of women exerts a positive influence on the environmental performance of a bank is not supported. The findings reveal that an increasing share of female directors enhances the environmental performance, but the effect stops once the share exceeds 50%. Further, they find that the validity of the critical mass theory depends on the gender of a firm's CEO. While the effect of the critical mass of women on the environmental performance is U-shaped for firms with female CEOs, the effect changes concavity and is found to be an inverted U-shape in the existence of male CEOs. An indication of the potential existence of homophily. Birindelli et al. (2018) have similar findings: A positive effect for women on the boards on a bank's ESG performance, but it is insignificant. Again, the effect stopped after a critical mass of approximately 50% women was reached. Consequently, only balanced boards would enhance the environmental performance. Therefore, they state that the critical mass theory is not supported for the banking industry.

This paper examines whether there is a change in the significance level or even the overall effect when the proportion of female directors increases. Several papers find that female directors can only influence decision-making if the threshold passes the critical mass. The following second hypothesis helps to determine whether Kanter's critical mass theory is supported for the banking sector and which percentage of female directors is most efficient for enhancing the ESG score:

H_2: Female board members increase a bank's ESG performance only when the proportion of women on the board is above a critical mass

2.1.3 Effect of the CEO's gender and gender homophily

Birindelli, Iannuzzi, and Savioli (2019) find women-led banks to support environmental initiatives more strongly. This advocates that the gender of a bank's CEO may impact the ESG performance. This assumption is grounded on similar reasons as a potential positive effect of female directors on the environmental performance. Based on the gender difference perspective, female CEOs significantly affect firms' practices. In more detail, women CEOs are more likely to pursue long-term over short-term strategies and focus on non-financial performance multiples rather than short-term growth (Matsa and Miller (2013), Brown, Brown, and Anastasopoulos (2002)). Yet, there does not exist a strict consensus about the gender of the CEO influencing the ESG performance. Among the little existing literature, Glass, Cook, and Ingersoll (2016), who observed U.S. firms, find a positive link between the two variables. Additionally, Liu (2018) identified significantly reduced environmental lawsuits in the presence of female CEOs in the same market but only for firms with an overall low percentage of female directors. Borghesi, Houston, and Naranjo (2014) present that female CEOs more frequently lead firms with a higher ESG performance and reversely that women CEOs are linked to higher socially responsible investments. Furthermore, contrary to female directors, women CEOs are found to be more risk-averse. This gap in values and characteristics may be due to different job requirements and divergent career paths. In addition, same as female directors, women CEOs are more long-term oriented and more environmentally sensitive than their male counterparts (Luo et al. (2018), Glass and Cook (2018), Glass, Cook, and Ingersoll (2016)).

A plausible explanation for the CEO's gender influencing the environmental score is the homophily theory. The homophily theory is based on the similarity attraction theory: People are more attracted to people similar to them (Byrne (1971), Ibarra (1995)). Thus, people with close characteristics and values are more likely to establish stronger social boundaries (Berger et al. (2013)). An example of a similarity is belonging to the same minority within a firm: Female directors. In the context of social minorities, Kanter's theory is again relevant. In contrast to the gender difference perspective, homophily claims that the influence of women on decision-making is based on the presence of other women within the board and institution (Glass and Cook (2018)). With gender homophily, one might overcome mistrust towards the minority of women and thus increase female directors' influence on the board's decision-making. The enhanced solidarity through gender homophily might explain the findings of Liu (2018): Female CEOs increase the percentage of women on the boards. Moreover, this is supported by Bell (2005): Women in executive positions are likely to be in the highest management position when a firm has a woman as its CEO. Similar results exist for the banking industry: "Women are more likely to be appointed to executive boards that are chaired by a female CEO" (Berger et al. (2013, p. 61)). A female CEO often cultivates a more women friendly culture (Tate and Yang (2015)). Contrariwise, one can ask whether the female

CEO established a women-friendly culture or reversely: the firm has a female CEO resulting from the first established women-friendly system.

There seems to be an overall consensus about the different effects of female compared with male CEOs. The differences seem to be bigger than the gap in values and characteristics between female and male directors. Based on the findings that female CEOs are more long-term oriented and more environmentally sensitive, the third hypothesis is the following:

H_3: Female CEOs positively affect a bank's ESG performance

Further, based on the existence of gender homophily, the influence of female directors is connected to the gender of the CEO. Therefore the fourth hypothesis is:

H_4: The critical mass theory is supported for banks with women CEOs

2.1.4 Differences in the effect between countries due to gender quotas

To understand the impact of gender quotas, the findings of Adams and Funk (2012) are relevant again. The characteristics and values of quota women are expected to lay between the findings on the general population of women and the group of female directors. With gender quotas, women may not need to demonstrate supplementary characteristics, such as more competitive behavior. Consequently, it is hard to predict the effect of gender quotas on ESG performance. Huse (2018, p. 1) claims: "It is necessary to move beyond superficial accounts and take better account of who the women are." Artificially made quotas are reasonable but might not help the financial sector develop towards a more sustainable future. This belief is supported by the findings of Birindelli, Iannuzzi, and Savioli (2019) on a negative effect between gender quotas and the ESG performance of banks. The results are explained by quota women not being chosen based on their skills, knowledge, and capabilities but only due to legal and ethical pressure. A higher percentage of female directors is associated with improved social behavior. However, if the process is driven only by regulatory reasons rather than expertise, the effect on the ESG performance is negative (Cucari, Esposito De Falco, and Orlando (2018)). In reverse, given that an increasing number of countries are introducing gender quotas, regulators must believe that more women on the boards are associated with positive effects. For example, gender quotas might lead to a more diverse group of female directors and thus, potentially positively affect a bank's environmental performance. Quota women have divergent values and characteristics compared to women who fought their way up to the leadership position. But still, some values would not be represented on the boards without quota women. Buallay et al. (2020) advise that regulators should establish gender quotas to reach the optimal level of gender diversity and accordingly enhance the bank's sustainable disclosure.

Data on gender quotas can be collected from the report of Deloitte Global Boardroom Program (2022). Whereas only legally mandatory quotas count as a quota (for example in Belgium, France, Italy, Germany, Austria and Portugal). Legislative quotas without sanctions for noncompliance classify as no gender quotas (e.g., Spain). Yet, even if some countries introduced mandatory quotas some years ago, the banks are mostly given some time to implement the quotas. Hence, it is probably too early to look at the effect of gender quotas on the ESG performance. Nonetheless, as a proxy for gender quotas, it is possible to group the firms based on the countries where their headquarter is located and subsequently test whether there are differences in the ESG performances of banks in different countries. This results in the following fifth hypothesis:

H_5: The country a bank's headquarter is located in affects the ESG performance of the bank

2.2 Discussion of the findings of previous literature

This section discusses findings from previous research on women and the extent to which these implications can be used for this paper. Overall, a lesson from previous literature is that the group of women is not homogenous. Hence, it is essential to distinguish between genders and also within groups of women. Nevertheless, many existing papers are missing this point, making it difficult to compare findings between studies.

It is questionable to what extent the previous results apply to the findings of this paper since most literature focuses on the general population of women. Specific literature on the characteristics and values of female directors and CEOs is rare. It is especially challenging to generalize typical characteristics of quota women. Moreover, the existence of a gender quota does not imply that all female directors are just quota women. Hence, even the subgroup of quota women is very heterogeneous, which makes it hard to predict the effect of gender quotas. The implications of the critical mass theory suggest that a single woman on the board does not change anything. That is why rightly set gender quotas might be helpful. Nonetheless, based on previous literature, one has to overcome the potential conflict that women are needed for an increased environmental performance, although the performance only rises until a critical point is reached. Finding the critical point that optimizes the ESG performance is crucial but rather tricky. The influence of adding female directors is positive if the additional women can bring in new perspectives. If the share of women exceeds a critical point, no new value is created by adding more female directors and consequently, the effect turns negative.

Most previous studies investigating women's values used surveys, which implies that the results are based on the subjective answers of the participants. This might be problematic because the self-image does not always coincide with a person's actual behavior. Other, it is also conceivable that the selection bias influenced the findings of survey data. Female directors with stronger social behavior might be more likely to respond to this type of survey questions. Hence, the results of a survey could be biased.

3 Data collection and classification of variables

This chapter explains the selection process of the banks included in the panel data. Other, details of the dataset are defined, such as the period and number of observations. Moreover, the variables used in the regression are introduced and classified as dependent, independent or bank-specific variables.

The Thomson Reuters Database was utilized to create the panel data of the regression. Several filters were used in the collection process to reach a relatively balanced dataset. First, exclusively European banks were considered. The definition of a European bank is based on where the bank's headquarter is located. Subsequently, the dataset was reduced by only including the groups "banks" and "capital markets" of the Global Industry Classification Standard (GICS). Lastly, only banks for which data on their ESG performances of at least the last 5 years was available, were used for the regression. The final dataset consists of 93 publicly listed banks from all over Europe.

The period covers 10 years (2012-2021), whereas the decision for this span is based on a comparison of several papers. A period after the financial crisis was deliberately chosen, as after the crisis, there has been a shift towards more long-term corporate strategies, and thus ESG has become more of a focus for banks (Shakil et al. (2019)). In summary, a total of 900 yearly observations on ESG scores of banks are assembled.

Dependent variables

In this thesis, the dependent variable is the ESG performance of a bank. The ESG score used by Thomson Reuters covers more than 450 ESG metrics to calculate the ESG score of firms and provides a score between 0 and 100 (Refinitiv (2021)). The score measures three dimensions: Environmental, Social, and Governmental. The environmental measurements check the bank's ability to reduce emissions, use natural resources in their business, and support research and development of sustainable services (Birindelli et al. (2018)). Birindelli et al. (2018) summarize the social and governmental dimensions as follows. The social measurements examine the bank's capacity to respect business ethics and create value-added services and products, whereas the governmental dimensions consider the bank's ability to act in the interest of its shareholders.

A score that exclusively measures a firm's environmental performance does not exist. Therefore, we must assume that the ESG performance is a valid proxy for the bank's environmental performance. Furthermore, the use of the ESG score from Reuter is based on its use in many previous studies (Shakil et al. (2019), Birindelli, Iannuzzi, and Savioli (2019), Shakil, Tasnia, and Mostafiz (2020), Di Tommaso and Thornton (2020), Buallay et al. (2020)).

Independent variables

The independent variables included in the models are the share of women on the board of directors and dummy variables for the gender of the CEO and a critical mass of women. Only two values are possible for dummy variables: 1 or 0. For example, if a bank has a female CEO the variable equals 1, otherwise 0.

The percentage of women is an indicator for gender diversity on the boards also used in previous studies (Cucari, Esposito De Falco, and Orlando (2018), Galbreath (2018), Cordeiro, Profumo, and Tutore (2020), Husted and Sousa-Filho (2019)). To test the critical mass theory, the percentage of women is divided into four groups (less than 30%, over 30%, between 30% and 50%, and over 50%). The data for the percentage of women on the boards was collected from the Thomson Reuters Database. Annual data on the CEOs' genders was found in the banks' annual reports and press releases. When a change in CEO position occurred throughout the year, the gender of the CEO at the beginning of the year was used in the panel data.

Bank specific variables

The bank-specific variables, used as control variables, were chosen based on findings of a significant influence on the ESG performance in existing studies (Albitar et al. (2020), Arayssi, Dah, and Jizi (2016), Arayssi, Jizi, and Tabaja (2019), Husted and Sousa-Filho (2019), Shakil et al. (2019), Velte (2016)). This thesis uses the following bank-specific variables in the regression model: Bank size, board size, return on equity (ROE), return on total assets (ROA), bank leverage, board background and skills, and the existence of a Corporate Social Responsibility (CSR) committee.

Bank size is measured through the proxy logarithm of total assets. For similar reasons, the logarithm of board size is used in the regression. Birindelli et al. (2018) also use logarithms to approximate the variables' normal distribution and overcome heteroskedasticity problems. The ROA is a proxy for the operational performance, whereas the ROE measures the financial performance (Buallay (2018), Esteban-Sanchez, de la Cuesta-Gonzalez, and Paredes-Gazquez (2017)). Leverage is the percentage of total debts compared to the bank's total equity. The variables board background and skills and the existence of a CSR sustainability committee are dummy variables. Namely, the value is 1 if the existence is given, otherwise the variable is 0. Data for the bank-specific control variables were collected with the Thomson Reuters Datastream.

4 Introduction of endogeneity and frequent sources of the bias

This section explains the concept of endogeneity, a frequent source for invalid results when ignored. Moreover, the selection bias is introduced, and the three primary sources of endogeneity are discussed based on an Ordinary Least Square (OLS) regression. Lastly, the section displays how this paper attempts to control for the bias.

4.1 The concept of endogeneity

Addressing endogeneity fully is still rare in economic research but has a considerable impact on the results. According to Antonakis et al. (2010), up to 90% of papers published in premier journals do not address endogeneity adequately; therefore, those papers' causal claims might be invalid. As a consequence, recently published journals have taken endogeneity more seriously by asking the researchers to fully address the problem in their papers (Ketokivi and McIntosh (2017), Zaefarian et al. (2017), Reeb, Sakakibara, and Mahmood (2012)).

Since the ESG performance of a bank depends on many parameters and the parameters themselves are most likely correlated with each other, the existence of endogeneity is probable. The single equation model is the following:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u, \qquad (1)$$

where y is the dependent variable, $(x_1, \ldots, x_k)$ are independent variables, and u is the unobservable error term. The constant parameters $(\beta_0, \ldots, \beta_k)$ still have to be estimated in the regression.

In the context of a regression model, endogeneity is a condition where the endogenous variable is correlated with the error term (Ullah, Akhtar, and Zaefarian (2018)). The problem is that it is impossible to observe the error terms. Hence, researchers cannot statistically ensure that endogeneity is fully resolved (Roberts and Whited (2013)). Consequently, endogeneity is a potential cause for different findings in research, which result in misleading conclusions and interpretations. For example, it might be that with the existence of endogeneity, the estimations tend no longer to be true if one increases the sample size (Ullah, Akhtar, and Zaefarian (2018)).

4.2 Sources of endogeneity

This subchapter introduces the selection bias due to its relevance in the context of gender studies. Additionally, the three main sources for the existence of endogeneity in research are discussed.

4.2.1 Selection bias

In previous literature on gender diversity on the boards, the selection bias is the most frequent source of endogeneity (Adams, Hermalin, and Weisbach (2010), Hermalin and Weisbach (1998)). The selection bias occurs when researchers choose their data set based on variables that lead the selected database to differ systematically from the general population (Angrist and Pischke (2010)).

A suitable example for this is firm size and the share of women. If we assume that large banks are fundamentally different from smaller ones, the probability of appointing women to the boards of directors also varies. Therefore every regression that does not include bank size in their model is biased (Brinkhuis and Scholtens (2018), Ahern and Dittmar (2012)).

Furthermore, regarding female directors, a selection bias occurs when the values and characteristics are systematically different from the ones of the general population of women. This is probable since female directors are more frequently associated with attributes that society connects to men. This is because women on the boards must first break the glass ceiling to reach this position (Adams and Funk (2012)).

4.2.2 Omitted bias

Omitted variables influence the independent variables but were not included in the regression for various reasons. Mainly, this is problematic when doing a regression in corporate finance. Some factors relevant to corporate behavior cannot be observed and are thus not included in the regression. Hence, the omitted variables affect the error term (Roberts and Whited (2013)).

The correct relation one should use is given by:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \gamma w + u. \quad (2)$$

In (2), the unobservable independent variable w is included. In contrast, the regression estimated by us is assumed to be the following, whereas the unobservable or the so-called omitted variable is not included:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + v. \quad (3)$$

In (3), the error term is now given by $v = \gamma w + u$, compared to u in (1). If the omitted variables are uncorrelated to the independent variables in the regression, omitted variables do not affect the results. Nevertheless, if they are correlated, it results in an endogeneity problem. Therefore, the error term u is correlated to the independent variables (Wooldridge (2002)). This is a problem when doing an OLS estimation, which in the case of correlated variables, will produce inconsistent estimations (Roberts and Whited (2013)).

The existence of the omitted bias is probable. Only a limited number of parameters is included in the regression, but variables correlated to gender diversity most likely exist, which the model does not contain.

4.2.3 Simultaneity bias

The simultaneity bias occurs when researchers are unsure whether the independent variable x produces the dependent variable y or the effect is the other way around (Wooldridge (2002)). This implies that the cause can plausibly be argued.

To demonstrate the simultaneity bias, (1) can be simplified by assuming there is only one independent variable:

$$y = \beta x + u, \tag{5}$$

$$x = \alpha y + v. \tag{6}$$

For example, in this paper's context, *y* is a bank's ESG performance, and *x* is the percentage of women on the board of directors. It can be assumed that gender diversity on the board of directors influences a bank's ESG performance. Meanwhile, it might be that a bank wants to reach a certain ESG level and consequently actively appoints more women onto the board of directors since they believe that women are more sensitive to ESG factors. An example of such a mechanism that increases the percentage of female directors is a gender quota.

Borghesi, Houston, and Naranjo (2014) find that it is more likely for firms with a higher ESG performance to have a female CEO. Other, female CEOs are additionally a source for higher socially responsible investments. Their results show that the cause can plausibly be argued, and the findings may therefore be biased. This claims that this thesis must consider the simultaneity bias while doing the regression.

4.2.4 Measurement error

Most empirical studies use proxies to quantify unobservable variables. Conceptual differences between the proxies and the unobservable parameters lead to a measurement bias. Whereas these errors can affect dependent and independent variables (Wooldridge (2002)). Regarding the independent variable: If the variable is correlated to the measurement error, the OLS regression provides inconsistent and biased results. If the independent variable is uncorrelated to the measurement error, the results of the OLS regression are unbiased (Wooldridge (2002), Kennedy (1998)). A second source of the bias is data collectors recording variables incorrectly. As a result, the measurement errors are included in the regression's error term. If the measurement error is uncorrelated with the unobservable variable, it must be correlated with the observable variable. As a consequence, the bias of this regression will strive toward zero (Wooldridge (2002), Kennedy (1998), Roberts and Whited (2013)).

The measurement bias is likely to exist in this paper. For example, the proxy percentage of women on the board of directors is used to measure gender diversity on the boards. Moreover, this paper uses the proxy ESG performance to quantify the environmental score of a bank. Additionally, since this paper uses data from databanks, the data collectors themselves probably used proxies, which can lead to measurement errors.

4.3 Dealing with endogeneity in the regression

The measurement bias is difficult to eliminate for this thesis since we are forced to use data from various providers. Databases can be compared to black boxes: It is unknown how they raised the data and which proxies they used in the process. This is especially problematic for the ESG measurements. Every Database uses a different definition; a standard ESG definition does not exist yet.

As mentioned before, it is impossible to eliminate endogeneity entirely, but awareness of the bias is essential. Overall, the different findings of previous papers on the effect of women on the boards may result from some studies not addressing endogeneity in their regression. Accordingly, in this paper, the GMM is used due to its consistent results in the presence of various sources of endogeneity.

Moreover, we should be aware of a potential selection bias. The GMM can only partly control for the omitted bias, the simultaneity bias, and measurement errors. However, not for the selection bias. It is up to the researcher to select the panel data so that the data set is not systematically different from the general population.

5 The General Method of Moments: Application and comparison with other methods

In this segment, the GMM is introduced and the assumptions that must be given for the GMM to produce valid results are highlighted. Moreover, the discrepancies between the difference and system GMM are described and it is evaluated in which context one of the methods produces more efficient findings. Furthermore, in the last subchapter, the exact GMM regression model used in Stata is specified.

5.1 Discussion and comparison of the OLS, fixed effects, and GMM regression

The GMM is utilized due to its consistent results in the presence of various sources of endogeneity. Another reason for choosing the GMM over the OLS is that the GMM is often used to address panel data. Meanwhile, the OLS is mainly used for survey data (Ullah, Akhtar, and Zaefarian (2018)). Nonetheless, in this paper, the OLS is run to compare the results, but one faces the problem that if just one variable in the regression is endogenous, the results provided by the OLS are inconsistent (Ullah, Akhtar, and Zaefarian (2018)).

Furthermore, a fixed effect regression is constructed. However, only to determine the efficiency of the GMM. Problems when using a fixed effects regression occur when one of the regressors is positively correlated to the error term, for example, due to a one-year shock in the past. If the sample period is long, a single shock does not affect the results (Roodman (2009)). Yet, since the period of the panel data is relatively short in this thesis, the fixed effects regression results might be biased. On top of that, the strict assumption of exogeneity must be fulfilled to use a fixed effect regression. Hence, the data on the bank's past and present dependent variables must not affect the present data of the independent variables (Schultz, Tan, and Walsh (2010), Wintoki, Linck, and Netter (2012)). For the data of this paper, this strict assumption is probably not valid. A bank's past and current ESG performance may affect the proportion of women on the board. This suggests that the static fixed effect model produces invalid results since the relationship is rather dynamic. Consequently, lagged values should be included to differentiate the dynamic GMM from the static fixed effect model (Ullah, Akhtar, and Zaefarian (2018)).

For example, the one year lagged value of the ESG score is simply the ESG performance of the previous year.

The GMM transforms data internally by including lagged values of the dependent variable and therefore controls for endogeneity. Transforming is the statistical process where the past value of a variable is subtracted from the present value (Roodman (2009)). Based on these transformations in the GMM, the results provided by the GMM can be significantly different from the ones in an OLS or fixed effects regression. In summary, the superiority of the more complex GMM is based on the assumption that the dataset covers a short period T and a large number of individuals N. If T was large enough, the bias becomes insignificant and more straightforward methods such as fixed effects or OLS would also work (Roodman (2009)).

5.2 Introduction of the difference and system GMM

One can differentiate between the difference GMM and system GMM, first introduced by Holtz-Eakin, Newey, and Rosen (1988), Arellano and Bond (1991), Blundell and Bond (1998), Arellano and Bover (1995).

Roodman (2009) summarizes that both methods are designed for use in situations with the following factors. The period T is small, and the number of individuals N is high. Furthermore, a linear function is given and the dependent variable is dynamic, meaning the value depends on the past realization of the variable. Additionally, the independent variables are not exogenous: They are correlated with the past and present error terms. Lastly, the methods can also be used in the presence of fixed individual effects, heteroskedasticity, and autocorrelation within the individuals.

The difference GMM was developed by Arellano and Bond (1991) and consists of only one so-called original equation. In a first step, all the regressors are transformed by differencing and using the general GMM. The difference method aims to correct endogeneity and the simultaneity bias in the OLS. This is done by omitting data from the panel data but has the limitation that the lagged level of variables may be a weak instrument.

To fix this issue, Arellano and Bond (1991), Blundell and Bond (1998) invented the system GMM. The system GMM improves efficiency and allows the use of more instruments through the additional assumption that the first differences of instrument variables are uncorrelated to the fixed effects (Blundell and Bond (1998), Arellano and Bover (1995)). The system GMM consists of two equations: The original of the difference GMM and an additional transformed one.

The system GMM is superior in use when having unbalanced panel data. Additionally, the system method is also advantageous when N (number of individuals, in the context of this paper, the number of banks) is much greater than the period T (Arellano and Bond (1991), Blundell and Bond (1998)). Contrariwise, Blundell and Bond (1998) found that if the dependent variable follows closely to a random walk, the difference GMM performs poorly. This is because the difference method tries to work around endogeneity by removing fixed effects, whereas the system GMM instruments the endogenous variables with variables uncorrelated to the fixed effect. Consequently, the system GMM leads to more precise results in these circumstances (Roodman (2009)). Yet, Shakil et al. (2019) use both methods in their regression and display that, given certain factors, the results of the two methods are the same.

A disadvantage when using the system or difference GMM through the program Stata is the complexity compared to an OLS regression. Hence, the GMM can easily lead to invalid estimates. For example, for the use of the system GMM, it is essential to ensure that the changes of the instruments are uncorrelated to the fixed effects (Roodman (2009)).

5.3 Requirements for applying the GMM

The following subchapter is based on the findings of Roodman (2009). Both GMMs face the problem that if the collection of instruments is large, it can overfit the endogenous variables. Unfortunately, there exists no consensus about how many instruments are too many. For example, Roodman (2009) raised the number of instruments from 5 to 10. As a result, the estimated result went from 0.8 to 0.86, whereas the true value was 0.8. Consequently, researchers should always test the robustness of the results by reducing the number of instruments and watching that the instruments never outnumber the number of individuals.

In addition, researchers must assume that the error terms can only be correlated within an individual but not across them. Thus, one should always incorporate time dummies to ensure that time-related shocks are not universal (Roodman (2009)). Moreover, lagged variables should be included since the relationship of the variables is dynamic. For example, the past values of the ESG performance probably influence the present ESG scores. The ESG score depends on many factors that must be built up over time, meaning that significant changes in the ESG scores within a year are hard to achieve. Therefore, it is probable that the present values are correlated to the ESG scores of the past. This is other supported by the assumption that potentially banks that invested in a more environmentally friendly manner in the past will also keep this mindset in the future.

When applying the GMM, researchers must execute two post-estimation tests: The Sargan/Hansen Test and the Arellano-Bond Test. The Arellano-Bond test is used to identify potential serial correlations of the error terms. These checks are necessary to test the validity of the instruments used in the regression (Ullah, Akhtar, and Zaefarian (2018), Shakil et al. (2019)). The Sargan/Hansen Test detects if the instrument proliferation overfits the endogenous variables. Yet, at the same time, this exact phenomenon also weakens the test to detect this problem and further also the validity of the system GMM's instruments (Roodman (2009)). Therefore, it is essential to do an instrument count. Another attempt to minimize this problem is the reduction of the lags used in the instruments. To ensure that the results are valid, the Hansen value must lay between 0.05 and 0.25. Hansen values higher than 0.25 can be a potential sign of trouble (Roodman (2009)).

5.4 Application of the GMM in Stata

The regression was performed using Stata. To better explain the assumptions made in the context of the regression and the selection of the variables, the code used for the general onestep difference GMM can be found below. Whereas the prefix *l1* means that the variable is lagged by one year:

Table 1: Stata code general regression for the onestep difference GMM

```
xtabond2 ESG l1.ESG Gender l1.Gender Log_boardsize l1.Log_boardsize Log_assets l1.Log_assets ROE l1.ROE ROA l1.ROA Debt l1.Debt  CSR Skills y*, gmm(l1.ESG l1.Gender l1.Log_boardsize l1.Log_assets l1.ROE l1.ROA l1.Debt, collapse) iv(Gender Debt CSR Skills Log_boardsize ROE y*) noleveleq nodiffsargan robust
```

ESG: ESG performance, Gender: Percentage of women on the board of directors, Log_boardsize: Logarithm of number of directors, Log_assets: Logarithm of the total assets, ROE: Return on equity, ROA: Return on total assets, Debt: % Total debt/total equity, CSR: Existence of CSR sustainability committee (Dummy), Skills: Board background and skills (Dummy), y: Year dummies*

Source: Own research

Variables listed in *gmm()* are internal instruments, the so-called endogenous variables are influenced by past and present error terms. The variables listed in *iv()* are the external instruments, which consist of the independent and bank-specific variables. Further, it is essential to include the collapse command since it assures that the number of instruments is lower than the number of groups. If this assumption was not true, the results might not be valid (Roodman (2009)).

The variables used for the regression were chosen based on the study of previous literature. But in contrast to Birindelli, Iannuzzi, and Savioli (2019), no country variables were included due to problems with multicollinearity. The multicollinearity between the variables is tested using the VIF. A high VIF means the variable should not be added, this is the case for data on a country's GDP. Country variables are crucial to include when the data sample contains data from very heterogeneous countries, for example, emerging markets. Nevertheless, this paper only uses data from European banks; consequently, the countries are relatively homogenous. This factor indicates that not including the variable GDP probably did not influence the results significantly.

Both the ROA and ROE are included. This was an essential step in the process of finding a well-fitted regression. If the ROA is omitted, the Sargan/Hansen Test values rise drastically. This implies that the instruments used are not valid and are consequently exogenous. Even if a high coefficient in the Sargan/Hansen Test theoretically implies that the null hypothesis of the used instruments being accurate is supported, Roodman (2009) states that values higher than 0.25 are a potential sign of trouble. Including the variable ROA assures that the Sargen/Hansen test statistics were below 0.25 and, at the same time insignificant.

For all the variables, except the dummy variables, one year lagged values are included. The possibility of not including lagged variables was considered, but the results showed that the parameters of the Sargan/Hansen test are significant in this case, which therefore rejects the null hypothesis of the used instruments being valid. Moreover, including at least one year lagged values is suggested by Roodman (2009) when having regressors that are not strictly exogenous.

Lastly, time dummies are included in the regression model. This is essential because one must assume that individuals or not correlated across the groups. By including time dummies, this assumption is more likely to be true (Roodman (2009)).

6 Presentation of the results

This chapter presents the results of four different regression models and discusses which hypotheses are supported by the findings. Before doing so, it is debated whether the results of the difference GMM or system GMM are more efficient.

6.1 Deciding between the difference and system GMM

Even though Shakil et al. (2019) find relatively similar results for the usage of the difference and system GMM, most of the time, one of the methods is superior in use. Blundell and Bond (1998) show that if the dependent variable is distributed close to a random walk, the application of the difference GMM is biased and inefficient. Mainly, this is problematic when period T is short. In such circumstances, the relatively poor performance of the difference GMM is due to the use of poor instruments, and one should address this using the system GMM. Additionally, an OLS and fixed effects regression is implemented to decide between methods, whereas the focus is on the effect of gender diversity on ESG performance.

Table 2: Comparison results of OLS, fixed effects and GMM

	Coefficient gender	P > | t | gender	Sargan/Hansen
Pooled OLS	*0.135*	*0.004*****	-
Fixed effects	*0.126*	*0.063**	-
Twostep difference GMM	*0.118*	*0.055**	*0.117 (Sargan)* *0.122 (Hansen)*
Twostep system GMM	*0.048*	*0.522*	*0.183 (Sargan)* *0.336 (Hansen)*
*Significant at 10% level **Significant at 5% level			

Source: Own research

Table 2 reveals that the coefficients for the variable gender in the OLS, fixed effects model, and the difference GMM are relatively similar. In contrast, the coefficient in the system GMM is massively lower and not significant at any level. This displays that one should always perform the difference GMM when using the system GMM. Strict assumptions must be fulfilled that the system GMM produces more efficient results than the difference method. If one has performed the difference GMM first, significant differences will stand out, indicating that the assumptions used for the system GMM may not be accurate.

The outcomes suggest that the difference GMM is more efficient for this paper. This is further supported by the higher p-value in the Sargan/Hansen test for the system GMM. The Hansen p-value of over 0.3 indicates the existence of a problem. Therefore, it is probable that in the system GMM, not all instruments are exogenous and uncorrelated to the fixed effects.

Moreover, the dependent variable, the ESG performance, is probably not following a random walk since the past might influence the present values. This indicates that the system GMM might not be superior in use. Other, the dataset is strongly balanced with a number of individuals that is not much greater than the period. As discussed, this indicates that the usage of the system GMM does not enhance the efficiency of the results.

Roodman (2009) suggests as a rule of thumb that the estimations of *l1.esg* by the OLS and fixed effects model provide the range in which the results of the GMM should lay. In the regression, the *l1.esg* estimated by the pooled OLS was 0.832, which is assumed to be the upper bound. The fixed effects model estimated the *l1.esg* to be 0.508, providing the range's lower bound. The results for the l1.esg estimated by the twostep difference GMM are in the plausible range. However, the result of the twostep system GMM is slightly below the lower bound. This strengthens the assumption that one should primarily focus on the findings of the difference GMM.

In the context of this paper: The system GMM may not be valid for testing whether more women on the board of directors influence the ESG score of a bank. In this case, this thesis mainly focuses on the results provided by the difference GMM. Regardless, for variables other than gender diversity, the results of the system GMM are very significant. That is why also the findings of the system GMM are partly discussed since overall, the two methods complement each other well.

6.2 Comparison of the results of different regression models

Table 3 contains results for the twostep difference and system GMM of four regression models. The first model is called the general regression, already explained in chapter 5.4. The second model adds a dummy variable for a critical percentage of 30% women on the boards. Next, in the third model, a dummy variable for the gender of the CEO is included instead. Lastly, model four combines the variables from models two and three.

Table 3: Comparison results of twostep difference and system GMM for different regression models

	Diff (1)	System (1)	Diff (2)	System (2)	Diff (3)	System (3)	Diff (4)	System (4)
Gender	*0.118 (0.055*)*	*0.048 (0.522)*	*0.140 (0.043**)*	*0 .098 (0.252)*	*0.111 (0.045**)*	*0.076 (0.338)*	*0.137 (0.027**)*	*0.128 (0.145)*
Board size	*-2.579 (0.461)*	*0.368 (0.940)*	*-2.396 (0.526)*	*-0.706 (0.886)*	*-2.109 (0.521)*	*-1.219 (0.829)*	*-2.352 (0.501)*	*-2.399 (0.652)*
Bank size	*-3.014 (0.378)*	*-7.793 (0.152)*	*-2.645 (0.412)*	*-7.318 (0.180)*	*-4.708 (0.078*)*	*-6.987 (0.283)*	*-4.705 (0.118)*	*-6.100 (0.335)*
ROE	*-1.508 (0.063*)*	*-1.599 (0.010**)*	*-1.486 (0.017**)*	*-1.651 (0.008**)*	*-1.514 (0.058*)*	*-1.399 (0.039**)*	*-1.560 (0.048**)*	*-1.481 (0.067*)*
ROA	*17.185 (0.534)*	*36.349 (0.092*)*	*15.499 (0.233)*	*35.972 (0.079*)*	*23.356 (0.397)*	*25.930 (0.279)*	*25.468 (0.353)*	*25.363 (0.290)*
Debt	*-0.078 (0.926)*	*-1.429 (0.054*)*	*-0.145 (0.861)*	*-1.325 (0.049**)*	*1.158 (0.000**)*	*-1.177 (0.402)*	*1.171 (0.000**)*	*-1.088 (0.407)*
CSR	*3.814 (0.186)*	*3.947 (0.040**)*	*3.732 (0.173)*	*3.829 (0.041**)*	*4.675 (0.095*)*	*2.755 (0.151)*	*4.575 (0.092*)*	*2.629 (0.158)*
Skills	*6.432 (0.014**)*	*4.509 (0.003**)*	*6.680 (0.010**)*	*4.559 (0.003**)*	*5.489 (0.031**)*	*4.205 (0.007**)*	*5.658 (0.026**)*	*4.313 (0.006**)*
Critical	-	-	*-0.362 (0.779)*	*-1.680 (0.246)*	-	-	*-0.731 (0.534)*	*-1.962 (0.197)*
CEO	-	-	-	-	*2.002 (0.238)*	*0.751 (0.602)*	*1.964 (0.242)*	*0.843 (0.581)*

**Significant at 10% level*
***Significant at 5% level*

Values equal the estimated coefficients of the variables, values in brackets represent the respective p-value

(1) General regression, (2) With dummy variable for critical mass of female directors, (3) With dummy variable for gender CEO, (4) With critical and CEO dummy variable

Source: Own research

The table hints at a negative effect of an increasing board size. This is not congruent with previous findings for the financial sector (Birindelli et al. (2018), Baselga-Pascual et al. (2018)). Yet, a possible explanation for this finding is that the variable's coefficient is insignificant. Moreover, it is conceivable that bigger boards pursue stronger group thinking since the ideas of individuals simply get overlooked. Smaller boards potentially give more room for innovative thinking. Fewer people are involved in the decision-making; consequently, the boards might be less sluggish. Therefore, smaller boards might be more open to new approaches. Similarly, the regression shows a negative connection between bank size and ESG score. This might be due to the more complex structures of bigger banks. It possibly takes longer for larger banks to restructure the whole company toward a more environmentally focus. Further, the ROE is negatively connected to the ESG performance, whereas the ROA strongly enhances the environmental score. But in contrast to the ROE, the coefficient of the ROA is not significant. Additionally, the table displays a robust positive effect of the variable skills that is significant in all four regression models. This hints at the importance of looking at the board members' skills and backgrounds rather than just their gender. This strengthens the argument of Huse (2018) regarding the great importance of who the women are. Also, a CSR committee is closely related to a higher ESG score. Though, this is not as surprising since this indicates that the bank pays attention to environmental issues. The existence of a CSR committee helps the credibility of taking sustainability issues seriously because the committee members have the skills and knowledge to deal with such challenges (Amran, Lee, and Devi (2014)). The indication of a positive relationship is in line with the finding of previous literature (Cucari, Esposito De Falco, and Orlando (2018), Ortiz-de-Mandojana, Aguilera-Caracuel, and Morales-Raya (2016)).

The findings on the effect of the variable debt are a reminder that researchers should always be cautious when interpreting results that are not significant. By looking at the coefficients in the first two models, one could assume a negative effect, yet it is not significant. Nevertheless, the coefficients of the difference GMM in models three and four are positive at a significance level of 1%.

6.2.1 Effect of female directors on ESG performance

Table 3 in subchapter 6.2 presents a positive coefficient of the variable gender that is significant at the 10% level in all variations of the difference GMM. Similar results for the banking sector are confirmed by previous papers of Birindelli, Iannuzzi, and Savioli (2019), Buallay et al. (2020), García-Sánchez, Martínez-Ferrero, and García-Meca (2018), Gangi et al. (2019). In contrast, there exist studies finding divergent results. For example, Glass, Cook, and Ingersoll (2016) show a very limited positive effect of a higher proportion of women on boards. Moreover, Deschênes et al. (2015) claim a negative relationship between the variables, but only for non-financial firms. Birindelli, Iannuzzi, and Savioli (2019) suggest that the divergent findings can be explained by country and industry effects leading to different outcomes. This further indicates a nonlinear relationship between women on the boards and environmental scores. Also, some researchers not controlling for endogeneity is a plausible cause for the contrasting implications. Addressing endogeneity fully is still rare in economic research but has a considerable impact on the results. There are significant differences in findings using the OLS method, fixed effects, or the GMM (Ullah, Akhtar, and Zaefarian (2018)).

Even though the limited validity of the fixed effects and OLS regression due to a potential bias, both methods find a positive coefficient for the variable gender. Together with the findings of a significant positive effect of a higher proportion of female directors on the environmental score in all four regression models, H_0 is rejected. Consequently, *H_1: A higher percentage of women on the board of directors has a positive effect on the ESG performance of a bank* is supported by the results.

6.2.2 Critical mass theory

Table 4 summarizes the findings on the effect of a critical mass of women. The values in the upper line represent the results of the twostep difference GMM; the estimations of the twostep system GMM can be found in the lower line.

Table 4: Effect of a critical mass of female directors

	Coefficient dummy	Coefficient gender	t-value dummy	t-value gender	P > | t | dummy	P > | t | gender
Without dummy	-	*0.118*	-	*1.94*	-	*0.055**
for critical mass	-	*0.048*	-	*0.64*	-	*0.523*
Dummy for	*-0.362*	*0.140*	*-0.28*	*2.06*	*0.779*	*0.043***
%women > 30	*-1.681*	*0.098*	*-1.17*	*1.15*	*0.246*	*0.251*
Dummy for	*0.477*	*0.118*	*0.45*	*1.79*	*0.656*	*0.076**
50 > %women > 30	*-0.779*	*0.071*	*-0.58*	*0.88*	*0.562*	*0.380*
Dummy for	*-3.360*	*0.131*	*-1.66*	*2.07*	*0.100**	*0.041***
%women > 50	*-3.350*	*0.064*	*-1.95*	*0.77*	*0.055**	*0.442*

**Significant at 10% level*
***Significant at 5% level*

Values upper line: twostep difference GMM, values lower line: twostep system GMM

Source: Own research

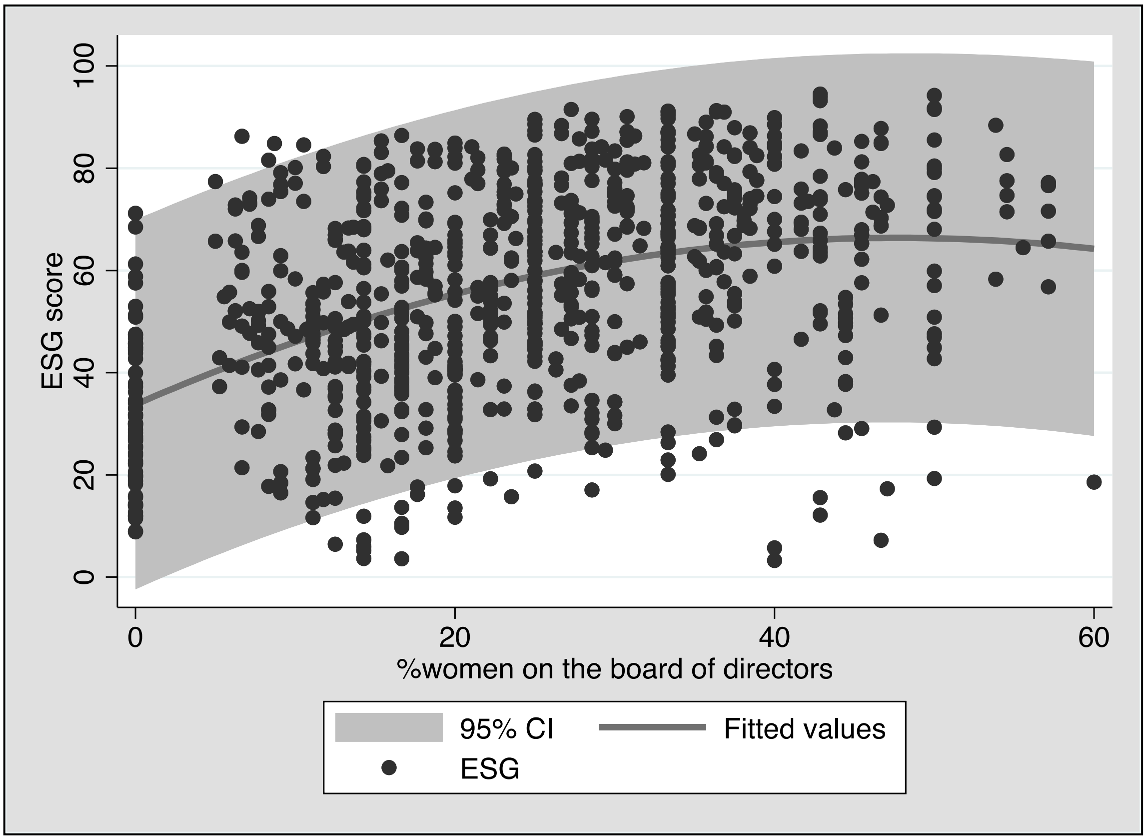
Without a dummy variable, the twostep system GMM presents a positive effect of an increasing number of female directors that is far from a significant level. The insignificance of the variable gender strengthens an assumed nonlinear relationship. By including a dummy variable for a critical mass of women, namely more than 30%, the coefficient of the variable gender in the system GMM increases, and the t-value nearly doubles. For the difference GMM, already without including a dummy variable, the coefficient of gender is significantly positive at a 10% level. But again, by including a dummy variable for a critical mass of women, the coefficient's value and the significance level rise.

One supposes the percentage of women is over the critical mass of 30%. In that case, the coefficient of the dummy variable in the twostep difference and twostep system GMM is slightly negative but not significant at any level. This leads to the assumption of a concave curve since the negative coefficient indicates that the effect of adding more women is decreasing with an overall higher percentage of female directors. This shows that the effect of gender diversity is already positive before a critical mass of 30% is reached. Therefore, Kanter's critical mass theory is generally not supported.

For the group where the percentage lies between 30 - 50%, the coefficient of the dummy variable is positive for the difference GMM and only slightly negative for the system GMM. This hints that the optimal percentage of female directors that maximizes the ESG score is settled in this range.

Finally, for data where the percentage of female directors exceeds 50%, the effect of the dummy variable on the ESG performance is again negative. Furthermore, the variable is now significant at the 10% level. This thesis already assumed that the effect has the form of a concave curve. After analyzing different sections, this thesis has now shown that the effect is an inverted U-shape, visible in the graphic below.

Figure 1: Effect of women on the board of directors on the ESG score



Source: Own research

Figure 1 indicates that if the percentage of female directors exceeds approximately 50%, the effect turns negative. The inverted U-shape suggests that gender balanced boards are most efficient for enhancing the ESG performance. This implies that female directors can only exert a positive influence as long as they can create new value. However, this is also probably the case for male directors. This assumption is additionally supported by the previous findings of Birindelli, Iannuzzi, and Savioli (2019), Buallay et al. (2020).

Furthermore, Figure 1 confirms the dual critical mass perspective (Schwartz-Ziv (2017)). The theory claims that boards are most active when they contain at least three women and three men. Moreover, these boards are more likely to take initiative, an essential condition for initiating the green revolution in the banking sector.

Aside from the discussion of whether there still exists a difference in values and beliefs between genders beyond the glass ceiling, the results show that including women to a board consisting exclusively of men enhances the ESG score. Gender balanced boards could likely improve efficiency since a more heterogeneous board could potentially reduce group thinking (Adams, Gupta, and Leeth (2009)). This might give room for induvial ideas toward developing a sustainable banking sector.

To conclude the findings of this chapter, *H_2: Female board members increase a bank's ESG performance only when the proportion of women on the board is above a critical mass* is rejected. The critical mass theory is not supported since the positive effect of female directors is visible even before a critical mass is reached. Instead, the effect of adding more women turns negative after a share of approximately 50% female directors is reached.

6.2.3 The effect of the CEO's gender and gender homophily

Table 5 on the next page displays the effect of adding a dummy variable for the gender of the CEO. The outcome reveals that adding a dummy variable for the gender of the CEO to the regression model increases the level of significance of gender diversity. Both, the difference and the system GMM, estimate a positive coefficient for the dummy variable gender of the CEO, but the effect is not significant at any level. Nonetheless, if additionally, a dummy variable for a critical mass of 30% women is included, the significance level and the coefficient of the variable gender increase. This indicates that the inclusion of the two dummy variables creates a more fitted regression model.

Table 5: Effect of women CEOs and gender homophily

	Coefficient dummy CEO	Coefficient gender	t-value dummy CEO	t-value gender	P > | t | dummy CEO	P > | t | gender
Without dummy for CEO	- -	*0.118* *0.048*	- -	*1.94* *0.64*	- -	*0.055** *0.522*
With dummy for CEO	*2.002* *0.751*	*0.110* *0.076*	*1.19* *0.52*	*2.04* *0.96*	*0.238* *0.602*	*0.045*** *0.338*
With dummy for CEO and critical mass	*1.964* *0.843*	*0.137* *0.128*	*1.18* *0.55*	*2.25* *1.47*	*0.242* *0.581*	*0.027*** *0.145*

*Significant at 10% level
**Significant at 5% level

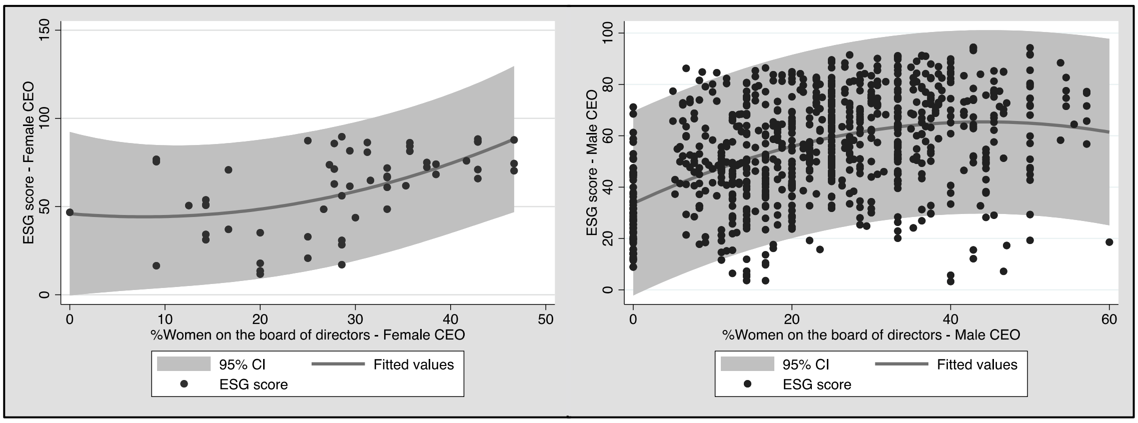
Values upper line: twostep difference GMM, values lower line: twostep system GMM

Source: Own research

Table 5 presents that the coefficient for the gender of the CEO is positive in both GMMs. This supports that women-led banks have ceteris paribus a higher ESG performance. A possible explanation for this finding is gender homophily. The theory claims that the influence of a female CEO is more substantial when simultaneously a critical mass of women on the board is reached and vice versa. In other words: Women CEOs entitle female directors to follow strong environmental initiatives (Birindelli, Iannuzzi, and Savioli (2019)). This paper finds that female directors, up to a certain level, and women CEOs increase the environmental performance. Thus, the regression model proposes that a mix of both factors maximizes a bank's ESG score.

Moreover, Figure 2 on the following page supports a potential difference in the effect of female directors when having a female compared to a male CEO. The effect of gender diversity on the boards is U-shaped for banks with a women CEO, whereas the effect is an inverted U-shape for banks with a male CEO. This finding is in line with the results of Birindelli, Iannuzzi, and Savioli (2019).

Figure 2: Difference in effect of gender diversity on the ESG score depending on the CEO's gender



Source: Own research

If a bank has a male CEO, the critical mass theory of Kanter does not seem to be supported since a positive effect of an increasing number of female directors is visible even before a critical mass of women is reached. For banks with a male CEO, gender balanced boards are associated with the highest ESG scores.

Nonetheless, the graphic shows that if a bank is led by a woman, the effect of female directors changes after reaching a critical mass of women. More precisely, the positive effect of gender diversity seems to be increased. Furthermore, Figure 2 implicates that Kanter's critical mass theory is supported for banks with women CEOs, even though the critical mass seems to be a bit lower than 30%. This demonstrates that a low percentage of female directors cannot influence the decision-making and hence, does not influence the ESG score. This can be due to male directors not recognizing the new perspectives that women can bring in, or it might be that the innovative ideas are simply not sufficiently expressed by the female directors (Joecks, Pull, and Vetter (2013)). However, if a critical mass of women is reached and simultaneously the CEO is a woman, the ESG score increases. A potential explanation for this only being valid for women CEOs is gender homophily. The inverted U-shape shows that titled and balanced groups establish more open discussions by asking different questions and providing new perspectives (Farrell and Hersch (2005), Burgess and Tharenou (2002), Apesteguia, Azmat, and Iriberri (2012)). Therefore the board of directors needs to reach a critical mass of women of about 30% to benefit from the advantages of a gender diversified board (Joecks, Pull, and Vetter

(2013)). This implies that studies, where the share of female directors is lower than 30%, are not valid for measuring the effect of gender diversity. A potential explanation for the negative results of Husted and Sousa-Filho (2019) in which the overall percentage of women was relatively low.

In theory, the left part of Figure 2 would suggest that more women always further strengthen the effect. Thus, the theoretical optimum in this model would be a percentage of 100% women. This is unrealistic since this thesis discussed that a board with uniform members is not efficient when searching for innovative approaches. Moreover, it reveals that one must be careful when interpreting these findings since the effect of the CEO's gender is not significant at any level. Other, it is essential to state that among the 910 observations on the gender of the CEOs, only 60 included a woman CEO. The relatively low number of observations makes it difficult to evaluate the validity of the findings. A scenario that combines the two figures above would be more realistic for women CEOs. More precisely, a curve with two extrema: One when a critical mass of female directors is reached, and the effect of adding more women to the board turns positive, and a second one, the maxima, when the effect of adding more women to the board turns negative.

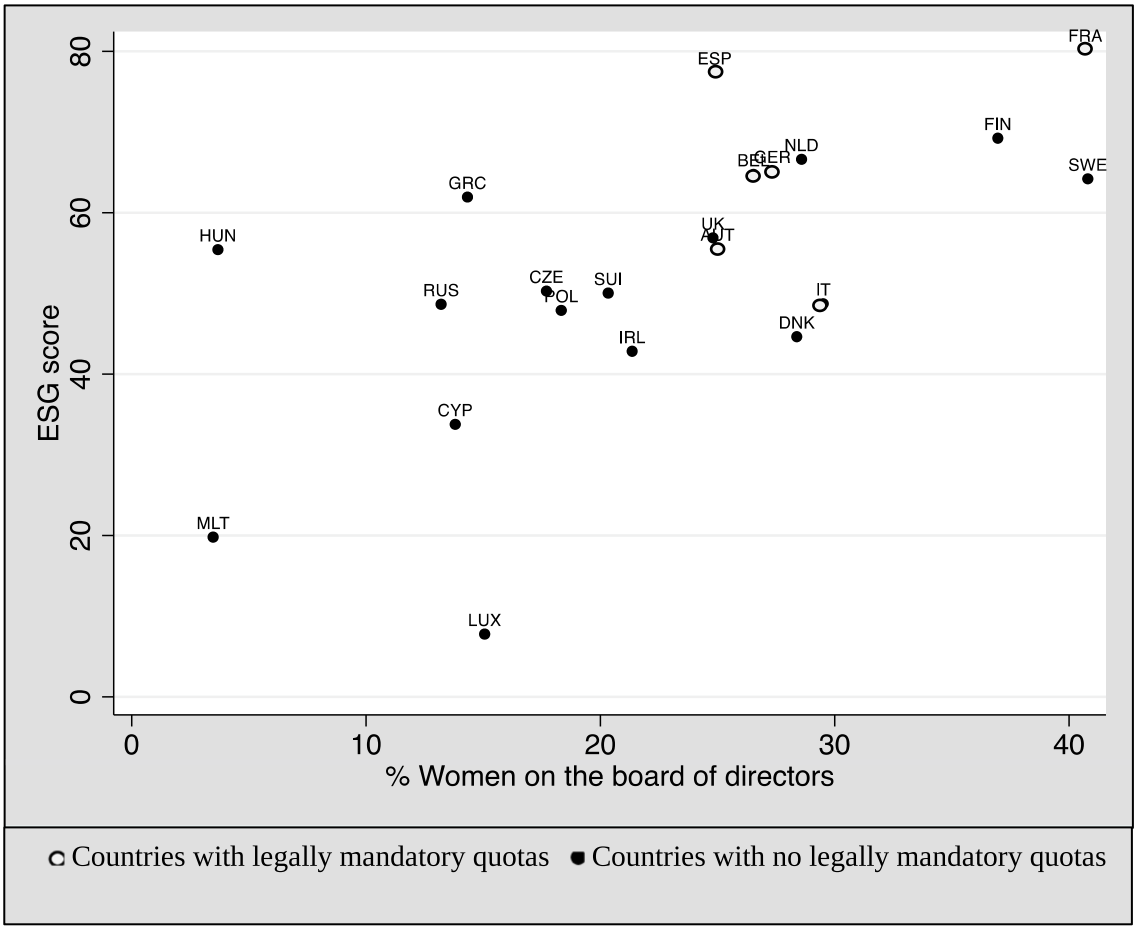
The positive coefficient of the variable gender of the CEO indicates that H_3*: Female CEOs positively affect a bank's ESG performance* might be supported. Yet, the effect is not significant at any level, meaning H_0 cannot be rejected. Therefore, H_3 is only indicated but not fully supported.

The findings regarding H_4*: The critical mass theory is supported for banks with women CEOs* are similar. The critical mass theory seems to be supported by the regression results and Figure 2. But again, the results are insignificant; consequently, the regression cannot reject H_0. The insignificant result might partly be due to small data available on banks with a female CEO and simultaneously a percentage of female directors higher than 30%.

6.2.4 Differences in the effect between countries due to gender quotas

Including a dummy variable for gender quotas in the regression is impossible. The possibility of developing different regression models for the countries and comparing them to each other was tested, but the results were not valid due to the small amount of data available per country. However, it is feasible to look at how the mean scores of the percentages of female directors and the ESG performances differ between countries. Figure 3 illustrates this by adding a country's mean percentage of women on the boards to the x-coordinate, whereas a country's mean ESG score is situated on the y-coordinate. Nonetheless, the limitation remains that the data samples per country are relatively small and unbalanced. It might be that the mean scores change when including more data per country.

Figure 3: Comparison of mean scores of female directors and environmental performances between countries



Source: Own research

At first, one could assume that the graph shows that if a country's mean percentage of female directors increases, the mean ESG score also rises. Yet, it must be remembered that the higher ESG score can also be due to other reasons since these results are not based on the GMM; it is simply a comparison of the mean scores. For example, it is conceivable that countries with a higher degree of innovation are more open to gender diversity which therefore implements a higher proportion of female directors. Moreover, the higher degree of innovation could imply that the banks are more open to sustainable solutions, which increases the ESG score. This suggests that the higher ESG scores might not only be due to having more female directors but are also based on an overall more open mindset in the country or various other reasons. Further, we should look closely at Reuter's definition of a firm's ESG performance. Likely, some variables measuring gender diversity are already included in the score. A potential explanation for the indicated connection between the two parameters. These factors show that Figure 3 should not be used to describe the effect of an increasing percentage of female directors on the ESG score. Nevertheless, it gives us a good overview of the implementation of ESG practices in the countries and their development of gender diversity on the boards.

France and Sweden have the highest mean percentages of female directors. France has a legally mandatory quota, Sweden does not. Among the top three countries with the highest share of women on the boards is also Finland which has no mandatory quotas as well. This might suggest that a higher percentage of women on the boards can also be achieved without quotas. For example, through a change in the overall mindset. Apart from these three leading countries, the percentage in the remaining countries is significantly lower. The low percentages are surprising since many previous studies showed positive effects of having gender diversified boards.

Figure 3 shows that the percentage of women on the boards and ESG scores are widely spread. This indicates that the group is not homogenous even if all banks are located in Europe. It seems that some banks still have not realized the urgency of a shift towards a greener future. With mandatory quotas, it is conceivable that the percentage of female directors will rise, although the associated positive effect on the ESG performance might not be as strong as the theoretical findings. Banks must discard old ways of thinking;

gender quotas probably cannot fix this problem. This suggests that strict gender quotas might not be the most efficient way to raise the environmental performance simply because the effects of adding more women to the boards are individual for every bank and country.

One attempt for the regulators to solve the problem might be implementing a minimal ESG score that all banks have to exceed as well as the target of an overall share of female directors for the financial sector. This allows every bank to implement the percentage of women that is most efficient for them in enhancing the ESG score. Further, this attempt would assure that regulators can achieve an optimal overall percentage of female directors. This macroeconomic attempt is already an efficient way to solve social problems. For example, emission certificates that help reduce a country's emissions to a certain level.

Due to the relatively small panel data, developing a different regression model per country was impossible. Therefore, *H_5: The country a bank's headquarter is located in affects the ESG performance of the bank* can neither be rejected nor supported. Yet, based on the existing literature, it is probable that the effect of having quota women differs from adding other female directors in general.

6.3 Discussion and limitations of results

By excluding banks from the dataset, it might be that the selection bias affects the results of this paper up to a certain level. Banks with no data on their ESG scores and percentages of female directors may pay less attention to sustainable solutions and gender diversity on the boards. Furthermore, the panel data consists exclusively of listed banks. Yet it might be that the results vary for smaller, not listed banks due to a lack of regulations. For example, in most countries, gender quotas are only mandatory for listed banks.

Joecks, Pull, and Vetter (2013) discuss this exact problem. Since the effect is rather an inverted U-shape than linear, the findings depend on whether the panel data consists of banks with an overall low or high percentage of female directors. For an overall low share of women on the boards, researchers exclusively examine the rising part of the inverted U-curve. Hence, the effect is assumed to be positive and upward biased. In the second case, when the researchers have data with an overall high percentage of women on the boards, they assume a negative effect. Nonetheless, this is less realistic for this study since our dataset's highest share of female directors is 60%. In conclusion, the positive effect of an increasing percentage of female directors is only supported for levels up to 60%.

It is tricky to compare the results of our paper to other studies. Even if we exclusively look at the difference GMM, the model produces significantly different results with the inclusion of just one additional variable. This is problematic since many papers do not specify their exact regression model. For example, the results vary depending on whether a one- or twostep GMM is applied. Moreover, the difference or system GMM lead to different results as well. It is conceivable that researchers present the models' findings that support their assumptions rather than the most efficient ones.

The divergent findings on the effects of specific control variables might also be due to us considering different regression models. Some coefficients change effect if one simply looks at the covariance matrix rather than the coefficients of the GMM regression. However, this is not concerning since this paper focuses on the effect of gender diversity on the boards. For this variable, the GMM, OLS, fixed effects model, and even the covariance matrix all show a positive effect of gender on environmental performance.

Besides the country where the bank is located and the choice of instruments potentially causing different findings, another factor is the choice of the ESG measurement. Every databank considers different measurements for their ESG scores. The results might change depending on how many of the included variables positively correlate with the proportion of women. Among the over 450 metrics included in Reuter's ESG score, some variables measuring gender diversity are likely already included since the score not only measures the environmental performance but also contains social and governmental factors. Therefore, the effect of gender diversity on the ESG performance might be upwards biased. It will be interesting for future research to compare the effects of gender on the ESG performance for different ESG measurements. In addition, the validity of the results is based on the assumption that the ESG score is an effective proxy for measuring the sustainability performance of a bank.

Lastly, it is essential to mention that this paper's results only apply to the banking sector. It might be that the shape of the effect of gender diversified boards is different for other sectors. This is due to the various special features of the banking sector. Banks are stronger regulated than other institutions due to their crucial role in the economy and often exhibit higher leverage. Furthermore, the shares of female directors are relatively lower compared to other sectors.

7 Conclusion

The literature and research review reveal a gap in value and characteristics between genders in general and even beyond the glass ceiling on the board of directors, yet the gap seems smaller compared to the general population. Furthermore, research supports differences within subgroups of women. Therefore, the findings on the general group of women cannot be applied to female directors since female board members must demonstrate supplementary characteristics that society classifies as male attributes. Hence, it is impossible to stereotype women based on specific features. Consequently, the effect of having a higher share of women on the boards might not only be due to the stereotypical characteristics of the general population of women. A better explanation is the outperformance of titled groups compared to uniform groups. Of course, the positive effect of an increasing number of women on the boards might be due to divergent characteristics compared to male directors. Nonetheless, we assume that a positive effect might also be visible if male directors are included, which can bring in new values and perspectives. Hence, rather than only looking at gender, we should additionally concentrate on the significant positive influence of the directors' skills and background. In the end, what is essential for an increased ESG score is that the additional director creates new value. Yet this is not automatically correlated to one's gender. In conclusion, we suggest focusing on overall diversified boards rather than just considering the gender of the board members.

The results of this thesis show that the effect of female directors on the ESG score is significantly positive. However, the inverted U-shape suggests that Kanter's critical mass theory is not generally supported for the banking sector since a positive effect is visible even before a critical mass of women is reached. The form of the effect further implies that gender balanced boards are most efficient for enhancing the environmental performance. Moreover, female directors not only increase the ESG score but assure that the bank performs legitimate ESG activities. Consequently, it might be that the effect of adding more women to the boards is even higher if one would only consider legitimate ESG practices and thus eliminate potential greenwashing.

This paper displays the critical mass theory's great importance when having a female CEO. The effect of higher shares of female directors on ESG scores is U-shaped but only for banks with a female CEO due to the possible existence of gender homophily. This suggests that regulators should not only help women break the first glass ceiling to reach the boards but, most importantly, shatter the second glass ceiling to increase the number of female CEOs in the banking sector. In contrast, when a bank is led by a man, the effect of having more women on the boards seems limited.

It was impossible to include gender quotas in our model and therefore investigate the differences in effects between countries. However, the assumed differences between quota women and female directors in general and the widely spread mean scores imply that static gender quotas might not be efficient. A higher percentage of women does not automatically enhance the ESG score if the bank's mindset does not change simultaneously. Hence, it might be more efficient to target a specific ESG score for all the banks and an overall percentage of female directors for the sector rather than implementing static gender quotas for whom the effect is unclear for individual banks.

In summary, our finding of a positive association between a higher proportion of female directors and a bank's environmental performance sheds empirical light on the importance of gender diversified boards. Our results are particularly relevant since only few previous papers have examined this effect while controlling for endogeneity. In contrast to extensive literature, which argues that the positive effect is due to stereotypical characteristics of women, we state that differences in skills and backgrounds caused by the double glass ceiling present a more valid explanation. Therefore, banks should focus on general board diversity, not just on the gender of their board members.

Future research should consider investigating the effect of gender quotas and the influence of high shares of women on the boards since the banking industry's overall percentage is still relatively low nowadays. Additionally, since this thesis states that the critical mass theory is not supported, future research should subsequently test the dual critical mass theory for the banking sector. Lastly, researchers should concentrate on giving concrete advice on implementing best higher shares of female directors rather than just showing the positive theoretical effects of having a gender diversified board.

8 References

Adams, Renée B., and Patricia Funk, 2012, Beyond the Glass Ceiling: Does Gender Matter?, *Management Science* 58, 219–235.

Adams, Renée B., Benjamin E. Hermalin, and Michael S. Weisbach, 2010, The Role of Boards of Directors in Corporate Governance: A Conceptual Framework and Survey, *Journal of Economic Literature* 48, 58–107.

Adams, Susan M., Atul Gupta, and John D. Leeth, 2009, Are Female Executives Over-represented in Precarious Leadership Positions?, *British Journal of Management* 20, 1–12.

Ahern, Kenneth R., and Amy K. Dittmar, 2012, The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation, *The Quarterly Journal of Economics* 127, 137–197.

Alberici, Adalberto, and Francesca Querci, 2016, The Quality of Disclosures on Environmental Policy: The Profile of Financial Intermediaries, *Corporate Social Responsibility and Environmental Management* 23, 283–296.

Albitar, Khaldoon, Khaled Hussainey, Nasir Kolade, and Ali M. Gerged, 2020, ESG disclosure and firm performance before and after IR: The moderating role of governance mechanisms, *International Journal of Accounting & Information Management* 28, 429–444.

Amran, Azlan, Shiau P. Lee, and Susela S. Devi, 2014, The Influence of Governance Structure and Strategic Corporate Social Responsibility Toward Sustainability Reporting Quality, *Business Strategy and the Environment* 23, 217–235.

Angrist, Joshua D., and Jörn-Steffen Pischke, 2010, The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics, *Journal of Economic Perspectives* 24, 3–30.

Antonakis, John, Samuel Bendahan, Philippe Jacquart, and Rafael Lalive, 2010, On making causal claims: A review and recommendations, *The Leadership Quarterly* 21, 1086–1120.

Aouadi, Amal, and Sylvain Marsat, 2018, Do ESG Controversies Matter for Firm Value? Evidence from International Data, *Journal of Business Ethics* 151, 1027–1047.

Apesteguia, Jose, Ghazala Azmat, and Nagore Iriberri, 2012, The Impact of Gender Composition on Team Performance and Decision Making: Evidence from the Field, *Management Science* 58, 78–93.

Arayssi, Mahmoud, Mustafa Dah, and Mohammad Jizi, 2016, Women on boards, sustainability reporting and firm performance, *Sustainability Accounting, Management and Policy Journal* 7, 376–401.

Arayssi, Mahmoud, Mohammad Jizi, and Hala H. Tabaja, 2019, The impact of board composition on the level of ESG disclosures in GCC countries, *Sustainability Accounting, Management and Policy Journal* 11, 137–161.

Arellano, Manuel, and Stephen Bond, 1991, Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *The Review of Economic Studies* 58, 277.

Arellano, Manuel, and Olympia Bover, 1995, Another look at the instrumental variable estimation of error-components models, *Journal of Econometrics* 68, 29–51.

Baselga-Pascual, Laura, Antonio Trujillo-Ponce, Emilia Vähämaa, and Sami Vähämaa, 2018, Ethical Reputation of Financial Institutions: Do Board Characteristics Matter?, *Journal of Business Ethics* 148, 489–510.

Bear, Stephen, Noushi Rahman, and Corinne Post, 2010, The Impact of Board Diversity and Gender Composition on Corporate Social Responsibility and Firm Reputation, *Journal of Business Ethics* 97, 207–221.

Bell, Linda, 2005, Women-Led Firms and the Gender Gap in Top Executive Jobs, Discussion Paper No. 1689, Haverford College and IZA Bonn.

Berger, Allen N., Thomas Kick, Michael Koetter, and Klaus Schaeck, 2013, Does it pay to have friends? Social ties and executive appointments in banking, *Journal of Banking & Finance* 37, 2087–2105.

Birindelli, Giuliana, Stefano Dell'Atti, Antonia P. Iannuzzi, and Marco Savioli, 2018, Composition and Activity of the Board of Directors: Impact on ESG Performance in the Banking System, *Sustainability* 10, 4699.

Birindelli, Giuliana, Antonia P. Iannuzzi, and Marco Savioli, 2019, The impact of women leaders on environmental performance: Evidence on gender diversity in banks, *Corporate Social Responsibility and Environmental Management* 26, 1485–1499.

Blundell, Richard, and Stephen Bond, 1998, Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics* 87, 115–143.

Borghesi, Richard, Joel F. Houston, and Andy Naranjo, 2014, Corporate socially responsible investments: CEO altruism, reputation, and shareholder interests, *Journal of Corporate Finance* 26, 164–181.

Branson, Douglas M., 2006, *No Seat at the Table: How Corporate Governance Keeps Women Out of America's Boardrooms* (New York University Press, New York).

Brinkhuis, Eline, and Bert Scholtens, 2018, Investor response to appointment of female CEOs and CFOs, *The Leadership Quarterly* 29, 423–441.

Brown, David H., Debra L. Brown, and Vanessa Anastasopoulos, Canada: The Conference Board of Canada, 2002, *Women on boards: Not just the right thing ... but the "bright" thing*.

Buallay, Amina, 2018, Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector, *Management of Environmental Quality: An International Journal* 30, 98–115.

Buallay, Amina, Reem Hamdan, Elisabetta Barone, and Allam Hamdan, 2020, Increasing female participation on boards: Effects on sustainability reporting, *International Journal of Finance & Economics* 27, 111–124.

Burgess, Zena, and Phyllis Tharenou, 2002, Women Board Directors: Characteristics of the Few, *Journal of Business Ethics* 37, 39–49.

Byrne, Donn E., 1971, *The Attraction Paradigm* (Academic Press, New York).

Charness, Gary, and Uri Gneezy, 2012, Strong Evidence for Gender Differences in Risk Taking, *Journal of Economic Behavior & Organization* 83, 50–58.

Chodorow, Nancy, 1974, Family structure and feminist perspective, in Michelle Z. Rosaldo, Louise Lamphere, and Joan Bamberger, ed.: *Women in culture and society* (Stanford University Press).

Cordeiro, James J., Giorgia Profumo, and Ilaria Tutore, 2020, Board gender diversity and corporate environmental performance: The moderating role of family and dual-class majority ownership structures, *Business Strategy and the Environment* 29, 1127–1144.

Croson, Rachel, and Uri Gneezy, 2009, Gender Differences in Preferences, *Journal of Economic Literature* 47, 448–474.

Cucari, Nicola, Salvatore Esposito De Falco, and Beatrice Orlando, 2018, Diversity of Board of Directors and Environmental Social Governance: Evidence from Italian Listed Companies, *Corporate Social Responsibility and Environmental Management* 25, 250–266.

Deloitte Global Boardroom Program, *Women in the boardroom a Global perspective*, 2022 (Deloitte).

Deschênes, Sébastien, Miguel Rojas, Hamadou Boubacar, Brigitte Prud'homme, and Alidou Ouedraogo, 2015, The impact of board traits on the social performance of Canadian firms, *Corporate Governance* 15, 293–305.

Dezsö, Cristian L., and David Gaddis Ross, 2012, Does female representation in top management improve firm performance? A panel data investigation, *Strategic Management Journal* 33, 1072–1089.

Di Tommaso, Caterina, and John Thornton, 2020, Do ESG scores effect bank risk taking and value? Evidence from European banks, *Corporate Social Responsibility and Environmental Management* 27, 2286–2298.

Eagly, Alice H., Mary C. Johannesen-Schmidt, and Marloes L. van Engen, 2003, Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men, *Psychological Bulletin* 129, 569–591.

Eagly, Alice H., and Steven J. Karau, 2002, Role congruity theory of prejudice toward female leaders, *Psychological Review* 109, 573–598.

Esteban-Sanchez, Pablo, Marta de la Cuesta-Gonzalez, and Juan D. Paredes-Gazquez, 2017, Corporate social performance and its relation with corporate financial performance: International evidence in the banking industry, *Journal of Cleaner Production* 162, 1102–1110.

Farag, Hisham, and Chris Mallin, 2017, Board diversity and financial fragility: Evidence from European banks, *International Review of Financial Analysis* 49, 98–112.

Farrell, Kathleen A., and Philip L. Hersch, 2005, Additions to corporate boards: the effect of gender, *Journal of Corporate Finance* 11, 85–106.

Galbreath, Jeremy, 2018, Is Board Gender Diversity Linked to Financial Performance? The Mediating Mechanism of CSR, *Business & Society* 57, 863–889.

Gangi, Francesco, Antonio Meles, Eugenio D'Angelo, and Lucia M. Daniele, 2019, Sustainable development and corporate governance in the financial system: Are environmentally friendly banks less risky?, *Corporate Social Responsibility and Environmental Management* 26, 529–547.

García-Sánchez, Isabel-María, Jennifer Martínez-Ferrero, and Emma García-Meca, 2018, Board of Directors and CSR in Banking: The Moderating Role of Bank Regulation and Investor Protection Strength, *Australian Accounting Review* 28, 428–445.

Gilligan, Carol, 1982, *In a Different Voice: Psychological Theory and Women's Development* (Harvard University Press, Cambridge).

Glass, Christy, and Alison Cook, 2018, Do women leaders promote positive change? Analyzing the effect of gender on business practices and diversity initiatives, *Human Resource Management* 57, 823–837.

Glass, Christy, Alison Cook, and Alicia R. Ingersoll, 2016, Do Women Leaders Promote Sustainability? Analyzing the Effect of Corporate Governance Composition on Environmental Performance, *Business Strategy and the Environment* 25, 495–511.

Hermalin, Benjamin E., and Michael S. Weisbach, 1998, Endogenously Chosen Boards of Directors and Their Monitoring of the CEO, *The American Economic Review* 88, 96–118.

HLEG - High-Level Group on Sustainable Finance, European Commission, 2018, *Financing a sustainable European economy*.

Hofstede, Geert, 1984, The Cultural Relativity of the Quality of Life Concept, *The Academy of Management Review* 9, 389–398.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen, 1988, Estimating Vector Autoregressions with Panel Data, *Econometrica* 56, 1371–1395.

Huse, Morten, 2018, Gender in the Boardroom: Learnings from world-leader Norway, *FACTBase Bulletin* 58, 1-16.

Husted, Bryan W., and José Milton de Sousa-Filho, 2019, Board structure and environmental, social, and governance disclosure in Latin America, *Journal of Business Research* 102, 220–227.

Ibarra, Herminia, 1995, Race, Opportunity, and Diversity of Social Circles in Managerial Networks, *The Academy of Management Journal* 38, 673–703.

Joecks, Jasmin, Kerstin Pull, and Karin Vetter, 2013, Gender Diversity in the Boardroom and Firm Performance: What Exactly Constitutes a "Critical Mass?," *Journal of Business Ethics* 118, 61–72.

Kanter, Rosabeth M., 1977, Being Better by Being Right: Subjective Group Dynamics and Derogation of In-Group Deviants When Generic Norms Are Undermined, *American Journal of Sociology* 81, 965–990.

Kennedy, Peter, 1998, *A Guide to Econometrics* (MIT Press, Cambridge, Mass).

Ketokivi, Mikko, and Cameron N. McIntosh, 2017, Addressing the endogeneity dilemma in operations management research: Theoretical, empirical, and pragmatic considerations, *Journal of Operations Management* 52, 1–14.

Kramer, Wicky V., Alison M. Konrad, Sumru Erkut, and Michele J. Hopper, *Critical mass on corporate boards: Why three or more women enhance governance*, *2016* (Wellesley Centers for Women, Wellesley)

Kyaw, Khine, Mojisola Olugbode, and Barbara Petracci, 2017, Can board gender diversity promote corporate social performance?, *Corporate Governance: The International Journal of Business in Society* 17, 789–802.

Liu, Chelsea, 2018, Are women greener? Corporate gender diversity and environmental violations, *Journal of Corporate Finance* 52, 118–142.

Luo, Jin-hui, Zeyue Huang, Xue Li, and Xiaojing Lin, 2018, Are Women CEOs Valuable in Terms of Bank Loan Costs? Evidence from China, *Journal of Business Ethics* 153, 337–355.

Matsa, David A., and Amalia R. Miller, 2013, A Female Style in Corporate Leadership? Evidence from Quotas, *American Economic Journal: Applied Economics* 5, 136–169.

Miller, Toyah, and María Triana, 2009, Demographic Diversity in the Boardroom: Mediators of the Board Diversity–Firm Performance Relationship, *Journal of Management Studies* 46, 755–786.

Niederle, Muriel, Carmit Segal, and Lise Vesterlund, 2013, How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness, *Management Science* 59, 1–16.

Ortiz-de-Mandojana, Natalia, Javier Aguilera-Caracuel, and Matilde Morales-Raya, 2016, Corporate Governance and Environmental Sustainability: The Moderating Role of the National Institutional Context, *Corporate Social Responsibility and Environmental Management* 23, 150–164.

Reeb, David, Mariko Sakakibara, and Ishtiaq P. Mahmood, 2012, From the Editors: Endogeneity in international business research, *Journal of International Business Studies* 43, 211–218.

Refinitiv, *Environmental, Social and Governance scores from Refinitiv,* 2021 (Refinitiv).

Roberts, Michael R., and Toni M. Whited, 2013, Chapter 7 - Endogeneity, in George M. Constantinides, Milton Harris, and Rene M. Stulz, ed.: *Empirical Corporate Finance1, Handbook of the Economics of Finance* (Elsevier).

Roodman, David, 2009, How to do Xtabond2: An Introduction to Difference and System GMM in Stata, *The Stata Journal* 9, 86–136.

Schultz, Emma L., David T. Tan, and Kathleen D. Walsh, 2010, Endogeneity and the corporate governance - performance relation, *Australian Journal of Management* 35, 145–163.

Schwartz-Ziv, Miriam, 2017, Gender and Board Activeness: The Role of a Critical Mass, *Journal of Financial and Quantitative Analysis* 52, 751–780.

Shakil, Mohammad H., Nihal Mahmood, Mashiyat Tasnia, and Ziaul Haque M., 2019, Do environmental, social and governance performance affect the financial performance of banks? A cross-country study of emerging market banks, *Management of Environmental Quality: An International Journal* 30, 1331–1344.

Shakil, Mohammad H., Mashiyat Tasnia, and Md Imtiaz Mostafiz, 2020, Board gender diversity and environmental, social and governance performance of US banks: moderating role of environmental, social and corporate governance controversies, *International Journal of Bank Marketing* 39, 661–677.

SKEMA Observatory on the Feminization of Companies, 2017, Gender diversity in the banking industry An international comparison, SKEMA Business School.

Tate, Geoffrey, and Liu Yang, 2015, Female leadership and gender equity: Evidence from plant closure, *Journal of Financial Economics* 117, 77–97.

Torchia, Mariateresa, Andrea Calabrò, and Morten Huse, 2011, Women Directors on Corporate Boards: From Tokenism to Critical Mass, *Journal of Business Ethics* 102, 299–317.

Townsend, Mike, 2016, *The Quiet Revolution: Towards a Sustainable Economy* (Greenleaf Publishing, London).

Ullah, Subhan, Pervaiz Akhtar, and Ghasem Zaefarian, 2018, Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data, *Industrial Marketing Management* 71, 69–78.

Velte, Patrick, 2016, Women on management board and ESG performance, *Journal of Global Responsibility* 7, 98–109.

Williams, Robert J., 2003, Women on Corporate Boards of Directors and their Influence on Corporate Philanthropy, *Journal of Business Ethics* 42, 1–10.

Wintoki, M. Babajide, James S. Linck, and Jeffry M. Netter, 2012, Endogeneity and the dynamics of internal corporate governance, *Journal of Financial Economics* 105, 581–606.

Wooldridge, Jeffrey M., 2002, *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge).

Zaefarian, Ghasem, Vita Kadile, Stephan C. Henneberg, and Alexander Leischnig, 2017, Endogeneity bias in marketing research: Problem, causes and remedies, *Industrial Marketing Management* 65, 39–46.

Zhang, Jason Q., Hong Zhu, and Hung-bin Ding, 2013, Board Composition and Corporate Social Responsibility: An Empirical Investigation in the Post Sarbanes-Oxley Era, *Journal of Business Ethics* 114, 381–392.

Database

Thomson Reuters Database/Datastream, EIKON, 19.04.2022

Appendix

Appendix A: List of European banks in panel data

Table 6: List of European banks in dataset

Company Name	Country of Headquarters	GICS Industry
3i Group PLC	United Kingdom	Capital Markets
ABN Amro Bank NV	Netherlands	Banks
Abrdn PLC	United Kingdom	Capital Markets
Alior Bank SA	Poland	Banks
Alpha Services and Holdings SA	Greece	Banks
Ashmore Group PLC	United Kingdom	Capital Markets
Azimut Holding SpA	Italy	Capital Markets
Banca Monte dei Paschi di Siena SpA	Italy	Banks
Banca Popolare di Sondrio ScpA	Italy	Banks
Banco Bilbao Vizcaya Argentaria SA	Spain	Banks
Banco BPM SpA	Italy	Banks
Banco de Sabadell SA	Spain	Banks
Banco Santander SA	Spain	Banks
Bank Handlowy w Warszawie SA	Poland	Banks
Bank Millennium SA	Poland	Banks
Bank of Cyprus Holdings PLC	Cyprus	Banks
Bank of Georgia Group PLC	United Kingdom	Banks
Bank of Ireland Group PLC	Ireland; Republic of	Banks
Bank Polska Kasa Opieki SA	Poland	Banks
Bank VTB PAO	Russia	Banks
Bankinter SA	Spain	Banks
Banque Cantonale Vaudoise	Switzerland	Banks
Barclays PLC	United Kingdom	Banks
BNP Paribas SA	France	Banks
Bper Banca SpA	Italy	Banks
Brait PLC	Malta	Capital Markets
Brewin Dolphin Holdings PLC	United Kingdom	Capital Markets
Caixabank SA	Spain	Banks
Close Brothers Group PLC	United Kingdom	Banks
CMC Markets PLC	United Kingdom	Capital Markets
Commerzbank AG	Germany	Banks
Credit Agricole SA	France	Banks
Credit Suisse Group AG	Switzerland	Capital Markets
Danske Bank A/S	Denmark	Banks
Deutsche Bank AG	Germany	Capital Markets
Deutsche Boerse AG	Germany	Capital Markets
EFG International AG	Switzerland	Capital Markets
Erste Group Bank AG	Austria	Banks
Eurobank Ergasias Services and Holdings SA	Greece	Banks
Euronext NV	Netherlands	Capital Markets
GAM Holding AG	Switzerland	Capital Markets
Getin Holding SA	Poland	Banks
Hargreaves Lansdown PLC	United Kingdom	Capital Markets

Hellenic Exchanges Athens Stock Exchange SA	Greece	Capital Markets
HSBC Holdings PLC	United Kingdom	Banks
IG Group Holdings PLC	United Kingdom	Capital Markets
ING Bank Slaski SA	Poland	Banks
ING Groep NV	Netherlands	Banks
Intermediate Capital Group PLC	United Kingdom	Capital Markets
Intesa Sanpaolo SpA	Italy	Banks
Investec PLC	United Kingdom	Capital Markets
IP Group PLC	United Kingdom	Capital Markets
Janus Henderson Group PLC	United Kingdom	Capital Markets
Julius Baer Gruppe AG	Switzerland	Capital Markets
Jupiter Fund Management PLC	United Kingdom	Capital Markets
Jyske Bank A/S	Denmark	Banks
KBC Groep NV	Belgium	Banks
Komercni Banka as	Czech Republic	Banks
Lloyds Banking Group PLC	United Kingdom	Banks
London Stock Exchange Group PLC	United Kingdom	Capital Markets
Man Group PLC	United Kingdom	Capital Markets
mBank SA	Poland	Banks
Mediobanca Banca di Credito Finanziario SpA	Italy	Banks
Metro Bank PLC	United Kingdom	Banks
MLP SE	Germany	Capital Markets
Moskovskaya Birzha MMVB-RTS PAO	Russia	Capital Markets
National Bank of Greece SA	Greece	Banks
Natwest Group PLC	United Kingdom	Banks
Nordea Bank Abp	Finland	Banks
OTP Bank Nyrt	Hungary	Banks
Partners Group Holding AG	Switzerland	Capital Markets
Permanent TSB Group Holdings PLC	Ireland; Republic of	Banks
Piraeus Financial Holdings SA	Greece	Banks
Powszechna Kasa Oszczednosci Bank Polski SA	Poland	Banks
Raiffeisen Bank International AG	Austria	Banks
Rathbones Group PLC	United Kingdom	Capital Markets
Ratos AB	Sweden	Capital Markets
Reinet Investments SCA	Luxembourg	Capital Markets
Santander Bank Polska SA	Poland	Banks
Sberbank Rossii PAO	Russia	Banks
Schroders PLC	United Kingdom	Capital Markets
Skandinaviska Enskilda Banken AB	Sweden	Banks
Societe Generale SA	France	Banks
St James's Place PLC	United Kingdom	Capital Markets
Standard Chartered PLC	United Kingdom	Banks
Svenska Handelsbanken AB	Sweden	Banks
Swedbank AB	Sweden	Banks
Sydbank A/S	Denmark	Banks
TP ICAP Group PLC	United Kingdom	Capital Markets
UBS Group AG	Switzerland	Capital Markets
UniCredit SpA	Italy	Banks
Valiant Holding AG	Switzerland	Banks
Virgin Money UK PLC	United Kingdom	Banks

Source: Own research

Appendix B: Correlation matrix

Table 7: Correlation matrix

	ESG	Gender	Board size	Bank size	ROE	ROA	CSR	Skills	Debt
ESG	*1*								
Gender	*0.3948*	*1*							
Board size	*0.3802*	*0.169*	*1*						
Bank size	*0.6436*	*0.2513*	*0.5441*	*1*					
ROE	*0.0085*	*0.086*	*0.0091*	*-0.0296*	*1*				
ROA	*-0.1625*	*0.0238*	*-0.1937*	*-0.4467*	*0.5139*	*1*			
CSR	*0.5496*	*0.2601*	*0.2599*	*0.4585*	*-0.0322*	*-0.1602*	*1*		
Skills	*0.0619*	*0.0406*	*-0.1984*	*-0.0301*	*0.0488*	*0.0483*	*0.0915*	*1*	
Debt	*0.1228*	*0.1262*	*0.0737*	*0.3044*	*-0.0178*	*-0.1754*	*0.0688*	*0.006*	*1*

Source: Own research

Appendix C: Stata code for regression models, correlation matrix and deciding between difference and system GMM (including fixed effects model and OLS)

```
*******************************************************Use "BA_dataset_regression"
*********************************** How dataset was prepared (do not run again)
reshape long yr,i(n_id v_id) j(year)
br
drop variable
reshape wide yr,i(n_id year) j(v_id)
move name year
rename yr1 skills
rename yr2 gender
rename yr3 boardsize
rename yr4 csr
rename yr5 esg
rename yr6 gdp
rename yr7 ceo
rename yr8 indep
rename yr9 roe
rename yr10 roa
rename yr11 assets
rename yr12 debt
rename yr13 critical
tab year, gen(dum_)

gen log_assets=log(assets)
gen log_boardsize=log(boardsize)

*********************************************** Test if dataset is balanced
xtset n_id year

**************************************************** Regression in Stata
*********** Upper regression is twostep difference GMM, lower is system GMM

******** First regression model
*** Onestep GMM
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets
log_assets roe l1.roe roa l1.roa csr skills debt l1.debt  y*, gmm(l1.esg l1.gender
l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y*  debt csr
skills log_boardsize roe ) noleveleq nodiffsargan robust small

xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets
log_assets roe l1.roe roa l1.roa csr skills debt l1.debt y*, gmm(l1.esg l1.gender
l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv( gender y* roe skills csr
log_boardsize debt, equation(level) ) nodiffsargan robust small
```

```
*** Twostep GMM
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt  y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y*  debt csr skills log_boardsize roe ) noleveleq nodiffsargan robust small twostep

xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv( gender y* roe skills csr log_boardsize debt, equation(level) ) nodiffsargan robust small twostep

******** Second regression model
*** With critical dummy
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt critical y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y*  debt csr roa skills log_boardsize roe critical) noleveleq nodiffsargan robust small twostep

xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt critical y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv( gender y* skills roe csr log_boardsize critical debt, equation(level) ) nodiffsargan robust small twostep

******** Third regression model
*** With CEO dummy
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt ceo y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.roa l1.debt, collapse) iv(gender y*  ceo csr skills log_boardsize roe debt) noleveleq nodiffsargan robust small twostep

xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt ceo y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv( gender y* roe skills csr log_boardsize debt ceo, equation(level) ) nodiffsargan robust small twostep

******** Fourth regression model
*** With CEO and critical dummy
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt critical ceo y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* debt csr skills log_boardsize roe critical ceo) noleveleq nodiffsargan robust small twostep
```

```
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets
log_assets roe l1.roe roa l1.roa csr skills debt l1.debt critical ceo y*, gmm(l1.esg
l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv( gender y*
roe skills csr log_boardsize debt critical ceo, equation(level) ) nodiffsargan robust small
twostep

*****************************************************************
*********************************************Generate correlation matrix
correlate esg gender log_boardsize log_assets roe roa csr skills debt

*****************************************************************
******************************Deciding between difference and system GMM
****Pooled OLS
reg esg gender log_boardsize log_assets roe roa csr skills debt y*, robust

****Fixed effects
xtreg esg gender log_boardsize log_assets roe roa csr skills debt y*, vce(robust) fe
i(n_id)

******** Estimating range for plausible results
*** Upper bound
reg esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets
roe l1.roe roa l1.roa csr skills debt l1.debt  y*, robust

*** Lower bound
xtreg esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets
log_assets roe l1.roe roa l1.roa csr skills debt l1.debt  y*, vce(robust) fe i(n_id)
```

Appendix D: Stata code to find the critical mass of women that is most efficient for enhancing ESG performance

```
*******************************************Use "BA_dataset_criticalmass "
********************************************** Test if dataset is balanced
xtset n_id year

**********************************************Regression critical mass
**********Upper regression is twostep difference GMM, lower is system GMM

******** Without dummy variable for critical mass (general regression)
***Onestep GMM
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets
log_assets roe l1.roe roa l1.roa csr skills debt l1.debt  y*, gmm(l1.esg l1.gender
```

l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* debt csr skills log_boardsize roe) noleveleq nodiffsargan robust small twostep

***Twostep GMM
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* roe skills csr log_boardsize debt, equation(level)) nodiffsargan robust small twostep

******** Percentage female directors over 30% (critical mass theory)
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt critical y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* debt csr roa skills log_boardsize roe critical) noleveleq nodiffsargan robust small twostep

xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt critical y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* skills roe csr log_boardsize critical debt, equation(level)) nodiffsargan robust small twostep

******** Percentage female directors between 30-50%
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt balanced y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* debt csr roa skills log_boardsize roe balanced) noleveleq nodiffsargan robust small twostep

xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt balanced y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* skills roe csr log_boardsize balanced debt, equation(level)) nodiffsargan robust small twostep

******** Percentage female directors over 50%
xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt morewomen y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* debt csr roa skills log_boardsize roe morewomen) noleveleq nodiffsargan robust small twostep

xtabond2 esg l1.esg gender l1.gender l1.log_boardsize log_boardsize l1.log_assets log_assets roe l1.roe roa l1.roa csr skills debt l1.debt morewomen y*, gmm(l1.esg l1.gender l1.log_boardsize l1.log_assets l1.roe l1.debt l1.roa , collapse) iv(gender y* skills roe csr log_boardsize morewomen debt, equation(level)) nodiffsargan robust small twostep

Appendix E: Stata code for figures showing shape of the effect of gender diversity on the ESG score depending on the gender of the CEO

```
**********************************************Use "BA_dateset_graph_CEO"
***************************************************Test if dataset is balanced
xtset n_id year

***************************************Differences in effects of gender diversity
******** Effect in general
twoway (qfitci esg gender, stdf) (scatter esg gender)

************************************************************************
********Effect for male CEO
twoway (qfitci esgceo genderceo, stdf) (scatter esgceo genderceo)

********Effect for female CEO
twoway (qfitci esgceo_f genderceo_f, stdf) (scatter esgceo_f genderceo_f)
```

Appendix F: Stata code for figure comparing mean scores of ESG measurement and share of female directors between countries

```
****************************************** Use "BA_dataset_graph_countries"
*********************************************** Graph differences by country
collapse (mean) esg gender, by(country)

graph twoway (scatter esg gender in 1/21, sort mlabel(country)) (lfit esg gender in 1/21,
sort mlabel(country))
```