STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic Year 2018-19

Towards Equal Participation in Leadership

An empirical study on the short-term effects of a mandatory gender quota on supervisory boards in Germany

Harini Annadanam (41222)
Aylin Shawkat (41199)

Abstract

With growing pressure for policy makers to address the underrepresentation of women in corporate leadership roles, it is important to analyze the effectiveness of board quotas as a demand-side intervention. This thesis examines the short-term effects of a mandatory gender quota on supervisory boards implemented in Germany in January 2016. We compile a dataset containing data on board composition and firm financial performance for the years 2010 - 2017, using two panel databases and a cross-sectional dataset. We identify the effects of the quota on subsequent female board share, and on firm financial performance, using a difference-in-differences framework. We find that the quota effectively increased the share of women on supervisory boards in the first two years following its introduction. We also find evidence that affected firms pre-emptively increase their female representation on supervisory boards beginning in 2014. Contrary to popular critiques of board gender quotas, we find that the quota has no significant impact on experience of board members and on firm performance. Furthermore, we find no evidence of spillover effects to the share of women on executive boards.

Keywords: Demand and supply of labor, labor discrimination, economics of gender, firm performance, firm employment decisions

JEL: J20, J70, J16, L25, M51

Supervisor: Abhijeet Singh
Date submitted: 13 May 2019
Date examined: 28 May 2019
Discussant: Paula Höhne Tarragona
Examiner: Magnus Johannesson

Acknowledgements

We would like to thank Abhijeet Singh for his guidance and support. His supervision provided the perfect balance of stern critiques, valuable recommendations and uplifting motivation, in addition to increasing our knowledge of idiomatic expressions.

We would like to thank Jan Sauermann for his constructive input, patience and encouragement. We greatly appreciate his generosity both with his time and in connecting us to the wonderful people at the Swedish Institute for Social Research (SOFI), which allowed us to see our research in the context of the gender economics community and discuss our thesis with interesting people in the field.

Furthermore, we wish to thank Pamela Campa for an inspiring conversation about the topic of gender quotas and female representation in leadership.

Lastly, we would like to thank Anna Dreber Almenberg for her positivity and reliability in all matters, and for being an inspiring role model for aspiring female economists.

Contents

List of Figures

List of Tables

1 Introduction

During his presidency, Richard Nixon was found saying "I don't think a woman should be in any government job whatsoever... mainly because they are erratic. And emotional. Men are erratic and emotional, too, but the point is a woman is more likely to be" (New York Times, 2001). This comment is illustrative of the widespread discrimination faced by women striving for leadership positions across the globe, and is not merely limited to the sphere of politics. In the business world, senior leadership positions tend to be dominated by men, leaving women severely underrepresented. Despite making up 58 percent of the American workforce and the majority of university graduates, only 22.5 percent of board seats in Fortune 500 companies are held by women - not to mention CEO seats, of which only 5 percent are occupied by women (Deloitte, 2018).

When women today strive for positions of power, they are often penalized for deviating from the stereotype of the nurturing and compassionate female. Assertive behavior, expressing disagreement and verbal self-promotion to obtain recognition for accomplishments are costly for women and may reduce their chances to advance professionally (Eagly & Carli, 2012). As a consequence, few women manage to climb the career ladder and enter senior leadership positions. This phenomenon has often been captured by the metaphor of the "glass ceiling", which describes the idea that there is a seemingly invisible barrier to women's advancement to leadership levels of the firm. However, this metaphor has been criticized for suggesting that women face only a single obstacle to their advancement at the penultimate stage of their career and that they generally have equal access to entry- and mid-level positions (Eagly & Carli, 2012). In reality, women face many more obstacles along their career trajectory, including wage and promotion penalties for childbearing (Correll, Benard, & Paik, 2007). The"glass ceiling" metaphor also wrongly suggests that the obstacle faced by women when striving for leadership positions is transparent, i.e. invisible. Many women are aware that the odds are stacked against their professional advancement and this may lead them to adjust their ex-ante aspirations accordingly. In this framework, sending a strong signal of equal opportunities through affirmative action policies is likely to have high returns.

In this thesis, we evaluate the short-run effects of a German policy reform that was implemented in January 2016 and introduced a gradual 30 percent gender quota on supervisory boards of large stock-listed companies. The *Law on Equal Participation of Women and Men in Leadership Positions in the Private and Public Sector* contains two components: the first is the mandatory 30 percent quota on supervisory boards and affects 106 firms; the second is a voluntary measure that requires a further 3,500 firms to announce targets for the gender balance of their management and supervisory boards. We use a difference-in-differences framework to estimate the effect of the mandatory gender quota, making use of the exogenous variation in time across policy implementation and the cross-sectional variation in the eligibility of firms for the mandatory quota. We collect data from three distinct sources and compile a panel dataset that includes information

on 335 firms, of which 89 firms are affected by the mandatory quota, over the period 2010-2017. We compile our dataset to include only firms that are affected either by the mandatory 30 percent quota or by the voluntary target measure. As a result, our control group includes firms affected by the voluntary target measure. Therefore, the results of our difference-in-differences analysis estimate the incremental effects of the mandatory quota policy relative to the baseline of the voluntary measure.

The German law is by no means the first of its kind: starting with a 40 percent gender quota for corporate boards in Norway, the past 15 years have seen a number of European countries implement gender quotas as an affirmative action policy to increase female participation in both government and business.[1] In economic theory, quotas are considered a distortion from the efficient market outcome. However, proponents of quotas argue that quotas actually *correct* inefficiencies that arise from the demand-side issue of discrimination in the hiring process, and the supply-side issue of the lower rate at which high-ability women select into contests (Reuben, Sapienza, & Zingales, 2014; Beaurain & Masclet, 2016). Ford and Pande (2011) present the argument that quotas can correct beliefs about female board members and reduce both inaccurate statistical discrimination and taste discrimination through increased exposure to female directors. Furthermore, gender quotas might have some spillover effects: they might improve the aspirations of younger women and provide increased incentives for them to invest in their careers.

We, therefore, evaluate the effects of the gender quota in Germany along four dimensions. First, we analyze the effectiveness of the quota in raising the share of women on supervisory boards in the first two years after coming into effect. Second, we assess how board characteristics have changed in firms subject to the quota. Third, we analyze the impact of the quota on firm performance – addressing the popular argument that quotas lead to lower average qualifications on boards and hence place affected firms at an economic disadvantage. Finally, we analyze whether the supervisory board quota had any spillover effects to other leadership bodies, such as the executive board.

We find that the quota successfully increased female representation on supervisory boards. Moreover, firms exhibit significant anticipation effects in the course of the public debate around the gender quota. We do not find any statistically significant impact on firm performance - neither in operating revenues nor in firm profitability. Furthermore, we cannot conclude that the higher female board composition had any impact on firms' risk attitudes. Finally, two years after the introduction of the quota, we find no evidence of spillover effects to the executive board.

Methodologically, our paper is closest to Matsa and Miller (2013) who study the impact of the Norwegian quota reform on corporate policy decisions and firm performance, using data on public and private firms in Norway and other Scandinavian countries. They

[1]Norway (2004), Belgium, France, Italy (all 2011), Germany (2015) and Austria (2017) all passed mandatory board quotas for corporate boards, while Spain (2007), Iceland (2010) and the Netherlands (2011) all passed target quotas without penalties in case of non-compliance.

exploit the natural experiment character of the Norwegian quota which exogenously increased the female board share on the boards of public companies in Norway but left other companies in their sample unaffected. Their difference-in-differences identification strategy compares the difference in outcomes between treated and control firms in the years prior to and following the implementation of the quota – similar to our identification strategy. In their first difference-in-differences regression, Matsa and Miller compare the treatment firms (i.e. public companies) to private Norwegian companies, thus holding constant the country and regulatory environment. They then compare treatment companies in Norway to public companies in other Scandinavian countries, to hold constant the type of corporation and potentially allow for greater comparability in governance styles. Finally, the authors combine these two approaches in a triple-difference-in-differences model to capture the change in difference between public and private companies in Norway compared to other countries before and after the quota.

They find that treatment firms undertook fewer workforce reductions, resulting in lower short-term profits due to higher labor costs. No effect was found on non-labor costs or revenues. Matsa and Miller attribute these effects to increased female representation since they find that other board attributes, such as average board member age or CEO experience, remain unaffected by the quota. While women have less CEO experience on average, the new appointees following the quota tend to replace younger and less experienced men on the board, leaving overall board characteristics unchanged. However, Eckbo, Nygaard and Thorburn (2016) fail to reject a value-neutral impact of the quota on firm performance, while Ahern and Dittmar (2012) find that firm performance (measured in Tobin's Q) and board experience dropped following the Norwegian quota. Overall, the literature on the gender quota for corporate boards in Norway yields divided results, which only adds relevance to the question of direction and magnitude of gender quota effects in Germany.

Our study contributes to the existing literature in important ways; while the effects of board quotas have been extensively analyzed in the context of Norway, it is difficult to generalize the results to Germany because of three key differences. Firstly, Germany has a special corporate governance structure with a two-tier board system that divides power between an executive board and a supervisory board. The 2016 German board quota focuses attention on the supervisory board, making it difficult to extrapolate the results of studies into the effect of gender quotas in countries with a one-tier board system. Secondly, the prevailing gender attitudes in Germany are very different from those in Norway: in the European Gender Equality Index, Germany ranks lower than the EU average while comparable Nordic countries, such as Sweden and Denmark, rank much higher. Prevailing gender attitudes define the climate within which the quota is implemented and, as such, determine its reception and potential effects. Lastly, the design of the quota in Germany differs from the Norwegian quota in an important way: while Norway imposed a hard deadline for reaching the 40 percent board representation of the underrepresented gender, the German quota has a gradual deadline by allowing firms to comply with the mandated 30 percent board share on a rolling basis, depending on

their scheduled election and appointment dates. Furthermore, election outcomes that do not comply with the quota law are punished with an "empty seat", meaning that the chair that should have been filled with the underrepresented gender must remain empty. In Norway, on the other hand, consequences were much harsher as non-compliant firms faced liquidation.

Beyond the fact that the results of quota studies in other countries cannot be extrapolated to the German setting for the reasons outlined above, we believe that the German policy warrants individual examination. While researchers have examined the likelihood of women on supervisory boards holding chairperson positions following the quota legislation (Bozhinov, Koch, & Schank, 2018), the link between the quota and firm performance has not yet been investigated, and nor has the causal impact of the quota in driving up female board share. With this thesis, we directly contribute to a deeper understanding of the effects of a major affirmative action policy in the world's fourth-largest economy.

The push for gender equality is one of the defining struggles of our time. Gender equality means that women's and men's rights, responsibilities and opportunities will not depend on whether they are born male or female. At its very core, it is a matter of fairness and equal opportunity. Holding education constant, women around the world face more obstacles and headwind in their career trajectory than their male peer group. Moreover, gender equality is also a matter of representation; if decisions about the design of society are primarily taken by representatives of 50 percent of society, policies are automatically skewed. Evidence for this can be found in Chattopadhyay and Duflo (2004), who use political reservations in India to study the impact of female representation on policy decisions. They show that leaders invest more in public goods that are directly relevant to the needs of their own genders. Extending this argument to the corporate leadership level, this implies that corporate governance policies are likely to be biased towards one gender. They may consequently place too little emphasis on issues that are more relevant to women, such as maternity leave, and fail to establish working conditions that reduce women's obstacles to corporate advancement. Lastly, as educational attainment of the genders has converged, recruiting decision-makers from only half of the possible candidate pool is, at its core, inefficient. The push for gender equality does not just benefit women, it benefits societies at large. It is estimated that if every country matched the progress toward gender parity of its fastest-moving neighbor, global GDP could increase by up to $12 trillion by 2025 (McKinsey, 2015).

The remainder of the paper is organized as follows: Section 2 lays out the policy framework in more detail. Section 3 describes the compilation of the data set, potential data limitations, and presents summary statistics. In section 4, we introduce our empirical specification, for which the corresponding results can be found in section 5. A discussion of our results is presented section 6. Section 7 concludes.

2 Policy Framework

2.1 Background

Germany has a history of promoting traditional marriage and the single-earner (read: male breadwinner) model through various policies introduced in the 1950s and 1960s, many of which continue to exist today. The joint taxation of married couples creates tax benefits for partnerships with a high degree of income inequality, providing the second earner with a disincentive to work. Moreover, non-working dependents are covered by the bread winner's health insurance. Furthermore, a law introduced in 1958, making the wife's right to work dependent on the compatibility of her employment decision with domestic duties, was only replaced in 1977 (Sprengholz, Wieber, & Holst, 2019). Since then, policy changes promoting more emancipated family models have been introduced, gender equality has entered public discourse, and women constitute 46.5 percent of the German labor force. However, the average current director on a corporate board in Germany grew up in a time when the notion of the male breadwinner was widespread.

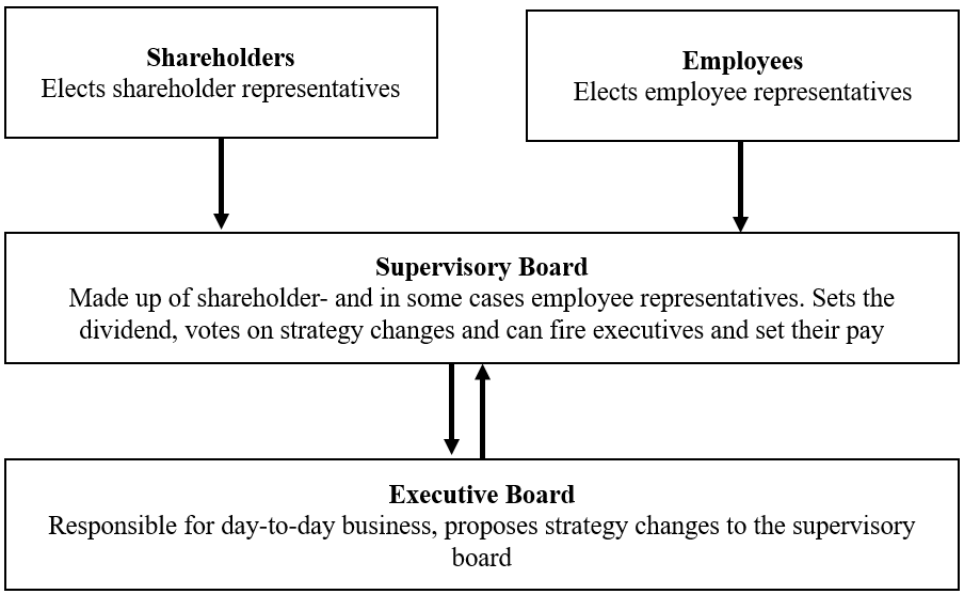
Against this backdrop, it is unsurprising that women in Germany remain severely underrepresented in corporate leadership. Executive boards of publicly traded German companies have a female board share of 8.6 percent and, prior to the introduction of the quota, a female share of only 18.6 percent on supervisory boards (Holst & Wrohlich, 2019). Germany currently has a gender gap of 21 percent in average gross hourly wage, placing it among the widest gaps in Europe. Adjusting this figure for differences in employment choice still leaves an adjusted gender wage gap of 6 percent (Federal Statistical Office of Germany, 2019). This is at odds with the disappearing gender gap in educational attainment, with women between 25 and 34 even outnumbering men in the completion of university studies. Counter to the implied invisibility of the glass ceiling, three quarters of women in Germany think that the labor market is unfair (German Federal Government, 2015).

The introduction of the *Law on Equal Participation of Women and Men in Leadership Positions in the Private and Public Sector* in March 2015 had several objectives: first, it intends to trigger a cultural change in the labor market with a strong signaling effect for aspiring women. It further aims at promoting the advancement of existing female candidates for board positions. Lastly, it aims to increase the representation of female employees at the leadership level: in case of the 106 largest German companies affected by the quota, the size of the corresponding workforce is significant - Volkswagen alone employs over 600,000 people. A quota should ensure that decisions about wages and employment conditions of this workforce are no longer made by an unrepresentative group of executive and supervisory directors (German Federal Government, 2015).

2.2 German Corporate Governance Structure

Before going into the details of the quota law, we need to lay out an idiosyncrasy of German corporate governance: its dual-board structure. In Germany, power is divided between the executive board and the supervisory board. Supervisory boards consist of shareholder and employee representatives and its members are elected by shareholders and employees respectively. This is designed to ensure that both shareholders and employees have an active say in corporate decision making. The supervisory board is responsible for monitoring the executive board, voting on strategy changes, and determining the dividend and executive pay. Furthermore, supervisory board members have the right to appoint and fire executives. Hierarchically, it is deemed superior to the executive board. The executive board is responsible for day-to-day business and strategy and proposes strategy changes to the supervisory board. Figure 1 illustrates the relationship between the two boards.

Figure 1: Germany's Two-Tier Board System



Source: Authors' creation

One important characteristic of supervisory boards is the relative number of seats held by shareholder representatives and employee representatives. This balance is determined by several laws based on the number of employees in a firm.[2] According to the laws, companies with more than 2000 employees are *parity codetermined*, meaning that they must have an equal number of employee and shareholder representatives on their supervisory boards. Companies with more than 500 employees but less than 2000 employees have *third codetermination*, meaning that the supervisory board consists of 33 percent employee representatives. Companies with fewer than 500 employees are not subject to codetermination on their supervisory board and hence consist only of shareholder rep-

[2]Specifically, see *Mitbestimmungsgesetz*, *Montan-Mitbestimmungsgesetz*, *Mitbestimmungsergänzungsgesetz* and *Drittelbeteiligungsgesetz*.

resentatives. The degree of codetermination - and, thereby, the number of employees - is one of the two sources of cross-sectional variation that determines whether a firm is subject to the supervisory board gender quota.

2.3 The Quota Legislation

The *Law on Equal Participation of Women and Men in Leadership Positions in the Private and Public Sector* affects German corporations in two ways: part one of the law imposes a mandatory and fixed quota of 30 percent for the underrepresented gender on stock-listed companies, with the legal form of *Aktiengesellschaft*[3] (AG), *Kommanditgesellschaft auf Aktien*[4] (KGaA) or *Société européenne*[5] (SE), and with more than 2,000 employees. Companies have to comply by working towards this quota whenever a new supervisory board member is appointed. Respective companies are asked to determine the status quo of the share of the underrepresented gender, identify terms and scheduled end-of-office of current supervisory board members, and actively begin searching for candidates of the underrepresented gender in a timely manner. 106 German companies currently fall under this law, corresponding to the top left box in Figure 2.

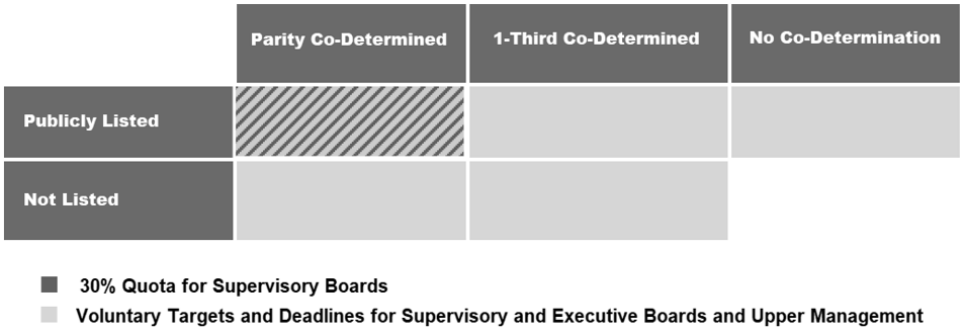
Part two of the law requires companies that are either publicly listed *or* that have at least 500 employees to set themselves target figures for the proportion of women on the supervisory board, the executive board or management level, and the two executive levels below board or management level. This law affects companies of the legal form AG, SE or German companies with limited liability (GmbH). In case the existing share of women falls below 30 percent, the target figure may not be lower than the status quo. However, in case of no female representation, zero-targets are permitted. Together with the target figures, companies must furthermore set themselves compliance deadlines. Non-complying companies do not face a penalty. This part of the law affects around 3,500 German companies, corresponding to the grey boxes in Figure 2. All remaining firms in Germany are exempt from the law, and are not required to comply with the mandatory quota nor to submit voluntary targets. The unshaded box in Figure 2 denotes these firms that are not listed and not codetermined, and are exempt from the quota law.

2.4 Policy Timeline

The passing of the quota law did not come as a surprise, in fact it was preceded by years of failed attempts and public debate. Gender quotas for supervisory boards entered public debate in Germany in 2008, when a number of German women's networks jointly demanded the introduction of a mandatory supervisory board quota of 40 percent in what was called the *Nürnberger Resolution*. They were unsuccessful. In March 2010, Deutsche Telekom voluntarily committed to having 30 percent of all middle and upper management positions held by women by 2015. Several other large German companies

[3]The U.S. equivalent is *Incorporated*.
[4]The U.S. equivalent is *Master Limited Partnership*.
[5]This is a public company registered in accordance with the corporate law of the European Union.

Figure 2: Treatment Status by Listed Status and Degree of Co-Determination



Source: Authors' creation

in the private sector such as BMW, Daimler, E.ON, Bosch and Airbus followed suit and committed to increasing female representation at the corporate leadership level to between 15 and 22 percent. In 2011, Germany's foreign minister first proposed a mandatory quota of 30 percent for supervisory boards, which was subsequently rejected in Cabinet by Chancellor Angela Merkel. The gender quota re-entered public debate in 2013 as part of coalition negotiations between the Social Democrats (SPD) and Christian Democrats (CDU) following the federal elections. A law mandating 30 percent female representation in German supervisory boards was once more proposed in Cabinet in December 2014 and was accepted. The *Law on Equal Participation of Women and Men in Leadership Positions in the Private and Public Sector* was passed by parliament in March 2015. It became binding in January 2016.

3 Data

3.1 Data Collection

We compile a dataset for our analysis by collecting data from three distinct sources. The final dataset used for analysis comprises a panel of 335 German firms over the period 2010 – 2017, of which 89 firms are affected by the quota. Descriptions of the individual sources and our strategy for compiling the dataset are outlined below.

3.1.1 Bureau van Dijk Amadeus

We extract data from Bureau van Dijk's Amadeus database, which contains financial and business information on Europe's largest 520,000 public and private companies by total assets. We selectively extract data on German firms with at least 250 employees in at least one of the three years between 2015 and 2017. The resulting panel dataset contains annual data on financial performance indicators at the firm-level for 16,069 firms over the period 2008 – 2017.

The Amadeus database also contains limited information on board members at the

individual level. For the last year that board-level information was collected for a given firm, the database contains director names, along with personal characteristics, including gender, age and educational background. Unfortunately, the database contains only a current snapshot of board members, giving information on current and historic board members at a specific point in time, without specifying the tenure of historic board members. Previous literature has made use of the archive of Amadeus data to aggregate these snapshots and compile a panel dataset of director-level characteristics (Matsa & Miller, 2013). Since our access is limited to a live subscription, we are unable to collect archived snapshots of Amadeus data. Therefore, we extract a second dataset from Amadeus with the current director information to be used to check the balance of our final dataset, but are unable to use the director-level data for our analysis.

3.1.2 BoardEx

The BoardEx database contains extensive information on the boards of publicly listed and selected notable private companies in all regions of the world. It contains more than 1.8 million profiles of public, private and not-for-profit organizations and more than 1.2 million people. The companies analyzed are selected by a number of criteria, but the majority are chosen due to being publicly-listed companies. The remainder are large private companies, the selection of which is mostly driven by client requests. All data are gathered from publicly available sources, so the companies analyzed and the level of detail may be limited by the amount of disclosure for a given company.

The BoardEx database contains several datasets. We first filter for German companies, then extract and combine two datasets. The first is a panel dataset at the level of individual board members, containing information on their personal characteristics and employment details including the company associated with the individual. The second is a panel dataset at the level of individual board members that contains detailed information on the company and characteristics of the board, individual remuneration, and other boards sat on. We combine the two datasets using the individual identifier specific to BoardEx. The result is a panel dataset at the individual board member level, with information on the company on whose board the individual is sitting, with annual observations over the period 2008 - 2018. The dataset contains information on 8,773 distinct board members, serving on boards of 579 distinct firms. Importantly, this dataset contains information on both supervisory and executive board directors, with variables including gender and age of the individual.

3.1.3 FidAR

We use data from "Frauen in die Aufsichtsräte" (FidAR), an independent Berlin-based women's advocacy group that aims to increase the proportion of women on German supervisory boards. FidAR provides a dataset containing a complete list of companies affected by the mandatory supervisory board quota. We use this list of companies to assign treatment status to firms in our final dataset.

3.1.4 Data Compilation

We combine the three distinct datasets described above for our analysis. A major problem is posed by the absence of a common unique identifier variable across the datasets. Each dataset contains unique identifiers that are specific to that data provider. As a result, we merge data using a two-pronged approach.

The BoardEx dataset contains an International Securities Identification Number (ISIN) for the majority of firms (443 firms). An ISIN number identifier is a 12-character alphanumeric code that serves for uniform identification of a security at trading and settlement. A given firm may have multiple ISIN codes if the firm has multiple traded securities. However, any given ISIN is unique to the firm. As a result, we can use ISIN codes to identify firms in our dataset. The Amadeus dataset also includes ISIN codes for 428 firms. We merge the two datasets using ISIN number, resulting in 284 merged firms. We then sense-check this process; we take the set of firms with an ISIN code that remain unmatched from BoardEx and manually search for these firms in the Amadeus dataset using company name to ensure that firms do not remain unmatched due to any mismatch in ISIN code.

Our second method is to merge firms by hand. We take the remaining firms from each dataset that do not have associated ISIN codes and manually search for firms from the BoardEx dataset in the Amadeus dataset using company name. Since the company names are in different formats and include varying amounts of detail in the text, we are unable to merge firms in Stata using company name. However, we are able to visually identify firms with matching company names and merge these firms by hand, resulting in a further 63 merged firms.

Finally, we assign a dummy variable denoting treatment status to firms using the FidAR dataset. The FidAR dataset contains a list of company names for the 106 firms that were affected by the mandatory quota. We create a treatment dummy variable with a value equal to one for these firms, use the Google Search engine to find a corresponding ISIN code for each of these firms, and merge the treatment dummy variable to the dataset containing Amadeus and BoardEx data using the ISIN code. We sense check any unmerged firms from the FidAR list, to ensure that we did not lose any firms by assigning an incorrect ISIN code. We are able to merge 89 treated firms, losing 17 firms due to limitations in the Amadeus and BoardEx data. The resulting dataset is a panel that contains annual data on 347 firms over the period 2008 - 2017, with 89 of these firms affected by the quota. We will hereafter refer to this dataset as the 'ABF dataset'.

3.1.5 Final Dataset

Following the above process, we separate the ABF dataset into a further two datasets. We first remove all non-supervisory board directors, and then collapse data on director-level characteristics to the firm-level, taking the mean values of director characteristics, such as age and a dummy variable for being female. We make one further restriction to

the data. Data from Amadeus include variables for number of employees and publicly-listed status. These variables allow us to drop firms that are not publicly-listed and have fewer than 500 employees – firms that would not meet the criteria to face the voluntary target measure. Excluding these firms allows us to estimate the impact of the mandatory quota relative to firms that were not affected by the mandatory measure but were affected by a voluntary measure. Finally, we drop the years 2008 and 2009, restricting our sample to the years 2010 - 2017. We do this in order to circumvent any problems that would be likely to arise from using 2008 as a base year, given that the values of firm financial outcomes are likely to be unrepresentative during the financial crisis. Our final dataset is a panel consisting of annual data on 335 firms between 2010 - 2017, of which 89 firms are firms affected by the quota. We call this dataset the 'ABF Supervisory dataset'.

We then repeat the above process, this time starting with the ABF dataset and removing all non-executive board directors. We call the final dataset the 'ABF Executive dataset'. Similarly, this dataset contains panel data at the firm-level on 335 firms over the period 2010-2017.

3.2 Data Limitations

Our study is limited by a lack of readily available data on German firms. We have tempered this problem by collecting and compiling information from a range of sources to create a dataset ready for analysis. However, this approach faces some limitations.

Firstly, our primary data source for board characteristics is the BoardEx dataset. The sampling method used by BoardEx involves collecting board information on all public firms, though this can be limited in some areas by the amount of disclosure, and notable private companies, with selection driven by client requests. This results in a dataset with few private firms relative to public.

As a result, this constrains our ability to use large private firms as a control group for the large public firms affected by the quota. Furthermore, we are concerned about the possibility of differential selection of firms into our matched dataset, due to the non-random sampling process used by BoardEx. We check the extent of this using a balance table to compare the characteristics of firms included in our matched set with those excluded, but present in the Amadeus dataset.

Another limitation stems from the fact that the BoardEx database contains multiple missing values for director characteristics. For instance, educational qualifications and remuneration information are missing for many directors. As a result, we are unable to use such characteristics in our regression analysis, and are unable to assess the impact of the quota on certain outcomes, such as the earnings of women on boards.

3.3 Summary Statistics

Table 1 presents the variables used for our analysis, including variables in both the ABF Supervisory and ABF Executive datasets. The table provides definitions and information on the original data source for the variable.

Using our ABF Supervisory dataset, we present summary statistics at the firm-level for the pre-treatment year 2013 in Table 2. We choose the year 2013, since choosing a year sufficiently in advance of the quota implementation allows us to present pre-treatment characteristics of firms with a high degree of confidence that these statistics are not affected by anticipation of the quota. We compare treatment firms with untreated firms, using the characteristics of supervisory board members only. Each table compares the group of treated firms with a different group of untreated firms. The treated group refers to firms affected by the mandatory quota legislation, while the untreated group comprises firms required only to report a voluntary target. We separate the latter group into five subgroups, according to the segments presented in Figure 2. The first column in the table presents summary statistics for the group of treated firms. The second row includes the group of publicly-listed and parity codetermined firms that are not affected by the quota; the third includes firms that are publicly-listed and not codetermined; the fourth, firms that are not publicly-listed but are parity codetermined; the final column presents firms that are not publicly-listed but third codetermined. This segmentation allows us to more clearly assess the similarity of treatment and control firms.

Table 1: List of Variables and Sources

Variable	Source Data	Description	Units
Outcome Variables			
Female Supervisory Board Share	BoardEx	Number of female supervisory board directors divided by total number of supervisory board directors	Proportion
Operating Revenue	Amadeus	Revenue generated from the company's primary business activities	Billion Euros
Log Operating Revenue	Amadeus	Logged Operating Revenue	Log Units
Profit Margin	Amadeus	Cents of profit generated by the company per Euro of revenue	Percentage Points
Return on Assets	Amadeus	A financial ratio indicating the profitability of the company relative to overall resources. It is calculated as net income divided by total assets.	Percentage Points
Liquidity Ratio	Amadeus	A measure of the company's ability to pay off current debt obligations without raising external capital. It is calculated as current assets that can be quickly liquidated divided by current liabilities.	Percentage Points
Age	BoardEx	Average age of supervisory board members	Years
Network Size	BoardEx	The number of individuals to which the average supervisory board director is connected through overlaps in start and end dates of all roles, including employment, education, or other activities	Number of connected individuals
Female Executive Board Share	BoardEx	Number of female executive board directors divided by total number of executive board directors	Proportion
Controls			
Number of Supervisory Board Directors	BoardEx	Total number of directors on supervisory board	Number of individuals
Number of Executive Board Directors	BoardEx	Total number of directors on executive board	Number of individuals
Treated	FidAR	Dummy variable taking the value of one for a company affected by the mandatory quota	Binary dummy taking the value one or zero
Year Dummies	NA	Dummy variable taking the value of one in a specific year	Binary dummy taking the value one or zero
Post-Treated	NA	Dummy variable taking the value of one in the post-quota period (2016 and 2017)	
Other Variables			
BoardID	BoardEx	Firm identifier	NA
Number of Employees	Amadeus	Number of employees employed by the company globally	Number of individuals
Publicly Quoted	Amadeus	Dummy variable taking the value of one if the company is listed on a stock exchange	Binary dummy taking the value one or zero
Natural Resources	Amadeus	Dummy variable taking the value of one if the company is in the natural resources sector	Binary dummy taking the value one or zero
Manufacturing	Amadeus	Dummy variable taking the value of one if the company is in the manufacturing sector	Binary dummy taking the value one or zero
Services	Amadeus	Dummy variable taking the value of one if the company is in the services sector	Binary dummy taking the value one or zero
Intellectual Services	Amadeus	Dummy variable taking the value of one if the company is in the intellectual services sector	Binary dummy taking the value one or zero

We can see that firms had significantly different shares of women on boards by treatment status, with firms eligible for the quota having an average pre-treatment female board share of 17.9 percent, while ineligible firms had a much lower pre-treatment average, across groups. The difference between means in female board share for treated and

Table 2: Pre-treatment Summary Statistics (2013)

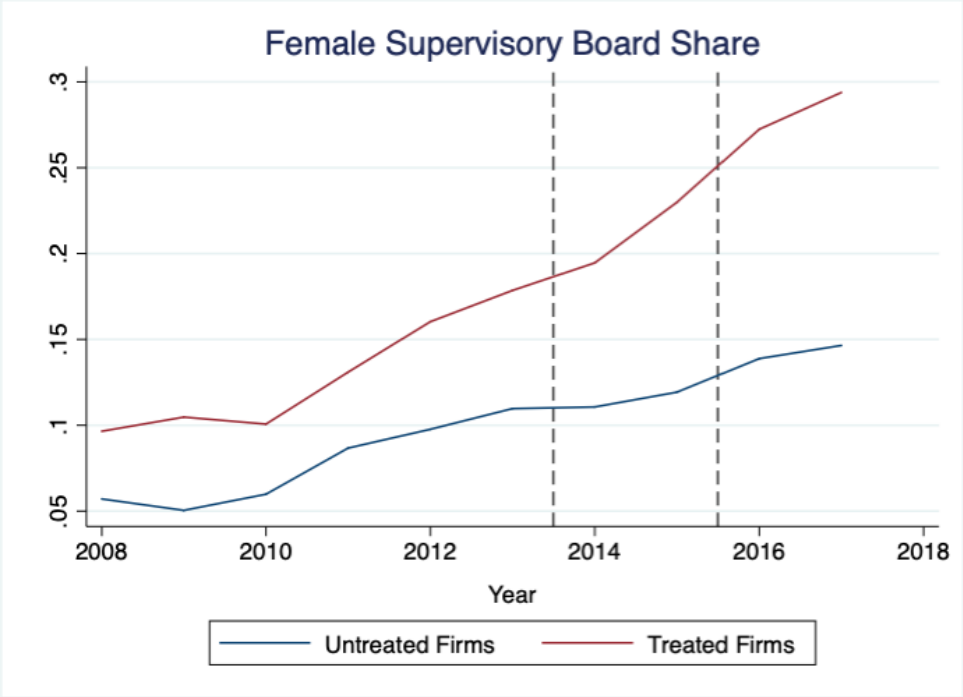
		Treated	Listed Parity	Listed Third	Listed Not Codetermined	Unlisted Parity	Unlisted Third
Female Supervisory Board Share	Mean	0.179	0.122***	0.076***	0.106**	0.144	0.073**
	Std. Dev.	0.098	0.138	0.117	0.244	0.107	0.101
	N	79	45	53	22	27	5
Age	Mean	62.0	63.7**	63.3	62.8	64.2**	64.0
	Std. Dev.	3.9	5.1	6.8	7.1	4.8	8.3
	N	79	45	53	20	27	5
Number of Employees	Mean	60173	9748***	925***	414**	34326	1127
	Std. Dev.	109570	25893	468	288	58399	574
	N	78	45	51	22	26	4
Manufacturing	Mean	0.46	0.16***	0.32	0.18**	0.26*	0.80
	Std. Dev.	0.50	0.37	0.47	0.39	0.45	0.45
	N	79	45	53	22	27	5
Natural Resources	Mean	0.03	0.02	0.02	0.00	0.00	0.00
	Std. Dev.	0.16	0.15	0.14	0.00	0.00	0.00
	N	79	45	53	22	27	5
Services	Mean	0.15	0.07	0.06*	0.09	0.15	0.00
	Std. Dev.	0.36	0.25	0.23	0.29	0.36	0.00
	N	79	45	53	22	27	5
Intellectual Services	Mean	0.37	0.76***	0.60***	0.73***	0.59**	0.20
	Std. Dev.	0.49	0.43	0.49	0.46	0.50	0.45
	N	79	45	53	22	27	5

untreated groups is only insignificant for the unlisted, but parity codetermined group; all other groups have a difference that is significant at the 5 percent level at least.

Average director age is significantly different between the treatment group and two of the untreated groups, but the difference is small, and we believe has no economic significance. We are more interested in the pre-treatment difference in number of employees. We find that the mean number of employees is much higher for treated firms than for untreated firms in 2013, across all groups of untreated firms. However, the difference is only statistically significant for the groups of publicly-listed firms. The number of employees in a company is likely to be strongly correlated with several of our outcome variables for a given firm. For instance, it is plausible that the number of employees would correlate with financial performance of firms, resulting in differential financial performance across treatment status based on this differential characteristic of treated firms. However, we do not believe that differences in the number of employees would drive differential trends in financial performance over time. As a result, we retain these groups of firms in our control group. It is noteworthy that most of the untreated groups are made up of a higher proportion of firms in the intellectual services sector than the equivalent proportion for treated firms.

Figure 3 displays the trends in female supervisory board share for firms by treatment status. This specific quota legislation was first discussed in the media in late 2014, at which point we might expect firms to begin to anticipate the quota and change their behavior accordingly. We display dashed lines on the figure indicating the beginning of the anticipation period in 2014 and the implementation of the quota in 2016. In line with this notion, the figure implies that trends begin to diverge in 2014, so we investigate the possibility that the treatment had anticipatory effects in advance of implementation in our empirical analysis. See Figure 4 for a counterpart of this figure using our regression coefficients.

Figure 3: Time Trends - Female Supervisory Board Share



Source: Authors' creation

3.4 Balance Checks

In Section 3.1.1, we mention that the Amadeus database contains information on board members at the individual level for the last year that board-level information was collected for a given firm. We extract this information for the full set of firms in the Amadeus database. This allows us to check whether firms in our final matched dataset are significantly different from firms that we are unable to include in our final dataset, due to lack of data on these firms in the BoardEx database. Data on supervisory board directors are available for a single year for each firm, ranging between 2014 and 2017. Restricting the sample to these years, and including each firm only once in the sample, we present summary statistics for the share of female directors in supervisory boards, the average age of supervisory board directors, and the number of employees in the firm, for the two groups of firms - matched firms that are included in our ABF dataset, and unmatched firms that we are unable to include. Examining these characteristics allows us to assess the validity of the concern that non-random selection causes our sample to be unrepresentative of the full German population of firms. Table 3 presents these figures.

We see that firms in the matched group do not have a significantly different share of women on their supervisory boards compared with the unmatched group. Similarly, average age of supervisory board directors is not significantly different across the two groups. Only in the number of employees in these firms do we see a significant difference between the two groups. Firms in the matched group are much larger than firms in the unmatched group, on average. While this does imply some selection on observables into

Table 3: Balance Check - Comparison of Matched and Unmatched Firms

		Matched	Unmatched
Female Supervisory Board Share	Mean	0.149	0.169
	Std. Dev.	(0.141)	(0.181)
	N	188	2342
Director Age	Mean	59.3	58.8
	Std. Dev.	(5.3)	(7.0)
	N	188	2304
Number of Employees	Mean	9797***	1269
	Std. Dev.	(35328)	(6525)
	N	203	12515

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

our sample of firms, we dismiss this as a major concern. Only large firms, with more than 2000 employees, are affected by the quota law. Given this, the apparent selection of larger firms into the matched set is not of great concern, since these firms are likely to be better control firms for the treated firms affected by the mandatory quota law. We are unable to comment on possible selection bias on unobservables, which might make our sample of firms unrepresentative of the population of firms and drive results. However, our balance checks do provide confidence that our dataset does not appear to present significant concern in terms of selection bias on observables, so we use this sample of firms for our analysis with relative complacence.

4 Empirical Specification

4.1 Outcomes

We estimate the effects of the quota policy on four broad categories of outcomes, in order to distinguish between various channels through which the policy might take effect. Firstly, we examine the success of the policy in achieving its primary goal – increasing the share of women on boards – by examining the effect on the gender composition of supervisory boards. This allows us to measure the degree of compliance by firms. Since the quota policy has a soft deadline in terms of permitting firms to take action gradually, measuring its short-term impact on female board share allows us to assess its efficacy relative to hard deadlines, such as the Norwegian reform.

Our second group of outcomes allows us to check for any changes in the characteristics of board members as a result of the policy. We examine how average board member age and network size (which may be seen as a proxy for experience) are affected, using director-level information from the BoardEx database. For each director, the network size variable counts the number of individuals to which the director is connected through overlaps in start and end dates of all roles, including employment, education, or other

activities. If two individuals hold roles at the same organization at the same time, they are connected, with each connection counted once.

The third set of regressions deals with financial performance. The aim is to examine the impact of the quota on the performance of firms to assess the validity of the argument that a quota would negatively affect firm performance, by necessitating the hiring of underqualified women. Specifically, we look at profit margin, operating revenue and return on assets. Moreover, addressing the commonly-held stereotype of women's higher average risk aversion, we examine the quota impact on liquidity ratio, which may be taken as a proxy of a firm's risk attitude.

Finally, we estimate any spillover effects of the policy on increasing the share of women in other areas of the firm, specifically on executive boards. One of the posited effect channels of quota policies that increase the representation of women in positions of power and leadership is to prompt further female-friendly changes within the firm. It may be that the proportion of women in management positions increases as a result of firms' increased awareness of the presence of highly qualified women in the labor market, engendered by their search for supervisory board members. Alternatively, since supervisory board members wield significant clout over the hiring process of executive board members, it may be that new female supervisory board members play a direct role in increasing representation of women in these positions, either through recommending women for the role or through pushing for more female-friendly policies within the firm, which then attract more women.

4.2 Methodology

We estimate the effect of the policy using a difference-in-differences (DD) framework. Our identification comes from the fact that the mandatory quota policy did not apply to all German firms, but rather to a subset of large German companies. We make use of the cross-sectional variation in quota eligibility to compare the change over time in the outcomes of treated German firms with the change over time of a control group of untreated German firms. The DD approach allows us to isolate the effect of the policy. Since we would expect the outcome variables to be changing over time due to industry-level or national-level trends, a solely time-series approach would likely conflate the effect of the policy with changes that would have happened regardless of the policy. Furthermore, since assignment to treatment is non-random, a purely cross-sectional estimate would contain selection bias due to time-independent variation between the treated and untreated groups. The DD strategy allows us to control for this time-independent variation to estimate the true policy effect on treated firms compared to control firms.

Our empirical strategy allows us to determine the intent-to-treat (ITT) effect of the policy. That is, we find the effect of the policy on the group of firms who are eligible to be affected by the quota policy. Furthermore, intention to treat and receipt of treatment are synonymous in this context, since a firm that is eligible for the quota is necessarily subject to it. The ITT is the same then as the average treatment effect on the treated

(ATT). However, it is important to note that the policy is not neutral with respect to the control group. The quota legislation mandates that a certain group of firms are affected by the mandatory 30 percent quota, while a second, larger group of firms is affected only by the requirement to announce targets and deadlines for gender balance. The remaining firms in Germany are not affected by any type of gender policy at this time. As a result, the effect we are able to estimate differs with our choice of control group. We restrict our sample of firms to exclude firms that are unaffected by any policy measure; that is, we exclude firms that are not publicly-listed and not codetermined. We do this because these firms are much more dissimilar to treatment firms than firms that are affected by the voluntary measures, and we cannot include firms from both groups since we expect the nudge to firms affected by the voluntary measure to place these firms on a different trend path to unaffected firms after the implementation of the quota law. Therefore, our DD estimation strategy allows us to estimate the incremental effect of the mandatory quota relative to the voluntary target measure. Henceforth, unless specified, our use of the term "untreated" refers to this control group.

We estimate several specifications of the regression model, beginning with the following:

$$Y_{i,t} = \alpha + \beta_1^{DD}\text{Treated}_i \times \text{Post}_t + \gamma\text{Treated}_i + \lambda_t + \epsilon_{i,t}. \quad (1)$$

$Y_{i,t}$ represents the outcome variable and the DD coefficient β_1^{DD} captures the effect of the quota. Treated_i is a dummy variable equal to one for firms affected by the mandatory quota, Post_t is a dummy variable taking the value of one in the post-reform period (that is, in 2016 and 2017) and λ_t represents time fixed effects. Lastly, the time-varying idiosyncratic error term is captured by $\epsilon_{i,t}$. We cluster standard errors at the firm level.

Our use of time fixed effects is important for two reasons. Firstly, it allows us to control for the time trend in outcomes across all firms. We expect that female board share is on an increasing trajectory over time in our sample period due to changing gender attitudes and a push for gender diversity across professional settings. Furthermore, since financial performance of all firms is likely to be influenced by the macroeconomic climate, we control for time to capture this effect. The year fixed effects tell us something more, however. They are interesting in themselves since they show us the trajectory of female board share for both treatment and control firms prior to treatment. In the years following treatment, the year dummies capture both the time trend and the effect of the voluntary target measure on control firms.

We then move to a second specification that replaces the Post_t term with separate dummy variables for the two post-reform years, in order to disentangle any heterogeneity in the effects of the quota over time.

The regression specification is as follows:

$$Y_{i,t} = \alpha + \beta_1^{DD}\text{Treated}_i \times 2016_t + \beta_2^{DD}\text{Treated}_i \times 2017_t + \gamma\text{Treated}_i + \lambda_t + \epsilon_{i,t} \quad (2)$$

We then augment the above by implementing a test for causality along the vein of

Granger (1969), drawing on the approach used by Autor (2003). We continue to include the two lagged years on the right-hand side, 2016 and 2017, to capture post-treatment effects, and augment this with five leads to capture any anticipatory effects of the quota [6]. These terms are now denoted by the summation terms below, where $D_{i,t,2016\pm\tau}$ indicates a dummy variable taking the value of one if a firm is treated in the year $2016 \pm \tau$.

$$Y_{i,t} = \alpha + \Sigma_{\tau=1}^{5}\beta_{-\tau}D_{i,t,2016-\tau} + \Sigma_{\tau=0}^{1}\beta_{\tau}D_{i,t,2016+\tau} + \gamma\text{Treated}_i + \lambda_t + \epsilon_{i,t} \tag{3}$$

Finally, we build on the above specification by including firm fixed effects to remove any bias stemming from time-invariant, firm-specific unobservables. Our DD coefficients would be biased if there were unobservable firm-specific characteristics that affect the outcome variable and are correlated with treatment status. An example of this pertains to the fact that only firms of certain legal forms are affected by the quota law; furthermore, it is unclear whether foreign companies that have their administrative seat in Germany are subject to the law. We might reasonably expect that legal form and origin country are drivers of financial performance. Our data do not include legal form of companies, and since our data comprise firms that operate within Germany and have supervisory boards within Germany, it is unclear whether the data include foreign firms. As a result, we might expect our DD terms to be biased in a regression of financial performance. Controlling for firm fixed effects partials out any such bias.

The following specification displays this addition, represented by the term μ_i.

$$Y_{i,t} = \alpha + \Sigma_{\tau=1}^{5}\beta_{-\tau}D_{i,t,2016-\tau} + \Sigma_{\tau=0}^{1}\beta_{\tau}D_{i,t,2016+\tau} + \lambda_t + \mu_i + \epsilon_{i,t} \tag{4}$$

4.3 Assumptions

A key identification assumption for the validity of our DD approach is that the treatment firms have parallel time trends to the control firms in the absence of treatment. In other words, it requires that, in the absence of treatment, the difference between the treatment and control group is constant over time, conditional on conditioning variables X. The assumption can be written as follows (Lechner, 2010):

$$\begin{aligned} &E(Y_1^0|X = x, D = 1) - E(Y_0^0|X = x, D = 1) = \\ &E(Y_1^0|X = x, D = 0) - E(Y_0^0|X = x, D = 0) = \\ &E(Y_1^0|X = x) - E(Y_0^0|X = x); \forall x \in X \end{aligned} \tag{I}$$

The term $E(Y_1^0|X = x, D = 1)$ refers to the expected value of the outcome variable in the second (post-treatment) period, in the absence of treatment, for the group of firms in the treatment group. This is, of course, a counterfactual term and is not an observed outcome. The term $E(Y_0^0|X = x, D = 1)$ represents the expected value of the outcome

[6] Note that we use the terms "lead" and "lags" to refer to interaction terms between the treatment indicator and year dummies, rather than lead and lagged terms in a time-series sense.

variable in the first (pre-treatment) period, in the absence of treatment, for the group of firms in the treatment group. The term $E(Y_1^0|X = x, D = 0)$ represents the expected outcome in the second period, in the absence of treatment, for firms in the control group. Finally, $E(Y_0^0|X = x, D = 0)$ denotes the expected outcome for firms in the control group in the first period, in the absence of treatment.

Our final regression specification can be used to verify the common trends assumption, since the regression includes leads of the treatment. The leads should not be significantly different from zero to confirm the presence of common trends across the two groups.

In order to identify a mean causal effect of the quota, a further set of assumptions must be satisfied. Two important assumptions for causal inference, in general, are the Stable Unit Treatment Value Assumption (SUTVA) and the Exogeneity Assumption (EXOG). The first implies that one, and only one, of the potential outcomes is observable for every member of the population. It requires that the response of a particular unit depends only on its assigned treatment status, and not on the treatments of others in the population. There is a possibility that this assumption is broken in this context. Since treated firms are large publicly-listed firms, it is plausible that they garner much attention in press and media. Media coverage of their response to the quota legislation might influence the behavior of control firms who, despite being unaffected by the mandatory quota, are nevertheless required to release voluntary targets. These firms may feel more or less pressured to increase their female board share by the visible progress of treatment firms in this direction. However, we do not believe this to be a serious concern due to the nature of the quota. The quota requires that treated firms must hire a female candidate in any board election following the implementation of the legislation in January 2016, until the 30 percent threshold for representation is achieved. As a result, firms are limited to compliance along a binary dimension. Treated firms either do or do not comply by hiring a woman and thus increasing female board share in elections from 2016 onwards. Therefore, there is little scope for differential compliance of treated firms, curtailing the effect of their behavior on control firms.

The second assumption (EXOG) requires that the components of the covariates X are not influenced by the treatment. The assumption is expressed formally as:

$$X^1 = X^0 = X, \forall x \in X. \tag{II}$$

The equation expresses that the values of the conditioning variables in the presence of treatment (X^1) are equivalent to the values of the conditioning variables in the absence of treatment (X^0). We do not use any conditional variables in our analysis, and thus satisfy this assumption by default. Since we use firm fixed effects to remove all time-invariant, firm-specific variation, and year fixed effects to capture variation along the time dimension, the only control variables that might improve our regression model would be firm-specific time-varying characteristics. Our dataset does not include any firm-specific time-varying controls that might hypothetically explain variation in any of

our outcome variables. Therefore, we do not include any conditioning variables in our analysis.

More specific to the strict DD framework is the assumption that the treatment has no effect on the pre-treatment population in the pre-treatment period. This assumption precludes the possibility that the treatment can influence pre-treatment outcome through anticipation effects. While this assumption is required to hold to determine the causal effect of the quota after implementation, we are interested in the dynamic causal effect of the quota on female board shares, including any anticipation effects that prompt firms to hire more female directors before implementation. Therefore, we relax the assumption that treatment has no effect in the pre-treatment period in our analysis.

Finally, some assumptions of the classical linear regression model must continue to hold in the DD framework to ensure unbiasedness of estimates and correct standard errors. We require the zero conditional mean assumption to hold to ensure unbiasedness of estimates. This assumption requires that the idiosyncratic error term has an expected value of zero, conditional on values of the independent variables (Wooldridge, 2013). Furthermore, the strict exogeneity condition must hold, which states that for each year, the expected value of the idiosyncratic error is uncorrelated with the explanatory variables in all years. Lastly, we require that conditional on X, the errors are uncorrelated across years. This is the assumption that there must be no serial correlation of the error terms, and can be written formally as:

$$\text{Corr}(\epsilon_{i,t}, \epsilon_{i,s}) = 0 \forall t \neq s. \tag{III}$$

Since we believe that the error terms will be correlated for a given firm over time, we cluster standard errors at the firm level to deal with potential serial correlation. The zero conditional mean assumption may be broken in the event that there are confounding gender-related policies being implemented within the timeframe of our analysis. For instance, if the government were to introduce a policy that aims to address some gender equality concern in the years immediately before or after the board quota policy studied here, our DD estimates may be biased. This would be the case if the second policy were to affect quota-treated firms differently from control firms. Furthermore, an additional policy of this type would bias the coefficients on the year dummies, which identify the time trend in female board share across all firms. However, while we find evidence of childcare policies implemented in 2013[7], we do not believe that this would bias our regression results here. Childcare policies are primarily targeted at women of or below childbearing age. Since the average board member age in our sample is 61, childcare policies are unlikely to affect the gender balance of boards, but rather to affect younger women in the firm.

[7]The Kinderfoerderungsgesetz, implemented in August 2013, mandates that parents have a legal right to childcare for all children below the age of four.

5 Results

Using the aforementioned regressions, we present and analyze the effects of the mandatory 30 percent gender quota for supervisory boards that came into effect in January 2016. For each of the different channels, we comment on the different regression specifications. Recall that we build up our regression specifications from looking at (1) a simple baseline DD regression, where the DD coefficient captures the difference in the difference between treatment and control groups before and after the introduction of the quota. We then move on to (2), attributing distinct DD terms to each post-introduction year - 2016 and 2017. In (3), we introduce interactions of the treatment dummy with lead years to capture any anticipation effects. Specification (4) introduces firm fixed effects.

5.1 Female Supervisory Board Share

Figure 4 plots the estimated DD coefficients from the results of our fourth specification in Table 4, with their respective confidence intervals. From visual inspection, we see that treatment firms already had a higher growth rate in female board share than firms in the control sample before the introduction of the quota. The DD coefficient is significantly different from zero in 2014, indicating that first significant anticipation effects emerge among treatment firms in 2014. This coincides with the increased public debate about gender quotas in the course of coalition negotiations following the 2013 German federal elections and the passing of the quota in Cabinet in late 2014.

In the presence of these anticipation effects, model specifications (3) and (4) including lead DD terms are more suitable for capturing the dynamics of female board share in treatment firms. Furthermore, the inclusion of lead terms allows us to verify the presence of parallel pre-treatment trends to an extent. In specification (4), we find that the lead DD terms are insignificant prior to 2014, providing some support for the assumption that treated and control firms have parallel trends prior to any anticipation of the quota.

The first specification displays the results of a simple DD regression evaluating the difference between treated and control firms after the quota was implemented in 2016. The results indicate that treated firms experienced an increase in female board share relative to control firms following implementation of the quota. Replacing the post-implementation DD term with year-specific DD terms for the post-implementation period in specification (2), we find that the quota increased female board share by 5.5 percentage points for treated firms compared with control firms over the period 2010 to 2016. Treated firms experience a further increase in 2017, resulting in a 6.5 percentage point increase in female board share compared to control firms, from the baseline in 2010. Upon introducing lead terms, we find that the DD estimates increase, indicating a stronger effect of treatment in the post-treatment period when allowing for anticipation effects. The magnitude of the coefficients on the year dummies are reduced, implying that some effect of the quota was being captured by the time trend in specification (2).

Specification (4) augments specification (3) by adding firm fixed effects. We find that the magnitudes of the estimated DD coefficients are slightly reduced in comparison with specification (3), while the magnitudes of the coefficients on the year dummy variables increases. The implication is that the DD coefficients in specification (3) may have been biased upward by picking up some unobserved time-invariant differences between treated and control firms. The group of treated firms may possess some unobserved characteristic that is correlated with higher female board share, relative to the control group. By controlling for firm fixed effects, we mitigate this bias, and attribute a slightly lower effect to the quota in increasing female board share. This is our preferred specification; we believe that the importance of correcting any bias in our estimates stemming from time-invariant differences between the treated and control groups outweighs the problem of lower efficiency in the estimates due to reduced variation when using fixed effects.

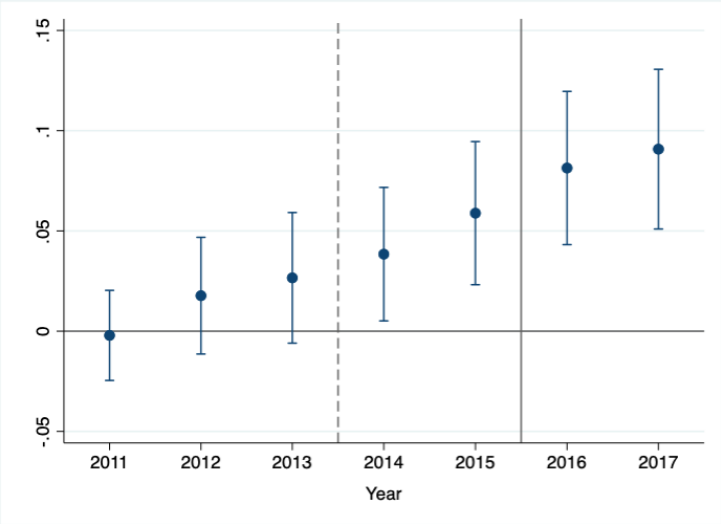
Specification (4) finds a steep adjustment process towards the 30 percent target after the quota became binding in 2016: the average supervisory board increased their share of women by 8.1 percentage points in 2016 with respect to the base year of 2010. In 2017, this effect increased even further to 9.1 percentage points. This treatment effect was further supported by a strong positive trend over time, which significantly drove up female board representation of both control and treatment firms over the full observation period. The coefficients on the year fixed effects indicate that female board share is increasing for all firms over the sample period, with statistically significant positive coefficients in all years. However, growth in female board share seems to increase in 2014, with average female board share increasing by nearly 5 percentage points between 2014 and 2017. This potentially points to the mild success of the voluntary target measure in inducing firms to increase female board share beyond the status quo.

We find that the values of the DD coefficients are significant in each year from 2014 onwards, indicating significant effects of the quota on female board share. Furthermore, the significance of coefficients on the year dummies indicates a significant positive trend in female board share over our sample period. We find that the DD coefficients are jointly significantly different from zero at the 1 percent level. We are unable to reject the null hypothesis that the DD coefficients for 2017 and 2016 are equal at the 10 percent level. We cannot, therefore, conclude that the quota had a differential impact in 2017 relative to 2016.

Overall, these results allow us to conclude that the quota was in fact successful in driving up the share of women on German supervisory boards. The early compliance of treatment firms further points to the interpretation that companies do not perceive an increased female board share as costly. With respect to the availability of qualified women on the labor market, the interpretation is less clear: early compliance could indicate that experienced and qualified women on the labor market are relatively scarce and that early compliers gain an early mover advantage by being able to choose first. On the other hand, it could also mean that there is, in fact, a broad availability of qualified women

and that the public debate increased the salience of qualified women in the candidate pool.

Figure 4: Female Supervisory Board Share



Source: Authors' creation

Table 4: Regression Results - Female Supervisory Board Share

	(1)	(2)	(3)	(4)
Treated x Post-Reform	0.0600***			
	(0.012)			
Treated x 2017		0.0650***	0.0932***	0.0908***
		(0.013)	(0.020)	(0.020)
Treated x 2016		0.0552***	0.0834***	0.0814***
		(0.012)	(0.019)	(0.019)
Treated x 2015			0.0608***	0.0589***
			(0.018)	(0.018)
Treated x 2014			0.0391**	0.0384**
			(0.017)	(0.017)
Treated x 2013			0.0274*	0.0266
			(0.016)	(0.017)
Treated x 2012			0.0189	0.0177
			(0.015)	(0.015)
Treated x 2011			-0.000257	-0.00212
			(0.011)	(0.011)
2017	0.113***	0.111***	0.0983***	0.100***
	(0.012)	(0.012)	(0.015)	(0.016)
2016	0.100***	0.102***	0.0886***	0.0906***
	(0.011)	(0.011)	(0.015)	(0.015)
2015	0.0920***	0.0920***	0.0692***	0.0712***
	(0.010)	(0.010)	(0.015)	(0.015)
2014	0.0718***	0.0718***	0.0555***	0.0562***
	(0.009)	(0.009)	(0.014)	(0.014)
2013	0.0642***	0.0642***	0.0517***	0.0525***
	(0.009)	(0.009)	(0.014)	(0.014)
2012	0.0520***	0.0520***	0.0424***	0.0437***
	(0.008)	(0.008)	(0.011)	(0.011)
2011	0.0339***	0.0339***	0.0307***	0.0319***
	(0.006)	(0.006)	(0.010)	(0.010)
Treated	0.0774***	0.0773***	0.0497***	
	(0.012)	(0.012)	(0.016)	
Constant	0.0401***	0.0402***	0.0529***	0.0672***
	(0.010)	(0.010)	(0.013)	(0.008)
R^2	0.165	0.165	0.170	0.146
Mean Dependent Variable	0.141	0.141	0.141	0.141
Mean Dependent Variable (2010)	0.078	0.078	0.078	0.078
Number of observations	2,009	2,009	2,009	2,009
Number of firms	326	326	326	326
Firm Fixed Effects	No	No	No	Yes

Standard errors are clustered at the firm level and presented in parentheses.

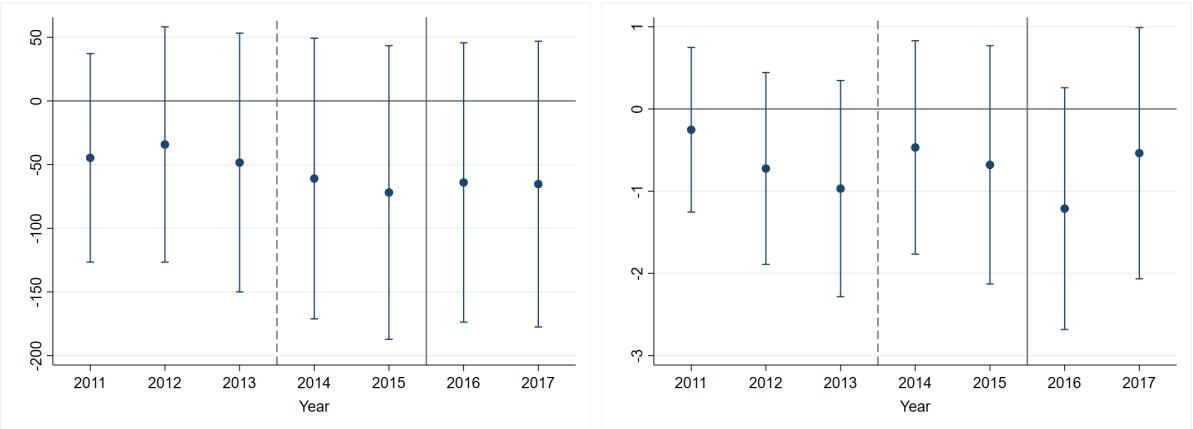
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

! The dependent variable is Female Supervisory Board Share. The unit is a proportion.

5.2 Board Characteristics

Our analysis of the impact of the quota on board characteristics partly addresses the question of availability of qualified women raised in the previous paragraph. One of the main objections to the quota was that the quota would negatively affect the quality of boards, as male board members would have to be replaced with women of less professional experience. We use Director Age and Director Network Size as proxies for experience since the number of professional contacts is correlated with the number of positions previously held. Including DD leads, time fixed effects, firm fixed effects and controlling for board size, we find that the quota had no significant effect on the average director network size on supervisory boards of treatment firms, nor did it lead to a significant decrease in average director age. This lends support to the interpretation that women are, in fact, not less professionally experienced than their male counterparts. This is even true for women of the generation that still saw significant gender gaps in educational attainment and grew up in a time of significantly more traditional gender attitudes.
We find negative year coefficients for the Director Age regression in all years of the observation period, pointing to a trend of appointing younger supervisory board members in all firms of our sample: in 2017, the average board member was six years younger than the average director in our base year of 2010, resulting in an average age of 58.7 years in 2017. The average Network Size on supervisory boards does not undergo any significant changes throughout the observation period. An F-test for the joint significance of the DD coefficients in the Network Size regression indicates that the DD coefficients are not jointly significantly different from zero. However, an equivalent F-test for the DD coefficients in the Director Age regression finds that the DD coefficients are jointly significantly at the 5 percent level, so we cannot rule out the possibility that the quota led to a reduction in average director age for treated firms.

Figure 5: DID Estimates - Female Board Shares and Director Characteristics



(a) Director Network Size

(b) Director Age

Source: Authors' creation

Table 5: Regression Results - Director Characteristics

	Network Size	Age
Treated x 2017	-65.31	-0.538
	(57.046)	(0.777)
Treated x 2016	-64.05	-1.212
	(55.766)	(0.748)
Treated x 2015	-71.91	-0.679
	(58.606)	(0.737)
Treated x 2014	-60.94	-0.468
	(56.022)	(0.660)
Treated x 2013	-48.39	-0.969
	(51.661)	(0.669)
Treated x 2012	-34.22	-0.724
	(46.989)	(0.593)
Treated x 2011	-44.69	-0.253
	(41.625)	(0.509)
2017	55.24	-6.112***
	(52.795)	(0.662)
2016	50.89	-4.870***
	(52.088)	(0.634)
2015	65.32	-4.329***
	(55.514)	(0.622)
2014	68.92	-3.284***
	(53.232)	(0.586)
2013	57.27	-1.958***
	(49.017)	(0.599)
2012	49.65	-1.261**
	(45.040)	(0.527)
2011	51.02	-0.838*
	(40.227)	(0.452)
Constant	375.8***	65.16***
	(30.651)	(0.352)
R^2	0.000	0.132
Mean Dependent Variable	412.6	61.8
Mean Dependent Variable (2010)	440.0	65.0
Number of observations	2,009	1,998
Number of firms	326	326
Firm Fixed Effects	Yes	Yes

Standard errors are clustered at the firm level and presented in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

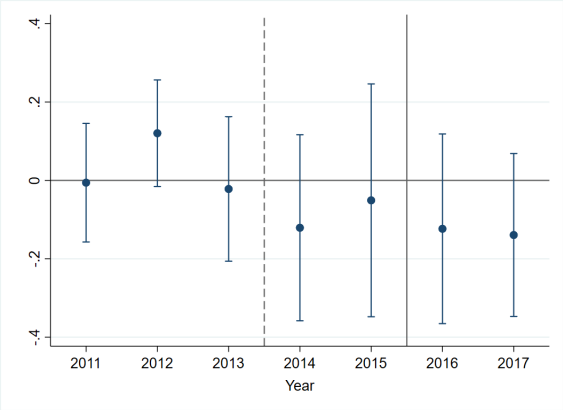
The dependent variables are Network Size and Average Board Member Age, respectively. The units are the number of individual connections and years of age, respectively.

5.3 Financial Performance

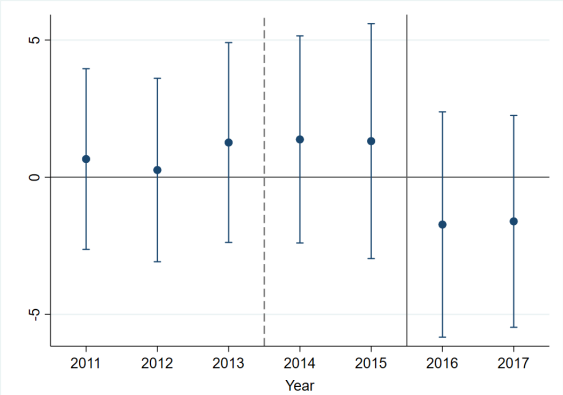
Figure 6 plots the DD coefficients for our four financial outcomes. Recall that Operating Revenue describes the revenue generated from a company's primary business activities. We take the logarithm of Operating Revenue because its distribution is more similar to a log-normal distribution (see Figure B.11). Furthermore, this allows us to interpret our coefficients as percentages. Despite having negative magnitudes, all of the DD terms, apart from the 2012 term, are statistically insignificant, leading us to conclude that the quota had no significant impact on firms' operating revenues. For a descriptive view of Log Operating Revenue by treatment status, see Figure B.12 in Appendix 2. A similar picture emerges from our analysis of Profit Margin. While visually inspecting graph b) in figure 6 suggests that the introduction of the quota caused profit margin to drop from positive to negative, this effect is not statistically different from zero. Hence, we conclude that the quota law had no significant effect on treated firms' profitability. It is, therefore, unsurprising that our second measure of profitability, Return on Assets, fails to exhibit any significant change following the policy change. Return on Assets provides an idea as to how efficiently a company's management is using its assets to generate earnings. Lastly, we address the common claim that women are more risk averse than men (Eckel & Grossman, 2008), by examining Liquidity Ratio, which is indicative of a firm's risk attitude. Liquidity ratios are used to determine a debtor's ability to pay off current debt obligations without raising external capital. As such, a higher liquidity ratio might proxy for more risk averse decision-making by firms. We do not find any support for reduced risk taking by firms following the introduction of the quota.

Overall, we find that none of the DD terms are statistically significant, allowing us to conclude that the effect of the quota on firm performance was not statistically different from zero. Furthermore, F-tests for the joint significance of the DD coefficients in regressions of Profit Margin, Return on Assets and Liquidity Ratio show that the DD terms are jointly insignificant, supporting the conclusion that the quota had no significant effect on these outcomes. An F-test for the joint significance of the DD coefficients on Log Operating Revenue finds that the coefficients are jointly significant at the 5 percent level, leaving open the possibility that the quota may have had a slight negative impact on operating revenue over the sample period. However, these effects only represent the early adoption phase and further significant effects might take some years to transpire. Moreover, firms have not yet reached full compliance and thus female board share might still be too low to actually change the dynamics and decisions of boards.

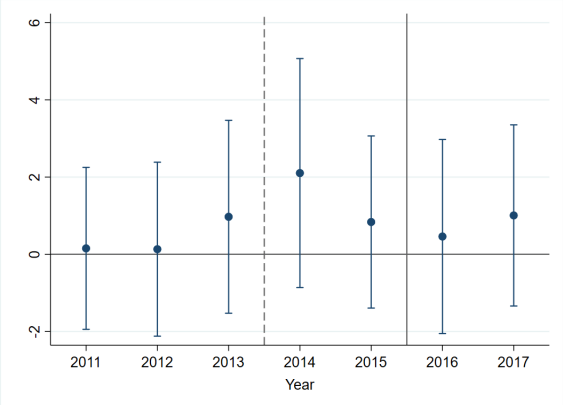
Figure 6: DID Estimates - Firm Financial Performance



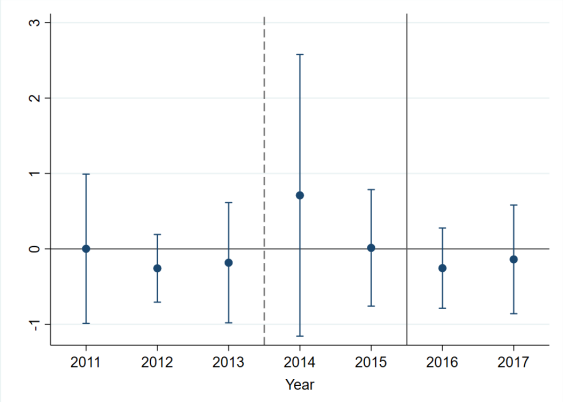
(a) Log Operating Revenue



(b) Profit Margin



(c) Return on Assets



(d) Liquidity Ratio

Source: Authors' creation

Table 6: Regression Results - Financial Performance

	ln(OpRev)	Profit Margin	ROA	Liq. Ratio
Treated x 2017	-0.139	-1.609	1.006	-0.139
	(0.106)	(1.962)	(1.193)	(0.366)
Treated x 2016	-0.123	-1.724	0.460	-0.255
	(0.123)	(2.087)	(1.278)	(0.270)
Treated x 2015	-0.0508	1.316	0.836	0.0135
	(0.151)	(2.176)	(1.133)	(0.393)
Treated x 2014	-0.121	1.377	2.104	0.710
	(0.121)	(1.918)	(1.507)	(0.949)
Treated x 2013	-0.0218	1.264	0.971	-0.183
	(0.094)	(1.851)	(1.270)	(0.405)
Treated x 2012	0.120*	0.262	0.132	-0.258
	(0.069)	(1.699)	(1.145)	(0.228)
Treated x 2011	-0.00581	0.662	0.153	0.00151
	(0.077)	(1.674)	(1.066)	(0.503)
2017	0.375***	1.994	-0.694	0.264
	(0.081)	(1.871)	(1.027)	(0.223)
2016	0.257***	1.071	-1.228	0.316
	(0.078)	(1.949)	(1.111)	(0.232)
2015	0.120	-1.199	-1.753*	0.307
	(0.102)	(1.932)	(0.934)	(0.210)
2014	0.148**	-1.477	-2.249**	0.492**
	(0.062)	(1.626)	(0.998)	(0.226)
2013	0.0939	-2.319	-1.868*	0.534**
	(0.059)	(1.687)	(1.021)	(0.247)
2012	0.0452	-0.771	-0.129	0.345*
	(0.064)	(1.609)	(0.943)	(0.193)
2011	0.0406	-0.280	0.0116	0.387
	(0.045)	(1.607)	(0.975)	(0.248)
Constant	-0.0581	7.160***	4.901***	1.905***
	(0.042)	(1.027)	(0.585)	(0.162)
R^2	0.035	0.020	0.019	0.018
Mean Dependent Variable	0.071	6.833	4.065	2.245
Mean Dependent Variable (2010)	0.397	6.556	3.992	1.843
Number of observations	1,885	1,854	1,864	1,867
Number of firms	315	308	309	310
Firm Fixed Effects	Yes	Yes	Yes	Yes

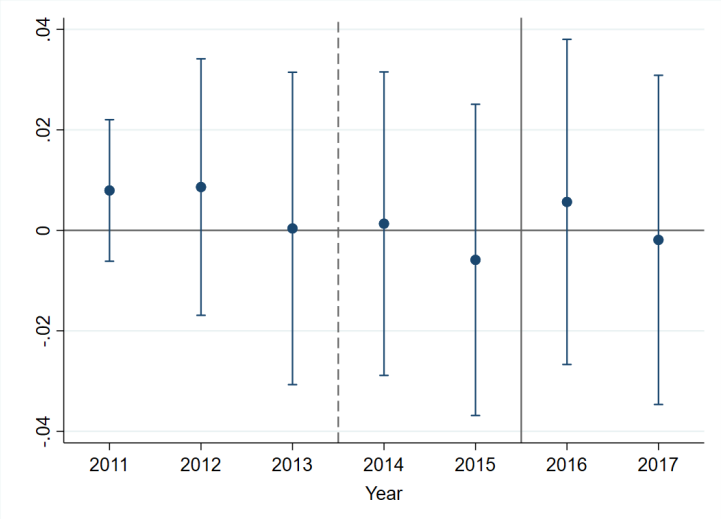
Standard errors are clustered at the firm level and presented in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The dependent variables are Log Operating Revenue, Profit Margin, Return on Assets, and Liquidity Ratio, in columns 1, 2, 3, and 4, respectively. The coefficients should be interpreted as percentages, percentage points, percentage points, and ratios, respectively.

5.4 Spillover Effects

In the final strand of our analysis, we evaluate the link between the quota legislation and potential spillover effects to the composition of executive boards. This link may exist for two reasons: if highly qualified women receive more attention in the wake of the public debate around the gender quota, then that may have directly increased their likelihood to be appointed to the executive board. Second, supervisory boards are directly responsible for the hiring and firing of executives. The quota may thus indirectly increase the share of women on executive board if the increased diversity of supervisory boards is reflected in their decisions regarding executive directors. Figure 7 plots the coefficients of the DD terms that capture the differential increase in executive board share between treatment and control firms. At a first glance, no early spillover effects exist and the quota had no significant short run effects on the composition of executive boards. An F-test confirms that the DD coefficients are jointly insignificant, confirming the null effect of the quota on this outcome. However, as Table 7 shows, we do find a positive and statistically significant time trend in our year dummies for the post-treatment period, meaning that all firms in our dataset slowly moved towards higher representation of women on executive boards. The female share of executive board members is 2.2 percentage points higher in 2016 relative to 2010, and a further 0.5 percentage points higher in 2017. We hypothesize that the increase in these years might be attributed to the voluntary measures that require firms to submit targets and deadlines for gender representation on their executive boards.

Figure 7: Female Executive Board Share



Source: Authors' creation

Table 7: Regression Results - Female Executive Board Share

	(1)	(2)	(3)	(4)
Quota x Post-Quota	0.00103			
	(0.010)			
Quota x 2017		-0.00251	-0.00129	-0.00190
		(0.011)	(0.016)	(0.017)
Quota x 2016		0.00454	0.00576	0.00565
		(0.011)	(0.016)	(0.016)
Quota x 2015			-0.00692	-0.00588
			(0.015)	(0.016)
Quota x 2014			0.000583	0.00132
			(0.015)	(0.015)
Quota x 2013			0.000628	0.000377
			(0.016)	(0.016)
Quota x 2012			0.00856	0.00861
			(0.013)	(0.013)
Quota x 2011			0.00764	0.00793
			(0.007)	(0.007)
2017	0.0256***	0.0267***	0.0262**	0.0269**
	(0.010)	(0.010)	(0.011)	(0.011)
2016	0.0235***	0.0225***	0.0220**	0.0218**
	(0.008)	(0.008)	(0.009)	(0.009)
2015	0.0133*	0.0133*	0.0152	0.0140
	(0.008)	(0.008)	(0.009)	(0.010)
2014	0.00917	0.00916	0.00887	0.00773
	(0.007)	(0.007)	(0.009)	(0.009)
2013	0.0147**	0.0147**	0.0144*	0.0141
	(0.007)	(0.007)	(0.009)	(0.009)
2012	0.0181***	0.0181***	0.0152*	0.0155*
	(0.007)	(0.007)	(0.009)	(0.009)
2011	0.000484	0.000484	-0.00219	-0.00195
	(0.004)	(0.004)	(0.005)	(0.005)
Quota	0.0167	0.0168	0.0155	
	(0.010)	(0.010)	(0.014)	
Constant	0.0125*	0.0124*	0.0130	0.0176***
	(0.007)	(0.007)	(0.008)	(0.006)
R^2	0.012	0.012	0.013	0.008
Mean Dependent Variable	0.032	0.032	0.032	0.032
Mean Dependent Variable (2010)	0.020	0.020	0.020	0.020
Number of observations	2,030	2,030	2,030	2,030
Number of firms	331	331	331	331
Firm Fixed Effects	No	No	No	Yes

Standard errors are clustered at the firm level and presented in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable is Female Executive Board Share. The unit is a proportion.

6 Discussion

In this thesis, we analyze the effect of the 30 percent gender quota on German supervisory boards, which came into effect in January 2016. The quota law required affected firms to comply on a rolling basis: depending on the office terms of directors on a supervisory board, vacancies after January 2016 must be filled with directors of the underrepresented gender until their board share reaches 30 percent. In the case of non-compliance, the seat in question would remain empty. Firms in our control group were not affected by this mandatory quota. However, they were subject to setting themselves voluntary targets and deadlines. We thus contribute to the existing literature by evaluating the effectiveness of the mandatory quota vis-à-vis voluntary targets. We find that the quota successfully increased female representation on supervisory boards of affected firms. Our estimates indicate that the quota increased the female supervisory board share by 8.1 percentage points compared to control firms between 2010 and 2016, and by 9.1 percentage points between 2010 and 2017. Moreover, treatment firms exhibit significant anticipation effects from 2014 onwards in the course of the public debate around the gender quota. We do not find any statistically significant impact on firm performance - neither in operating revenues nor in firm profitability. Furthermore, we conclude that the quota had no significant impact on treated firms' risk attitudes. Finally, two years after the introduction of the quota, we find no evidence of spillover effects to the supervisory board.

6.1 Comparison to Previous Findings

It is difficult to draw a direct comparison between our findings and the existing literature for a number of reasons. First, the different quota legislations that have been passed over the last 15 years differ in their contexts of country-specific regulations, gender attitudes and corporate governance structures. Furthermore, the different quota laws differ in their designs, both with respect to the mandated percentage under the quota and required compliance timeframe. Lastly, even within a given country, studies have produced ambiguous and inconclusive results: in Norway, Ahern and Dittmar (2012) and Matsa and Miller (2013) show negative short term effects on valuation and short term profits, while Dale-Olsen, Schone and Verner (2013) and Eckbo et al. (2016) fail to reject a value-neutral impact of the quota with respect to valuation and operating performance. Our findings in this paper are in line with the latter. This is an important finding as it goes against the popular objection raised by employer unions in Germany that treatment firms would incur an economic disadvantage from hiring a mandated share of women to their supervisory boards. It lends support to the idea that women do not have a lower aptitude for leadership and do not perform worse than men when in positions of power. This is further supported by our finding that average director experience (as proxied by age and network size) remains unaffected by the quota. Especially in the German context, where gender attitudes are historically more traditional, this finding can provide support for a cultural shift towards gender equality.

In terms of spillover effects to executive boards, we do not find any early evidence of this effect. This is in line with Maida and Weber (2019) who study the early effects of the Italian quota law passed in 2011. The law introduced a gradual gender quota for boards of directors of publicly listed companies. Similar to our results, they show that while the reform substantially raised the female membership on corporate boards, there is no evidence of spillover effects on the representation of women in top executive or top earnings positions. A similar picture emerges in Norway, where, seven years after the board quota policy came into full effect, it had no significant impact on women in affected companies beyond hiring women into the boardroom as part of the quota (Bertrand, Black, Jensen, & Lleras-Muney, 2018). However, evidence from US companies that firms with a higher share of women on their board are more likely to hire female top executives suggests that spillover effects may be dependent on the cultural context (Matsa & Miller, 2011).

Since the introduction of the quota is still very recent, we do not preclude the possibility of spillover or trickle-down effects further in time. Studies of women's legislative behavior often refer to the concept of the "critical mass", which describes the relationship between the percentage of female representation and the passing of policies that are beneficial to the female gender. The thirty percent threshold has often been identified as the minimum threshold at which to expect diversity to influence decision making (Child & Krooks, 2008). The fact that the average board share in German firms is still below this cut-off point offers a possible explanation of the lack of spillover effects to the composition of executive boards, and possibly also for the lack of effects on firm performance in either direction.

6.2 Policy Implications

Generally, quota legislations across Europe differ along four main dimensions. The first is the absolute share of women mandated by the quota, which typically is either 30 or 40 percent. The second is the severity of the penalty in case of non-compliance, ranging from no penalty (Netherlands) to liquidation of the firm (Norway). Third, the design of the compliance deadline, which can be either hard, gradual (but mandated), or rolling. Lastly, depending on the corporate governance structure, a gender quota can be designed to target different boards. In the German case, the 30 percent supervisory board quota was to be implemented according to the rolling replacement of existing board members, starting in January 2016. In case of non-compliance, the seat in question would remain empty. The results indicate that the German gender quota was a successful and cost-effective policy to increase female board representation without putting affected firms at an economic disadvantage. Over our observation period from 2010 to 2017, firms in our treatment group increased their average share of women on the supervisory board from 10.1 percent to 29.4 percent.

Our study allows us to determine the relative effectiveness of mandatory quotas versus voluntary targets. The results of our analysis clearly show that firms eligible for the mandatory quota showed a much steeper increase in female board share than control

firms, who merely received a nudge in the form of voluntary targets. We thus conclude that governments who are seriously committed to reducing inequalities in corporate leadership positions should choose mandatory quotas over voluntary targets - especially in the absence of compliance costs to firm performance.

One of the findings of our analysis is the absence of any short run spillover effects to the executive board. These may take a few more years to transpire, though previous research from Norway and Italy suggests that they may not transpire at all (Bertrand et al., 2018; Maida & Weber, 2019). In the absence of such spillover effects, the German government should consider extending their affirmative action policy to the executive board to ensure that day-to-day business decisions are taken by a more representative body of executive directors. Through the increase in the proportion of women in leadership positions, quotas might improve the representation of women's corporate policy interests. By applying the quota legislation solely to supervisory boards, the impact of the quota in representing women's interests is constrained. While supervisory board members are able to affect firm outcomes through monitoring executive board decisions, governments that wish to increase female representation in corporate decision-making within a two-tier board structure ought to target the executive board which possesses much greater leeway to effect change.

6.3 Validity

6.3.1 Internal Validity

The internal validity of our study refers to our ability to correctly identify the causal effect of the German quota policy without bias stemming from confounding effects. We discuss the potential threats to the internal validity of our study below.

One potential threat to internal validity is the presence of confounding events that might affect the outcomes of control and treated firms differently. The difference-in-differences framework allows us to eliminate the effects of confounding events that affect both groups equivalently. However, it may be reasonable to assume that the 2008 financial crisis affected the financial performance of treated firms differently to that of control firms; treated firms are publicly-listed and thus, on average, more sensitive to movements in global financial markets. If treated firms were affected by, and dynamically reacted differently to, the financial crisis, then our pre-treatment parallel trends assumption is violated, leaving our difference-in-differences coefficients biased. If, for example, treatment firms were on a stronger recovery path, then the effect of the quota on firm performance with respect to 2010 would be upward biased. While this is a possibility, we posit that our results do not show significant evidence of any such bias. Our empirical specification includes the use of lead DD terms that capture the difference between treatment and control firms prior to implementation of the quota. These terms are not statistically significant in the years prior to 2014, after which point we might expect some anticipation effects of the policy, for almost all of our regressions. Since there is no statistical test that

will verify the presence of parallel trends pre-treatment, our regression specification is a close proxy for testing the parallel trends claim, and we find confirmation of the claim here.

Movement of firms between the treated and control groups would jeopardize internal validity of our results. However, by the nature of our study design, treatment status is assigned as a time-invariant firm characteristic, precluding the possibility that firms can move between groups after treatment status has been determined. We assign all firms that were eligible to be affected by the mandatory quota legislation in 2016 to the group of treated firms for the full duration of our sample period. All remaining firms are considered untreated. It is possible that firms that met the eligibility criteria for the quota in the years preceding 2016 could avoid the quota obligation by taking certain measures. That is, firms that were affected by the quota and perceived it to be costly, could try to avoid compliance by changing their legal form or reducing their workforce below 2000 employees. However, we believe this possibility to be remote and unlikely to affect our results. First, the number of firms affected by the quota law remained stable (Holst & Wittenberg, 2019). Second, the "empty seat" penalty for non-compliance is relatively soft and rather symbolic in nature. In Norway, where firms faced the much harsher penalty of liquidation in case of non-compliance, no firms are found to have changed their legal status or changed their board size in order to manipulate compliance (Eckbo et al., 2016).

We have discussed the possibility of selection bias in the sampling procedure used to create our dataset. The BoardEx database is selective in its inclusion of private companies, and our ability to match firms between the BoardEx and Amadeus datasets further constrains this selection problem. The composition of our control group is predominantly affected by this non-random selection, with 89 of the 106 potential treatment firms included in our final dataset and subsequent analysis. Since it was not possible to verify the principles guiding selection of firms into the BoardEx database, we are unable to determine whether there is a specific characteristic of these firms that might drive the estimated effects of the quota. The problem here is that the firms included in our final dataset, and hence, used as control firms, are unrepresentative of the population of German firms.

One potential method to deal with this would be to implement inverse probability weighting to make our sample more representative of the population of German firms. We could do this by weighting the outcome measures by the inverse of the probability of a firm with a given set of covariates being matched into our final dataset. Since the Amadeus dataset contains data on a much more exhaustive set of firms than our final dataset, we could use information on firms from this dataset as "unmatched" firms to compare with the "matched" firms in the final dataset used for our analysis. We could estimate the probability of a firm being matched using a probit regression of matched status on observable characteristics of firms:

$$p(x) = P(M = 1|X = x). \tag{4}$$

We could then use the predicted probabilities to assign a weight of $1/p(x)$ to matched firms and a weight of $1/(1-p(x))$ to unmatched firms. This would then add a larger weight to firms who are underrepresented in the sample and a lower weight to overrepresented firms to make the sample group more similar to the population.

However, we choose not to do this here, on the basis that the group of treated firms is, itself, not representative of the population of German firms. Since treated firms are large and publicly-listed, they differ severely on observable characteristics from the vast majority of German firms in terms of size, governance, and financial performance. As a result, in estimating the effects of the quota on treated firms, we are likely to require a group of firms much more similar to treated firms. We check for the balance in observables between matched and unmatched firms in section 3.4 and find that matched firms are much larger than unmatched firms. As a result, these firms are more similar to treated firms, so we find no justification for making these firms more representative of the population of firms along this dimension.

In identifying the causal impact of the quota on treated firms, our DD framework requires that treated and control firms have parallel trends before the treatment period. This implies that treated and control firms must be relatively similar. A further problem of non-random selection into the dataset is that our control group of firms is not representative of the group of *treated* firms. A measure to mitigate non-random selection based on observables and to provide a better control group for treated firms is propensity score matching. Propensity score matching corrects for bias arising from differences between treated and control firms caused by a factor that predicts treatment, rather than the treatment itself, by mimicking randomization. Similar to inverse probability weighting, propensity score matching would involve estimating the probability that a given firm is subject to the mandatory quota, based on the values of observable firm characteristics, using a probit model. Treated firms are then matched with control firms using the estimated propensity scores, with the aim of approximating a random experiment.

We do not use propensity score matching in our analysis for two reasons. The first is related to the effect that we identify. Propensity score matching would be a relevant measure to estimate the average causal effect of treatment (ATE), by proxying a random experiment setup with randomized selection into treatment. The effect we estimate in this paper is the average incremental causal effect of the mandatory quota on treated firms, relative to the voluntary target measure applied to control firms (ATT). Furthermore, we believe that the effect of the quota is likely to be driven by other characteristics of treated firms. For instance, the effect of the quota on firm financial performance is likely to be driven by changes to board member dynamics resulting from the quota. As a result, were treatment to be randomized across the population of German firms, we believe that the effects of the quota would be very different to the effect on the treated population in this instance. Moreover, it is not clear that the ATE is a more interesting measure of the policy effect than the ATT. We argue that it is not. The ATT, which is here the same as the intent-to-treat effect (ITT), is informative regarding the effect of the policy on the

group it was intended to treat. This is a more informative measure of policy effectiveness for policy-makers. Based on this argumentation, we do not believe that using propensity score matching to estimate the ATE would add much to our analysis.

The second reason we do not use propensity score matching is related to practical concerns. One-to-one or even one-to-many matching algorithms typically require a larger control group than is available in our dataset. Daw and Hatfield (2018) find that when the pre-treatment outcome level is correlated with assignment to treatment, an unmatched analysis is unbiased, but that matching units on pre-treatment outcome levels produces biased estimates. This is the case for our study, in which we find that the treated group have a significantly higher pre-treatment share of female supervisory board members, for instance (see Table 2). For these practical reasons, we choose not to tackle the issue of potential selection bias using propensity score matching.

As displayed in Figure 2, there are five groups of firms that are affected by the voluntary target measure, each differentiated by the publicly-listed status of firms in the group and by the codetermination requirement on supervisory boards. In order to provide full coverage of the set of firms affected by the voluntary target measure and due to the small number of firms in each group, we do not exclude any group of firms from our analysis. However, future research using a more comprehensive dataset, such as the Amadeus archive (as in Matsa & Miller, 2013), could include a more representative sample of firms that were affected by the voluntary quota, mitigating the problem of selection bias. This would provide more conclusive evidence for identification of the average causal effect of the quota policy on treated firms relative to firms affected by the voluntary target.

6.3.2 External Validity

External validity is a concern for quasi-experimental studies of policy interventions. The results of our analysis are context-specific, making it important to assess the generalizability of the results discussed here to other settings. We argue that previous literature on board quotas is not directly generalizable to the German context, and hence, that our study adds to the field of literature on board quotas by examining the effects of a gender quota in a new context. Likewise, it is reasonable to expect that the transferability of our findings to other contexts is constrained. This is based on four key features that differentiate this policy from those implemented in Norway or in other European countries. The two-tier corporate board structure and prevailing gender attitudes in Germany provide a specialized institutional background for the quota. The quota policy is itself soft in terms of compliance deadline, which differentiates this policy from the Norwegian reform. Finally, the mandated threshold for gender representation is, at 30 percent, lower than its Norwegian predecessor. We discuss how these features affect the generalizability of our results in turn.

For instance, our regressions on financial performance outcomes suggest that the quota has no significant effect on financial performance of treated firms. A potential explana-

tion for this is that supervisory board directors have limited impact on a firm's financial performance, due to their distance from day-to-day management decisions. Under a one-tier board structure, it is plausible that a quota might affect financial performance; if the quota in Germany had been implemented on the executive board in place of the supervisory board, and if the characteristics of female directors are significantly different from male directors, we might expect a greater impact on financial performance.

However, the effects of the German quota may be informative for countries with similar corporate governance structures. Countries that have two-tier board structures and have not already introduced laws mandating minimum gender representation on boards include Croatia, Czech Republic, Denmark, Estonia, Finland, Hungary, Latvia, Poland, Romania and Slovakia. We believe that our results could inform the expected effect of future board quotas in these countries, provided that some further criteria are fulfilled. Austria, which also has a two-tier board structure, passed a law to promote gender balance on supervisory boards in 2017. The quota in Austria came into effect in January 2018. It applies during board member elections and calls for listed companies and companies employing more than 1,000 employees to achieve a composition of 30 percent women on their supervisory boards. Given its similarity to the German quota, and the similarity in culture between the two countries - the importance of which will be discussed below - we propose that our results have the requisite external validity to be extrapolated to make predictions about the impact of the Austrian quota.

The quota in Norway was mandated in December 2005 and required that firms comply with the 40 percent threshold for the underrepresented gender by January 2008. In this sense, it represents a "hard" deadline since firms that fail to achieve gender balance by this date are penalized. The quota in Germany has a "soft" deadline, in that it allows firms to comply by increasing the proportion of female board members gradually, requiring only that firms below the 30 percent threshold increase their gender balance with each board member election, rather than setting a deadline by which firms must have achieved the mandated board share. This distinction does not, in isolation, imply that the two quota legislations would have different impacts on firms' behaviour with regard to meeting the requirements of the policy. Norwegian board directors have a recommended tenure period of two years, so the Norwegian quota legislation is equivalent to the German policy in practice, by permitting firms to limit replacement of board members to their pre-existing election schedule. However, supervisory board members in Germany have a maximum tenure period of five years. As a result, the lower frequency of board elections in Germany implies that the movement towards greater gender balance on boards in Germany would be slower than that in Norway. Parallel to this argument, we might expect that a similar policy implemented in a country with different frequency of board elections would have a different dynamic effect to the German policy.

Another important criterion for the generalizability of results to other countries is the comparability of prevailing gender attitudes. Gender attitudes towards female leaders are likely to influence the effect of the policy in two ways. Firstly, they might determine

the amount of pre-emptive action taken by firms in anticipation of the policy. If appointing female directors is not viewed as costly, we might expect firms that anticipate being affected by the policy to begin to increase their share of female directors in advance. Meanwhile, firms in countries with more traditional gender attitudes might be averse to taking any action before they are legally obliged to do so. Secondly, the effect of the quota on financial performance might vary with gender attitudes. The channel through which we expect the quota to affect financial performance is through the increase in the proportion of female board members and the resulting change in board level dynamics. The effectiveness of new female board members is likely to be compromised by skepticism towards these directors from existing board members if more traditional gender attitudes prevail. This may reduce board effectiveness and harm firm performance, regardless of the individual female director's aptitude for her position. Our results are, therefore, more generalizable to countries with similar gender attitudes.

Lastly, for the comparability of results, the mandated threshold for the board share of the underrepresented gender must be of similar magnitude to the German 30 percent threshold. The effect of the policy on the female share of supervisory board members is driven by the magnitude of the threshold. Since firms are required to hire female directors only if they are below the threshold, the requirement to take action terminates when firms reach this threshold. Policies that require firms to reach a greater share of female board members will continue to have an effect after firms reach the 30 percent threshold, so we might reasonably expect such policies to be more disruptive to the status quo than the German quota and hence lead to different outcomes in firm performance. Furthermore, as stated above, the 30 percent threshold has been described as the "critical mass" needed to change legislative behavior of a given board. Regulatory frameworks that provide for a higher mandated female board share are likely to engender a higher degree of corporate policies designed to benefit women at large, as in Duflo et al. (2004), and increase the likelihood of spillover effects on the representation of women on executive boards or other management positions.

7 Conclusion

Following the example of the Norwegian quota in 2005, an increasing number of governments across Europe used or considered using gender quotas as an affirmative action policy to increase female representation in corporate leadership, more specifically on corporate boards. This thesis examines the short-term effect of a mandatory gender quota on supervisory boards implemented in Germany in January 2016. We compile a panel dataset containing data from three different sources on board composition and firm financial performance for the years 2010 to 2017. Using a difference-in-differences framework, we estimate the effects of the quota policy on four broad categories of outcomes, in order to distinguish between various channels through which the policy might take effect.

We find that, in the first two years following its introduction, the quota effectively increased the share of women on supervisory boards with respect to control firms who were only subject to voluntary targets and deadlines: our estimates indicate that the quota increased the female supervisory board share by 8.1 percentage points compared to control firms between 2010 and 2016, and by 9.1 percentage points between 2010 and 2017. Contrary to popular concern that the quota would lead to the hiring of less qualified women, we show that the change in average board experience following the introduction of the quota is not significantly different from zero. We also present evidence that firms did not incur an economic disadvantage from higher female representation in leadership as revenue, profitability and risk attitudes remained unchanged in the first two years after implementation of the quota. This is further supported by the finding that eligible firms pre-emptively increase their female board representation beginning in 2014, suggesting that firms themselves did not perceive compliance as costly. While members of supervisory boards have significant clout over the appointment and firing of executives, we do not find any evidence of spillover effects of the gender quota to the composition of the executive board. In summary, we conclude that the German board quota was a cost-effective way to increase female participation at the leadership level.

The push for gender equality has made it to the top of political agendas in many countries around the world and is one of the United Nations sustainable development goals. At its very core, gender equality is about fairness and equal opportunities. Beyond that, it is also about representation. The underrepresentation of women in leadership positions of both politics and business has serious implications for the types of policies passed by the public and private sectors, which may reinforce the status quo if left unaddressed by affirmative action policies. It is, therefore, of utmost importance to understand the effectiveness of affirmative action policies at increasing female participation on corporate boards and the effects of such policies beyond their impact on board composition. For instance, the finding that board quotas seem to lack spillover effects to executive boards and other management positions may inform policymakers of the need to extend supervisory board quotas to the executive level.

Our analysis sheds light on the short-run effects of the quota and opens the field for a range of follow-up research. Replicating our study with a more comprehensive panel dataset as time passes and more data become available may prove informative in identifying the medium- to long-term dynamic effects of the quota legislation. Since the quota was intended to send a strong signal to aspiring female leaders, further research in the vein of Bertrand et al. (2018), who to study the link between the quota legislation in Norway and women selecting into business or management programs, could be indicative about the success to this end. Furthermore, analyzing whether the quota corresponds to more female-friendly corporate policy that allows more women to climb the corporate ladder is of great importance. It would help in identifying the obstacles faced by women throughout their career trajectory and advise on the necessity and desired scope of further affirmative policy.

In conclusion, we find it essential for future studies to continue shedding light on the various channels through which affirmative action may promote gender equality at the leadership level. As governments around the world seek ways to improve opportunities and representation of women, empirical research must play a role in guiding policy decisions with evidence.

References

Ahern, K. R., & Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, *127*(1), 137–197.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, *21*(1), 1–42.

Beaurain, G., & Masclet, D. (2016). Does affirmative action reduce gender discrimination and enhance efficiency? new experimental evidence. *European Economic Review*, *90*, 350–362.

Bertrand, M., Black, S. E., Jensen, S., & Lleras-Muney, A. (2018). Breaking the glass ceiling? the effect of board quotas on female labour market outcomes in norway. *The Review of Economic Studies*, *86*(1), 191–239.

Bozhinov, V., Koch, C., & Schank, T. (2018). The second glass ceiling: Women's role in supervisory boards of german firms. *Schmalenbach Business Review*, 1–27.

Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, *72*(5), 1409–1443.

Childs, S., & Krook, M. L. (2008). Critical mass theory and women's political representation. *Political Studies*, *56*(3), 725–736.

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, *112*(5), 1297–1338.

Dale-Olsen, H., Schøne, P., & Verner, M. (2013). Diversity among norwegian boards of directors: Does a quota for women improve firm performance? *Feminist Economics*, *19*(4), 110–135.

Daw, J. R., & Hatfield, L. A. (2018). Matching and regression to the mean in difference-in-differences analysis. *Health Services Research*, *53*(6), 4138–4156.

Deloitte. (2019, January). *Missing pieces report: The 2018 board diversity census of women and minorities on fortune 500 boards.* online.

Eagly, A. H., & Carli, L. L. (2012). Women and the labyrinth of leadership. *Contemporary Issues in Leadership*, 147–162.

Eckbo, B. E., Nygaard, K., & Thorburn, K. S. (2016). Does gender-balancing the board reduce firm value? *Working Paper*.

Eckel, C. C., & Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics*, *1*, 1061–1073.

Federal Statistical Office of Germany. (2019, March). *Press release no. 098 14 march 2019.*

German Federal Government. (2015, March). *Im wortlaut: Schwesig "das ist moderne gleichstellungspolitik".* online.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.

Holst, E., & Wittenberg, E. (2019). Geschlechterquote für aufsichtsräte bringt nicht automatisch mehr vorständinnen mit sich. *DIW-Wochenbericht*, *86*(3), 35–35.

Holst, E., & Wrohlich, K. (2019). Increasing number of women on supervisory boards of major companies in germany: Executive boards still dominated by men. *DIW Weekly Report*, *9*(3), 17–32.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, *4*(3), 165–224.

Maida, A., & Weber, A. (2019). Female leadership and gender gap within firms: Evidence from an italian board reform.

Matsa, D. A., & Miller, A. R. (2013). A female style in corporate leadership? evidence from quotas. *American Economic Journal: Applied Economics*, *5*(3), 136–69.

McKinsey. (2015). *A ceo's guide to gender equality.*

New York Times. (2001, September). *Book says nixon considered a woman for supreme court.* Online.

Pande, R., & Ford, D. (2011). Gender quotas and female leadership. *Background Paper for World Development Report, World Bank*.

Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, *111*(12), 4403–4408.

Sprengholz, M., Wieber, A., & Holst, E. (2019). *Gender identity and wives' labor market outcomes in west and east germany between 1984 and 2016* (Tech. Rep.). SOEPpapers on Multidisciplinary Panel Data Research.

Wooldridge, J. M. (2013). Introductory econometrics: A modern approach 5th edition. *Mason, OH: South-Western*.

A Appendix 1

Table A.8: Regression Results - Network Size

	(1)	(2)	(3)	(4)
Treated x Post-Reform	-18.79			
	(21.993)			
Treated x 2017		-21.12	-68.18	-65.31
		(23.973)	(56.271)	(57.046)
Treated x 2016		-16.49	-63.57	-64.05
		(21.371)	(55.041)	(55.766)
Treated x 2015			-71.46	-71.91
			(57.894)	(58.606)
Treated x 2014			-59.98	-60.94
			(55.199)	(56.022)
Treated x 2013			-47.78	-48.39
			(51.053)	(51.661)
Treated x 2012			-33.02	-34.22
			(46.436)	(46.989)
Treated x 2011			-43.47	-44.69
			(41.119)	(41.625)
2017	34.43	35.14	56.79	55.24
	(35.491)	(35.795)	(51.997)	(52.795)
2016	28.12	27.42	49.11	50.89
	(34.715)	(34.757)	(51.343)	(52.088)
2015	34.36	34.36	63.33	65.32
	(33.168)	(33.177)	(54.789)	(55.514)
2014	41.69	41.69	67.23	68.92
	(31.360)	(31.369)	(52.378)	(53.232)
2013	34.55	34.55	56.14	57.27
	(28.541)	(28.549)	(48.387)	(49.017)
2012	31.04	31.04	47.84	49.65
	(26.109)	(26.116)	(44.464)	(45.040)
2011	29.45	29.46	49.33	51.02
	(22.911)	(22.918)	(39.706)	(40.227)
Treated	22.49	22.53	68.97	
	(46.520)	(46.536)	(62.291)	
Constant	371.2***	371.2***	350.1***	375.8***
	(41.466)	(41.478)	(53.211)	(30.651)
R^2	0.000	0.000	0.000	0.000
Mean Dependent Variable	412.6	412.6	412.6	412.6
Mean Dependent Variable (2010)	440.0	440.0	440.0	440.0
Number of observations	2,009	2,009	2,009	2,009
Number of firms	326	326	326	326
Firm Fixed Effects	No	No	No	Yes

Standard errors are clusted at the firm level and presented in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable is Network Size. The unit is the average number of individual connections.

Table A.9: Regression Results - Age

	(1)	(2)	(3)	(4)
Treated x Post-Reform	-0.286			
	(0.425)			
Treated x 2017		0.0935	-0.527	-0.538
		(0.488)	(0.764)	(0.777)
Treated x 2016		-0.661	-1.281*	-1.212
		(0.421)	(0.738)	(0.748)
Treated x 2015			-0.746	-0.679
			(0.728)	(0.737)
Treated x 2014			-0.530	-0.468
			(0.653)	(0.660)
Treated x 2013			-1.047	-0.969
			(0.663)	(0.669)
Treated x 2012			-0.806	-0.724
			(0.590)	(0.593)
Treated x 2011			-0.325	-0.253
			(0.506)	(0.509)
2017	-6.310***	-6.426***	-6.141***	-6.112***
	(0.488)	(0.503)	(0.649)	(0.662)
2016	-5.218***	-5.104***	-4.820***	-4.870***
	(0.466)	(0.466)	(0.624)	(0.634)
2015	-4.591***	-4.591***	-4.269***	-4.329***
	(0.393)	(0.393)	(0.612)	(0.622)
2014	-3.501***	-3.501***	-3.242***	-3.284***
	(0.358)	(0.358)	(0.578)	(0.586)
2013	-2.320***	-2.320***	-1.898***	-1.958***
	(0.366)	(0.366)	(0.593)	(0.599)
2012	-1.522***	-1.522***	-1.183**	-1.261**
	(0.323)	(0.323)	(0.523)	(0.527)
2011	-0.947***	-0.947***	-0.766*	-0.838*
	(0.272)	(0.272)	(0.448)	(0.452)
Treated	-0.394	-0.400	0.213	
	(0.526)	(0.527)	(0.758)	
Constant	65.04***	65.04***	64.76***	65.16***
	(0.488)	(0.488)	(0.619)	(0.352)
R^2	0.130	0.131	0.131	0.132
Mean Dependent Variable	61.8	61.8	61.8	61.8
Mean Dependent Variable (2010)	65.0	65.0	65.0	65.0
Number of observations	1,998	1,998	1,998	1,998
Number of firms	326	326	326	326
Firm Fixed Effects	No	No	No	Yes

Standard errors are clustered at the firm level and presented in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable is Average Board Member Age. The unit is years of age.

Table A.10: Regression Results - Log(Operating Revenue (billion EUR))

	(1)	(2)	(3)	(4)
Treated x Post-Reform	-0.0993			
	(0.063)			
Treated x 2017		-0.101	-0.120	-0.139
		(0.063)	(0.105)	(0.106)
Treated x 2016		-0.0974	-0.115	-0.123
		(0.069)	(0.122)	(0.123)
Treated x 2015			-0.0472	-0.0508
			(0.150)	(0.151)
Treated x 2014			-0.115	-0.121
			(0.120)	(0.121)
Treated x 2013			-0.0206	-0.0218
			(0.093)	(0.094)
Treated x 2012			0.120*	0.120*
			(0.069)	(0.069)
Treated x 2011			-0.00503	-0.00581
			(0.077)	(0.077)
2017	0.346***	0.346***	0.356***	0.375***
	(0.067)	(0.070)	(0.081)	(0.081)
2016	0.239***	0.238***	0.248***	0.257***
	(0.068)	(0.066)	(0.077)	(0.078)
2015	0.0934	0.0934	0.112	0.120
	(0.078)	(0.078)	(0.102)	(0.102)
2014	0.100*	0.100*	0.139**	0.148**
	(0.055)	(0.055)	(0.062)	(0.062)
2013	0.0818*	0.0818*	0.0909	0.0939
	(0.046)	(0.046)	(0.059)	(0.059)
2012	0.0809**	0.0809**	0.0440	0.0452
	(0.040)	(0.040)	(0.064)	(0.064)
2011	0.0351	0.0351	0.0385	0.0406
	(0.037)	(0.037)	(0.045)	(0.045)
Treated	2.573***	2.573***	2.591***	
	(0.208)	(0.208)	(0.209)	
Constant	-0.948***	-0.948***	-0.957***	-0.0581
	(0.120)	(0.120)	(0.125)	(0.042)
R^2	0.303	0.303	0.303	0.017
Mean Dependent Variable	0.071	0.071	0.071	0.071
Mean Dependent Variable (2010)	0.397	0.397	0.397	0.397
Number of observations	1,885	1,885	1,885	1,885
Number of firms	315	315	315	315
Firm Fixed Effects	No	No	No	Yes

Standard errors are clustered at the firm level and presented in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable is Log Operating Revenue. Operating Revenue is measured in billion Euros.

Table A.11: Regression Results - Profit Margin

	(1)	(2)	(3)	(4)
Treated x Post-Reform	-2.617***			
	(0.943)			
Treated x 2017		-2.431**	-1.928	-1.609
		(0.977)	(1.900)	(1.962)
Treated x 2016		-2.777***	-2.273	-1.724
		(1.038)	(2.028)	(2.087)
Treated x 2015			0.835	1.316
			(2.123)	(2.176)
Treated x 2014			0.776	1.377
			(1.871)	(1.918)
Treated x 2013			0.813	1.264
			(1.814)	(1.851)
Treated x 2012			-0.0862	0.262
			(1.672)	(1.699)
Treated x 2011			0.324	0.662
			(1.660)	(1.674)
2017	2.570*	2.509*	2.274	1.994
	(1.312)	(1.318)	(1.807)	(1.871)
2016	1.710	1.761	1.526	1.071
	(1.369)	(1.389)	(1.891)	(1.949)
2015	-0.428	-0.429	-0.765	-1.199
	(1.176)	(1.176)	(1.877)	(1.932)
2014	-0.563	-0.562	-0.881	-1.477
	(0.982)	(0.983)	(1.572)	(1.626)
2013	-1.532	-1.531	-1.864	-2.319
	(1.003)	(1.004)	(1.648)	(1.687)
2012	-0.385	-0.384	-0.420	-0.771
	(0.944)	(0.945)	(1.581)	(1.609)
2011	0.227	0.228	0.0621	-0.280
	(0.957)	(0.957)	(1.593)	(1.607)
Treated	0.867	0.866	0.372	
	(0.980)	(0.981)	(1.703)	
Constant	6.141***	6.142***	6.368***	7.160***
	(1.136)	(1.136)	(1.584)	(1.027)
R^2	0.008	0.008	0.008	0.007
Mean Dependent Variable	6.833	6.833	6.833	6.833
Mean Dependent Variable (2010)	6.556	6.556	6.556	6.556
Number of observations	1,854	1,854	1,854	1,854
Number of firms	308	308	308	308
Firm Fixed Effects	No	No	No	Yes

Standard errors are clustered at the firm level and presented in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable is Profit Margin. The unit is percentage points.

Table A.12: Regression Results - Return on Assets

	(1)	(2)	(3)	(4)
Treated x Post-Reform	-0.195			
	(0.643)			
Treated x 2017		0.207	0.528	1.006
		(0.732)	(1.225)	(1.193)
Treated x 2016		-0.542	-0.218	0.460
		(0.712)	(1.301)	(1.278)
Treated x 2015			0.257	0.836
			(1.160)	(1.133)
Treated x 2014			1.481	2.104
			(1.527)	(1.507)
Treated x 2013			0.474	0.971
			(1.300)	(1.270)
Treated x 2012			-0.329	0.132
			(1.184)	(1.145)
Treated x 2011			-0.291	0.153
			(1.104)	(1.066)
2017	0.0631	-0.0685	-0.230	-0.694
	(0.801)	(0.814)	(1.064)	(1.027)
2016	-0.564	-0.453	-0.614	-1.228
	(0.850)	(0.874)	(1.141)	(1.111)
2015	-1.063*	-1.062*	-1.203	-1.753*
	(0.619)	(0.620)	(0.969)	(0.934)
2014	-1.116	-1.114	-1.633	-2.249**
	(0.734)	(0.734)	(1.031)	(0.998)
2013	-1.174*	-1.172*	-1.376	-1.868*
	(0.690)	(0.691)	(1.060)	(1.021)
2012	0.269	0.271	0.335	-0.129
	(0.633)	(0.633)	(0.990)	(0.943)
2011	0.399	0.401	0.459	0.0116
	(0.623)	(0.624)	(1.017)	(0.975)
Treated	1.364*	1.362*	1.048	
	(0.749)	(0.749)	(1.236)	
Constant	3.800***	3.801***	3.953***	4.901***
	(0.811)	(0.811)	(1.072)	(0.585)
R^2	0.009	0.009	0.009	0.007
Mean Dependent Variable	4.065	4.065	4.065	4.065
Mean Dependent Variable (2010)	3.992	3.992	3.992	3.992
Number of observations	1,864	1,864	1,864	1,864
Number of firms	309	309	309	309
Firm Fixed Effects	No	No	No	Yes

Standard errors are clustered at the firm level and presented in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable is Return on Assets. The unit is percentage points.

Table A.13: Regression Results - Liquidity Ratio

	(1)	(2)	(3)	(4)
Treated x Post-Reform	-0.263			
	(0.246)			
Treated x 2017		-0.200	-0.156	-0.139
		(0.191)	(0.356)	(0.366)
Treated x 2016		-0.317	-0.273	-0.255
		(0.319)	(0.261)	(0.270)
Treated x 2015			-0.0259	0.0135
			(0.371)	(0.393)
Treated x 2014			0.674	0.710
			(0.944)	(0.949)
Treated x 2013			-0.228	-0.183
			(0.394)	(0.405)
Treated x 2012			-0.273	-0.258
			(0.215)	(0.228)
Treated x 2011			0.0114	0.00151
			(0.499)	(0.503)
2017	0.319	0.299	0.274	0.264
	(0.256)	(0.232)	(0.212)	(0.223)
2016	0.336	0.354	0.329	0.316
	(0.219)	(0.240)	(0.223)	(0.232)
2015	0.331*	0.331*	0.326*	0.307
	(0.196)	(0.196)	(0.198)	(0.210)
2014	0.742**	0.742**	0.523**	0.492**
	(0.366)	(0.366)	(0.219)	(0.226)
2013	0.489**	0.489**	0.559**	0.534**
	(0.209)	(0.209)	(0.251)	(0.247)
2012	0.273**	0.273**	0.358**	0.345*
	(0.131)	(0.131)	(0.178)	(0.193)
2011	0.385	0.385	0.376	0.387
	(0.237)	(0.237)	(0.243)	(0.248)
Treated	-0.463	-0.463	-0.505*	
	(0.395)	(0.395)	(0.275)	
Constant	2.131***	2.132***	2.155***	1.905***
	(0.288)	(0.287)	(0.250)	(0.162)
R^2	0.009	0.009	0.010	0.006
Mean Dependent Variable	2.245	2.245	2.245	2.245
Mean Dependent Variable (2010)	1.843	1.843	1.843	1.843
Number of observations	1,867	1,867	1,867	1,867
Number of firms	310	310	310	310
Firm Fixed Effects	No	No	No	Yes

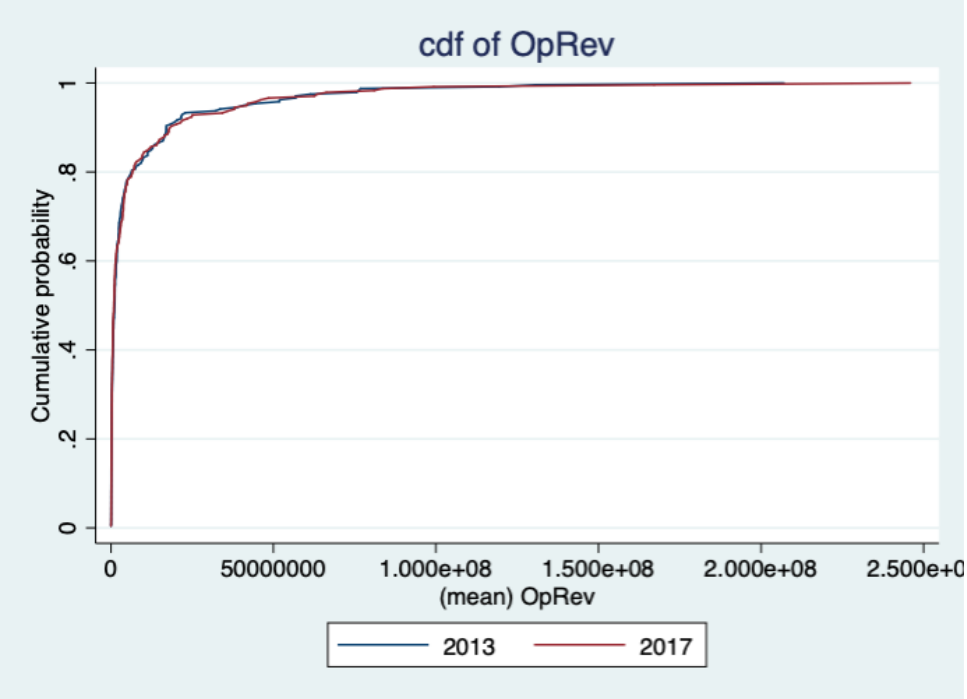
Standard errors are clustered at the firm level and presented in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable is Liquidity Ratio. The variable is unitless, taking a fraction value.

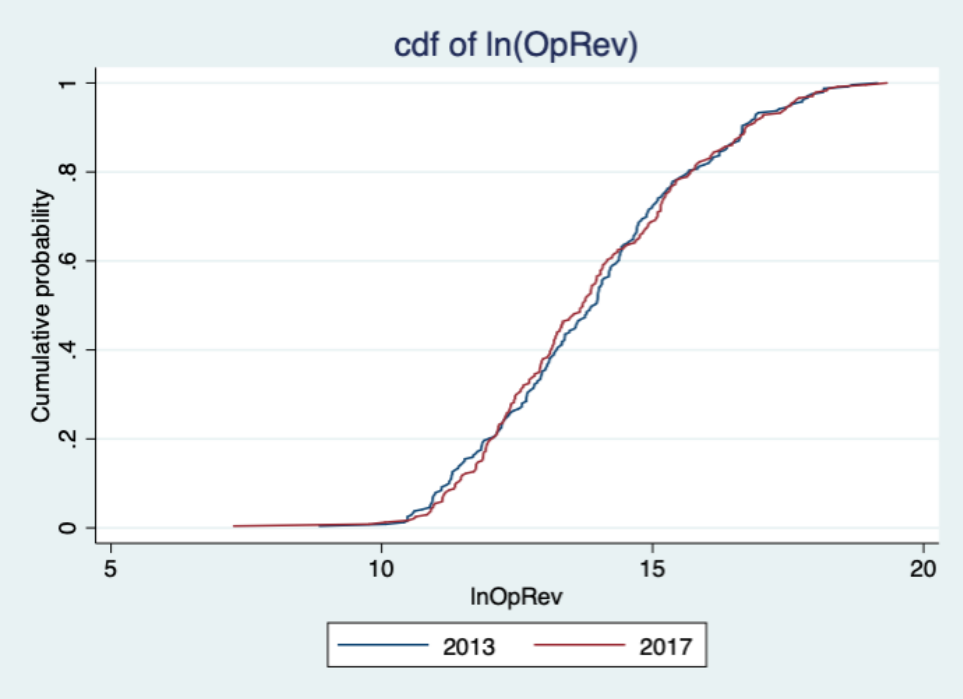
B Appendix 2

Figure B.8: Cumulative Distribution Function of Operating Revenue



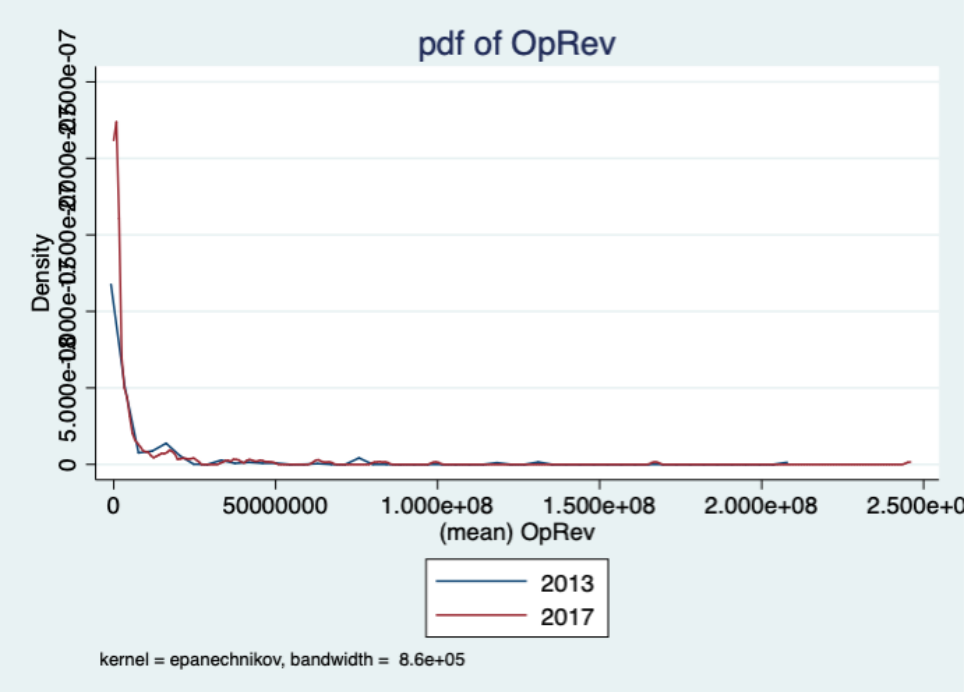
Source: Authors' creation

Figure B.9: Cumulative Distribution Function of Logged Operating Revenue



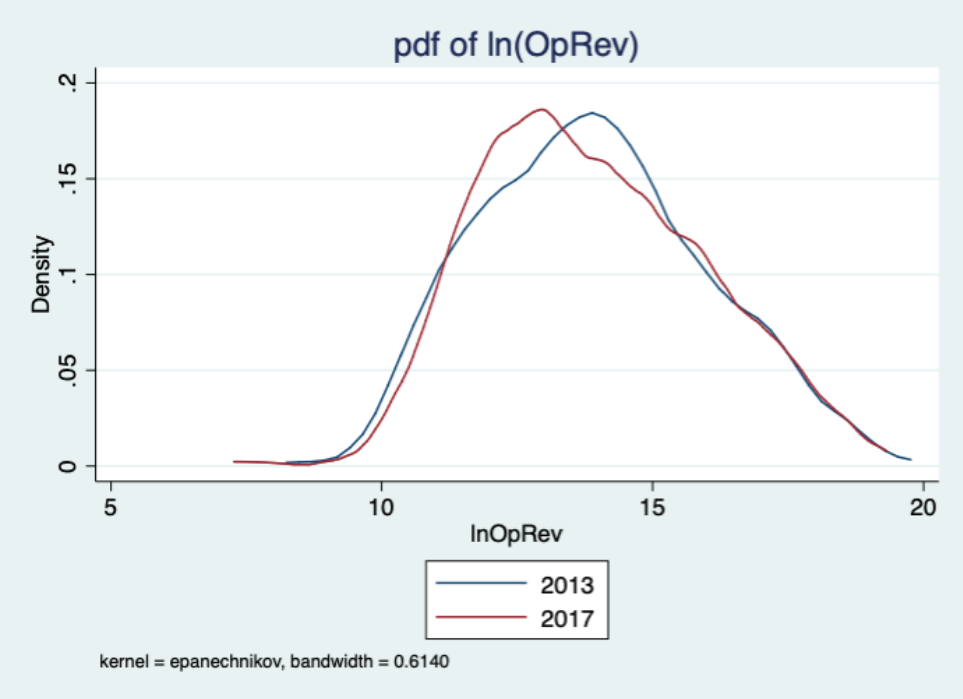
Source: Authors' creation

Figure B.10: Probability Distribution Function of Operating Revenue



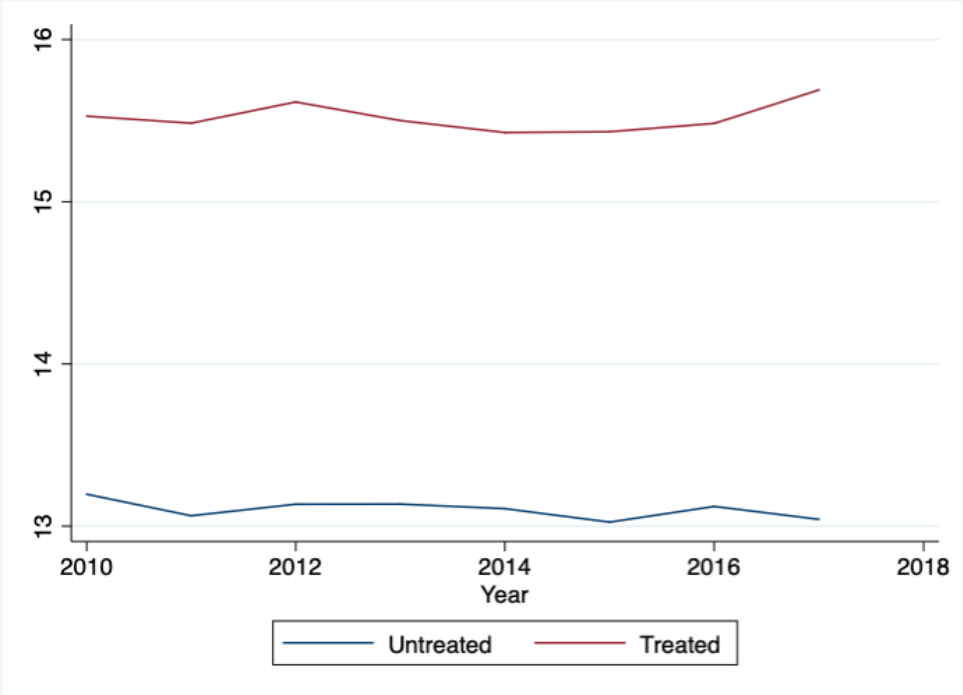
Source: Authors' creation

Figure B.11: Probability Distribution Function of Logged Operating Revenue



Source: Authors' creation

Figure B.12: Log Operating Revenue by Treatment Status



Source: Authors' creation