The effect of a German female member of parliament on the number of female candidates: a Regression Discontinuity Design Approach.

Bachelor Thesis Policy Economics by Isabel Orlemans *

Supervisor: P.R. Bose

Second assessor: M. Bastiaans

5th of August 2022

This thesis examines the effect of a directly elected female member of parliament on the number and proportion of female candidates in the next election in Germany. Research on this topic is important due to the high societal gains of an increase in female political representation. A Regression Discontinuity Design with data from the elections of 2002-2017 indicates an insignificant effect of the election of a female member of parliament on the number and percentage of female candidates within constituencies and districts, except for a decrease of 1 to 3.5 female candidates within her constituency. However, the insignificant effect on the proportion of female candidates within districts is significant when examining the effect on the subgroups of females with and without a previous job in politics, with respectively a negative and positive effect. These findings show the need for further research on the heterogeneous effects and mechanisms and also indicate that policymakers should minimalize or even positively shift the negative and zero effects.

* Student (student number: 541425) of Erasmus School of Economics (ESE), Erasmus University Rotterdam.

Table of Contents

1. **Introduction**

Improving gender equality is a major issue on the global agenda at present. This is not surprising given the social relevance of this issue. Half of the population is female and besides, economic theory shows that an improvement toward gender equality has a huge societal impact due to its increase in economic efficiency and growth (e.g. Heathcote et al., 2010; Morais Maceira, 2017). One of the most important factors to improve gender equality is female political representation, especially since it allows women to directly partake in public decision-making. Worldwide, women are generally underrepresented in federal governments. To be more precise, 26% of all members of parliament worldwide were women in 2021 and only in five countries the parliament consisted of at least 50% women (The World Bank, 2021). The female political underrepresentation is surprising due to its potential social benefits. A higher share of female politicians is beneficial for the total society, with different but positive effects on developing and developed countries (Hessami & da Fonseca, 2020). To be more precise, in developing countries a higher share of female politicians raises the provision of public goods, especially healthcare and education goods, while in developed countries public policies' spending patterns are unaffected, but a higher share does affect the exact public policy choices and parliamentary deliberations. Moreover, a higher share of women improves the institutional quality due to a reduction in corruption and rent extraction. Therefore, research on the political representation of women is socially beneficial and therefore also socially relevant.

Gender equality is also one of the key challenges for the German government, in both the national and international context. Therefore, the German government supports a large number of projects aimed to decrease gender inequality (German Federal Foreign Office, 2022). Nevertheless, German public sector directors are divided on the possible implementation of gender equality rules and most of them consider their companies to do well regarding gender equality. In contrast, The German Confederation of Trade Unions doubts if the latter is realistic, and also argues that gender equality rules are beneficial (Die Welt, 2022). In comparison to the other European countries, Germany ranks 10th out of 27 in 2021 in the gender equality index by the European Institute for Gender Equality (2021). As already argued, female political representation is an important factor in improving gender equality. Currently, the German parliament consists of 34.92% women and only 32.59% of all election candidates in 2021 were female (IPU Parline, 2022). By 2020 the German population consisted of 50.6% (The World Bank, n.d.) and thus women are underrepresented in the German federal government. The German share of female members of parliament is comparable to the proportion of women in national parliaments in the whole European Union, for which 33% of all members of parliament are women (The World Bank, 2021). Thus, even though the German government wants to increase gender equality, women are politically underrepresented in parliament, showing the importance of studying this subject.

One determinant of female political representation is the number of female electable candidates since it could positively affect the number of female members of parliament, this is in line with economic theory

and empirical evidence. Such an effect can be explained through four mechanisms; the empowerment, role model, contagion and incumbency effect. In the first-mentioned mechanism, a female MP uses her power in her party to influence its candidates' selection patterns, thereby affecting the number of female candidates (Kittilson, 2010). In the second mechanism, the female MP is a role model for others and thereby changes their beliefs, due to women politically underestimating themselves or having non-egalitarian beliefs about women in politics (Fox et al., 2010; Ladam et al., 2018) Further, in the third-mentioned mechanism, the female MP affects the number of female candidates through other parties since these change their candidate selection patterns since as a reaction to an elected female from another party (Matland & Studlar, 1996). In the last-mentioned mechanism, the female MP benefits from her incumbent status by running for re-election and thereby increasing the number of female candidates in the next election (Jankowski et al., 2019). Thus, the four above-mentioned mechanisms, which are extensively discussed in section 3.1, can explain an effect of a female MP on the number of female candidates in the next election and therefore, the research question is as follows:

What is the effect of a directly elected female member of parliament on the proportion and number of female candidates in the next federal election in Germany in 2002-2017?

This research question is answered with a Regression Discontinuity Design (RDD) as the main strategy. To measure the effect as aimed for in the research question, four outcome variables are considered; the absolute number and proportion of female candidates in the next election, within constituencies and districts. In contrast to the expected positive effect, no economically significant effect is found, except for a decrease in the number of female candidates within her constituency. Specifically, a female MP decreases the number of female candidates by 1 to 3.5 candidates. Further, in the additional analysis, the effects on the subgroups of females with and without a previous job in politics in districts are examined. Interestingly, the insignificant effect on the total proportion of these two groups becomes significant for both subgroups. More precisely, the election of a female MP increases the share of female candidates without a former job in politics by around 2% and decreases the share of female candidates with a former job in politics between 1.3% and 2.4%. This implies that the effects within constituencies could also be significant when estimating the effect for the same subgroups.

This thesis contributes to academic research in several ways. Most importantly, as related literature shows, the effects differ between countries (Broockman, 2014; Bhalotra et al., 2017; Jankowski et al., 2019; Valentim, 2021). Therefore, a country for which this research question has not been studied yet, Germany, is studied in this thesis. Furthermore, the data period is more recent, which could affect the results due to time shocks. Moreover, an additional analysis on the effects on female candidates with and without a previous job in politics is performed, which is not as extensively discussed in the related literature. To summarize, this thesis differs from related literature in several aspects, with the country of study as the most important difference, showing the scientific relevance of this thesis.

The rest of this thesis is structured as follows. In section 2, the German election setting is discussed. Afterwards, the literature review is presented with subsections on mechanisms and related literature, in section 3. Hereafter, in section 4, the data and methodology are reviewed. In sections 5 and 6, the results of respectively the main and additional analysis are presented, followed by robustness checks for the first in section 7. The discussion can be found in section 8. Section 9 concludes.

2. Institutional background

This section starts with a description of the German federal elections and their voters. Thereafter, the voting ballots and the allocation of the seats in parliament are discussed. To end, candidate selection within their parties is discussed.

The German federal election is called "Bundestagwahl" and elects the German parliament for a four-year term. All Germans who are at least eighteen years old on the election day have the right to vote if they are German residents or have been normally residents. However, some Germans are excluded from this right, for example, because of a court decision. Voters can cast their ballots at polling stations or can vote by mail (Federal Ministry of the Interior and Community, 2017). These are cast in one of the 16 districts that consist together of 299 constituencies. The number of constituencies differs between districts. The district with the smallest number of constituencies is Saarland, with only 4 constituencies. In contrast, the district with the highest number of constituencies is Nordrhein-Westfalen, which has 64 constituencies. In this thesis, constituencies and districts are distinct terms and are thus not similar.

Each ballot has two votes, the first vote is for a candidate in the voter's constituency and the second one is for a party in the corresponding district. For the first vote, each party is allowed to nominate one candidate in each constituency. For the second vote, each party must provide a ranked list of candidates. The candidate that is directly elected to the Bundestag must receive the most first votes and is thus elected using relative majority voting. Thus, for each constituency, there is one directly elected member of parliament (MP). Hereafter, the other seats of the Bundestag are allocated. These are allocated to the parties proportionally according to the second votes for each district. However, the number of seats won directly by the first votes is deducted from the remaining seats. These remaining seats are assigned to candidates in the order they appear on the party's list in the corresponding district. For parties to enter the Bundestag, they must receive at least five per cent of second votes (Federal Ministry of the Interior and Community, 2018).

The direct candidate selection is similar to the selection of list candidates. Besides the candidate selection, the candidate rank is determined according to the same procedure. In a party's assembly, a democratic and secret election must be held to select the candidates for both the constituencies as well as for the districts. For both, the candidates are selected primarily by their parties' members. Candidates may be proposed by a member who is allowed to vote in the candidate selection. These candidates may not be candidates from other parties since 2009, but do not necessarily have to be a member of the party.

Furthermore, the constituency candidates may also be chosen by an association of voters. In the setting of the first votes in Germany, an elected female can affect the proportion of female candidates in the next election in both the constituency as well as the district since she could affect the selection of direct candidates. In addition, she could also influence the list rank, but this is beyond the scope of this thesis. These potential mechanisms are more extensively discussed in section 3.

3. Literature review

This section starts with a discussion of the possible mechanisms which could explain an effect of a female MP on the number of female candidates in the next election. For these, a short description, literature findings and expectations in the German setting are reviewed in subsection 3.1. Similar research to this thesis is already performed for related settings like minorities and female mayors, but also in more comparable settings like other countries. This is discussed in subsection 3.2.

3.1 Mechanisms

An increase in female candidates is important for increasing female political representation, which is, as already mentioned, highly beneficial for society (Hessami & da Fonseca, 2020). Several mechanisms can explain how the election of a female MP could cause such an increase in the proportion and number of female electable candidates. First, is the empowerment effect, which affects other (new) candidates due to the influence of the female MP – i.e. using their power in the party to influence candidates' selection patterns (Kittilson, 2010). An example of this effect is when female MP influences which candidates are nominated. This is closely related to networking. Second, is the role model effect, in which a female MP affects other (new) candidates due to changing beliefs as a result of exposure to female politicians – i.e. the woman is a role model for other individuals - and are therefore more likely to run for office. These belief changes exist due to women politically underestimating themselves or having non-egalitarian beliefs about women in politics (Fox et al., 2010; Ladam et al., 2018). Third, is the contagion effect. In this mechanism, the female MP does not influence other women but affects other parties since their candidate selection patterns change in reaction to an elected female from another party (Matland & Studlar, 1996). These authors define contagion as a process in a multiparty system, in which one party stimulates other parties to adopt their policies or strategies, and therefore call the effect the contagion effect. Fourth, is the incumbency effect, in which the female MP benefits from her incumbent status since she is more likely to run for re-election, therefore increasing the number of female candidates in the next election (Jankowski et al., 2019). In this context, the incumbent is defined as an individual who has a particular position, especially an elected one (Cambridge Dictionary, 2022). Thus, the four above-mentioned mechanisms could explain an effect of a female MP on the number of female candidates in the next election.

Literature provides evidence for the existence of these mechanisms in various settings. To start with the empowerment effect, female candidates are more likely to be nominated when the gatekeeper - those

responsible for candidate recruitment - or a party elite is a woman rather than a man – e.g. the local party president (Cheng & Tavits, 2009; Kittilson, 2010). Further, in line with the role-model effect, when female citizens are represented by women, they are more active in politics, with bigger effects in the early stages and on younger individuals (Atkeson, 2003; Wolbrecht & Campbell, 2007; Fridkin & Kenney, 2014; Gilardi, 2015), or their political interest increases (Bühlmann & Schädel, 2012). Moreover, female representatives can affect political interests positively for both men and women (Karp & Banducci, 2008), or even negatively (Foos & Gilardi, 2019). Closely related to the role-model effect is that prior exposure to a female chief counsellor improves the female leader's effectiveness and weakens gender role stereotypes in both public and domestic spheres (Beaman et al., 2009). Interestingly, this effect is also present in party elites, providing evidence for the contagion effect. Moreover, women incumbents do increase the participation of female candidates in their own and other parties (Palmer & Simon, 2005). To end with the incumbency effect, in which incumbents theoretically run for re-election when their political skill is equal to or higher than the population's average political skill and when the incumbents' benefits of leaving politics are lower than the benefits of staying in politics. In addition, the literature suggests that women can benefit from the incumbent advantage relatively more (Jankowski et al., 2019), but it is also feasible that women and men are equally likely to benefit from this advantage (Allik, 2015; McGregor et al., 2017; Fiva & Røhr, 2018). Thus, literature on these mechanisms provides evidence for these mechanisms to be present, although this evidence is sometimes somewhat mixed. This implies the importance of these mechanisms and studying the effects of the election of a female MP.

In the German first votes setting, all four mechanisms could be present. In the case of the empowerment effect, a female MP uses her power to influence the number of female candidates within her party only. Since each party can have one directly electable candidate in a constituency and she is likely to run for re-election, this effect is only present in her district and not her constituency. To continue to the role model effect, a female MP affects candidate nominations of her party and other parties, in both her constituency as well as her district, since she can be considered a role model for all candidates. This is thus not only limited to her party. Next, the contagion effect, only affects other parties in her constituency and district, with a bigger effect in her constituency since these other parties are confronted with her parties' female representation more directly. To end with the incumbency effect, it is expected, that she only affects the number of female candidates when using her incumbency advantage within her constituency since she will likely run for re-election in the same constituency. However, since the number of female candidates in her constituency is also included in the district numbers in this thesis, the effect will also be present in districts. Thus, it is expected that the empowerment effect affects only the proportion of female candidates in districts, while the role model, contagion and incumbency effect affects the proportion within both the constituencies and districts. Furthermore, it is expected that at least one of these mechanisms positively affects the outcome variables in the German setting. Therefore,

the hypothesis states that the election of a female MP increases the number and proportion of female candidates within her constituency and district.

3.2 Related literature

A regression discontinuity design, the main approach of this thesis, can also be implemented in other settings. For example, for estimating the effect of electing an ethnic minority candidate. Just like ethnic minorities, females are politically underrepresented (e.g. Bloemraad & Schönwälder, 2013; Childs & Hughes, 2018; Kantola & Lombardo, 2019**).** In close mixed-ethnic races, black victories are more likely than black losses and are characterized by a higher voter turnout and probability of future black victories, in municipal elections (Vogl, 2014). The election of a black mayor does not affect the adopted policies, but does affect the black representation among police officers and positively affects the income and relative unemployment of the African American population (Hopkins & McCabe, 2012; Nye et al., 2014). Thus, the election of an underrepresented minority candidate positively affects the future election of these minorities, positively affects their wealth and increases their representation also outside of politics.

In Germany the effects of a female representative on female candidates are also studied, using the same estimation strategy as in this thesis, for mayors instead of MPs. These council elections are held once every five years using an open-list electoral rule – i.e. preference voting. In these elections, female candidates are relatively more advanced regarding their initial rank if the mayor is female (Baskaran & Hessami, 2018). Due to these findings of a positive effect in Germany, it is especially interesting to study whether a positive effect is present for female MPs and the federal election candidates as well.

The RDD has already been used to measure the same effect of an elected female on the number of female candidates. Four of these papers are similar to this thesis, the main difference is, however, the studied country. To be more precise, these papers consider the United States, India, Poland and the United Kingdom. The findings are somewhat mixed. A positive effect of an elected female candidate on the number of female candidates is present, but this effect is mainly driven by the incumbent effect (Bhalotra et al., 2017; Jankowski et al., 2019). In addition, Jankowski et al. find some support for the contagion effect and find no evidence for the empowerment effect in Poland. However, the effect on the number of respectively federal and local female candidates is insignificant for the United States and the United Kingdom (Broockman, 2014; Valentim, 2021). Due to these mixed findings for different countries, it is especially relevant to study whether a female MP significantly affects potential female candidates.

The country of study can affect the effects of a female MP, which makes it especially interesting whether the effect is present in Germany as well. Besides a different culture, politics and its systems also differ between the mentioned countries and Germany. For example, Germany had a female chancellor for four of the five studied election periods, while the United States, India and the United Kingdom had no women as head of state and Poland had only one period a female prime minister. In addition to a

difference in the country, this thesis differs in several other aspects from the above-mentioned papers. First, the data's period is more recent, which could influence the results due to time shocks. An example of a shock which could affect the number of female candidates is the social movement on sexual harassment #MeToo, which started in 2016. Second, this thesis focuses on both the constituencies as well as districts, while the related papers focus on constituencies and neighbouring constituencies. Due to the German electoral system, it is likely that a female MP also affects non-neighbouring constituencies in her district. Studying neighbouring constituencies is also interesting, but is beyond the scope of this thesis. Third, in this thesis, an additional analysis is presented on the effects on female candidates with and without a previous job in politics, which is not extensively discussed in the other papers. One could expect a heterogeneous effect for these subgroups, for example, because the role model effect affects potential female candidates differently.

4. Data and Methodology

This section starts with a description of the data, including a discussion of the descriptive statistics in subsection 4.1. Afterwards, the description of an estimation strategy using ordinary least squares and its endogeneity issues are presented in subsection 4.2. Hereafter, the main estimation strategy, the regression discontinuity design, are discussed in subsection 4.3. Since this is the main estimation strategy, the last subsection, subsection 4.4, describes the endogeneity issues and robustness checks of this approach.

4.1 Data

To answer the research question, two data sources are merged using the programming language Python. The first dataset contains data on the characteristics of all federal election candidates, provided by the German government. These characteristics include names, gender, previous job and year of birth. Since these characteristics include the candidates' gender, this data allows for creating the outcome variable as well as determining whether an election race is between candidates of different genders. This data will be matched with the second dataset, which includes election outcome data and is collected from the website bundeswahlleiter.de (The Federal Returning Officer, 2021). Since every party has at most one candidate in each constituency, the datasets are matched on the constituencies and parties. The characteristics dataset allows for studying all six waves of federal elections in the period from 2002 to 2021. Since this thesis aims to study the effect of a directly elected female on the proportion and number of female candidates, not all election races are considered. For a race to be considered in this study, the top two candidates, that is the two candidates with the highest number of first votes, have to be of different genders. Further, to measure the outcome variable on time t, the treatment variable on the gender of the winning candidate for election on time t -1 is needed. Therefore, the data on the election of 2021 is only used to measure the outcome variable, and thus the effects of a female MP for the five elections in the period 2002-2017 are studied.

As can be seen in the descriptive statistics of Table 1, the total sample consists of 633 mixed-gender election races, of which 182 are close mixed-gender election races. If not stated otherwise, a race is considered close when the winner of the race won with a relative vote margin of at most 10%. A bandwidth of 10%, 5% or 1% are all intuitive, however, the sample size is too small for the last two maximum relative vote margins. The formal definition of the relative vote margin can be found in subsection 4.2.

In close mixed-gender election races, men won 51.6% and women 48.4% of the races. These percentages suggest that the selection of a male or female is random in close mixed-gender election races, due to these percentages both being close to 50%. In contrast, male candidates win more often in all mixed-gender elections. This same pattern holds for the vote percentages of the candidates; males have higher vote percentages in mixed-gender elections but have similar vote percentages in close mixed-gender elections. Furthermore, all samples' averages have a year of birth close to 1960, although women are slightly younger than their male opponents. The majority of the mixed-gender elections candidates have a former job in politics, which is slightly higher for the sample with all mixed-gender elections.

Table 1: Descriptive statistics characteristics of mixed-gender elections

	All mixed- gender elections			**Close mixed-gender elections**		
	All	**Male**	**Female**	**All**	**Male**	**Female**
Won	0.5	0.654	0.346	0.5	0.516	0.484
Vote percentage	0.372	0.407	0.336	0.370	0.372	0.369
Relative vote margin	0	0.093	-0.093	0	0.004	-0.004
Year of birth	1960.233	1960.088	1960.378	1959.64	1959.033	1960.253
Former job in politics	0.559	0.580	0.5387	0.514	0.5	0.527
N	1266	633	633	364	182	182

Remarks: this table shows the means of characteristics for both all mixed-gender elections and close mixed-gender elections. A mixed-gender election is defined as close if the relative vote margin $\delta_{c,i,t}$ (difference between the vote percentage for the two candidates with the highest percentage within a constituency divided by its sum) is within the range of -0.10 and 0.10. The means are rounded and are shown for the whole sample as well as for the male and female subsamples. The dummy variable won indicates whether the candidate won the election. The dummy former job in politics is 1 if the German government defined the job category of the candidate as 71 or 76, which includes for example a member of the Bundestag or a mayor. N indicates the number of candidates in the sample.

4.2 Ordinary Least Squares

As a starting point, an ordinary least squares (OLS) regression will be performed. To define the corresponding regression equation, the relative vote margin and the treatment definitions are discussed first. For an election at time t in an individual constituency c, the relative margin of victory $\delta_{c,i,t}$ of candidate i is defined in Equation 1. Thus, if the candidate won the election, $\delta_{c,i,t}$ is positive, while it is negative if the candidate lost the election. In the latter case, the candidate is also referred to as the runner-up. In addition, a relative vote margin of zero indicates a tie, this is of no concern for elections due to its small probability of happening and does not occur in the data. In this setting, the value of zero is therefore a natural threshold. The formal definition of the relative vote margin can be found in

Equation 1 and a histogram of this variable in the German mixed-gender elections can be found in Figure 1.

(1) $\delta_{c,i.t} = \frac{v_i - v_o}{v_i + v_o}$, where v_i and v_o are the vote percentages of resp. candidate i and opponent o.

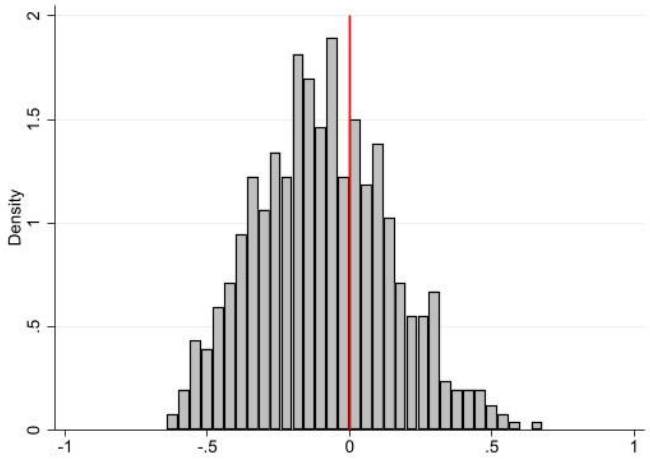

Figure 1: Histogram relative vote margin, $\delta_{c,i,t}$

The treatment status depends on the relative vote margin of an election between a female and male candidate. Namely, the mixed-gender election is treated when the female candidate won the mixed-gender election race and is used as a control group when the relative vote margin is negative. The latter implies that the male candidate won the mixed-gender race. The formal definition of the treatment status can be seen in Equation 2. From this follows the OLS regression equation in Equation 3.

(2) $T_{c,i,t} = \begin{cases} 1 \text{ if } \delta_{c,i.t} > 0 \text{ and gender is female} \\ 0 \text{ if } \delta_{c,i.t} < 0 \text{ and gender is female} \end{cases}$

(3) $Y_{c,i,t+1} = \alpha + \beta T_{c,i,t} + \epsilon_{c,t+1}$

In which, $Y_{c,i,t+1}$ is the outcome variable, measured for the next election at time $t + 1$. In this thesis, four different outcome variables are studied to answer the research question. These enable exact measurement of the effect aimed for in the research question. The first outcome variable measures the share of candidates that are female, in the next election in the corresponding constituency. The second outcome variable measures instead of the share, the absolute number of female candidates in the next election in the corresponding constituency. To investigate the effect including area spillovers, the third and fourth outcome variables measure the same as the first two, but for the corresponding district instead of the corresponding constituency. This is interesting since a female candidate could also affect the candidate (selection) in neighbouring constituencies through the same mechanisms. However, not all neighbouring constituencies are in the same districts and not all constituencies in districts are neighbours. Ideally, an outcome variable that only measures the effect on neighbouring constituencies is studied, but this is beyond the scope of this thesis. These four outcome variables are respectively

referred to as share constituency, absolute constituency, share district and absolute district. Furthermore, the coefficient of interest, β, measures the effect of an elected female on the outcome variable $Y_{c,i,t+1}$. It is expected that β statistically differs from zero since this implies a true effect of the elected female on the outcome variable, as aimed for in the research question. Lastly, α is the constant and $\epsilon_{c,t+1}$ the error term in a given constituency at time $t+1$.

Since the percentage of candidates in mixed-gender races who had a former job in politics is relatively high, which is presented in Table 1, it is interesting to additionally analyse two components of the share district variable. Namely, the share of candidates without as well as with a former job in politics within a district, and are respectively referred to as share new district and share former district. This enables to study whether the effect of a female MP is heterogeneous for these subgroups, and also enables to possibly target policies more efficiently. Examining the effects on the constituency is also relevant, but is behind the scope of this thesis. These outcome variables are constructed using the job category classification of the German government. To be more precise, a female candidate is considered to have a former job in politics when the job category is 71 or 76 and is considered to have no former job in politics otherwise. Examples of jobs that fall within these categories are a MP and a mayor. The share is defined as the number of female candidates within the classification, divided by the total number of candidates. As a consequence, the sum of these two shares is equal to the share district variable. Moreover, like for the main outcome variables, the female candidate herself – i.e. incumbency- is also included in the share former district variable when she runs for re-election. Thus, besides the effect on other female candidates, incumbency is also included. The analysis of these additional variables will be similar to the analysis of the main outcome variables and are discussed in section 6.

This regression is performed on the total sample of mixed-gender election races, as well as close mixed-gender election samples. In this case, close has multiple definitions and is defined as a relative vote margin of at most 15, 10, 7.5 and 5% respectively. It needs to be noted that the sample with a bandwidth of 5% results in a too small sample size of 88. The other sample sizes are big enough for causal inference and can be found in section 5.

Nevertheless, an OLS regression in general suffers from omitted variable bias (OVB), which causes the estimated effect not to be causal. Research by Bhuller et al. (2020) is an example of this since it shows that the sign of the effect differs between using an unbiased instrumental variables design and OLS. Thus, omitted variables bias the estimation, and can even change its sign. Relevant omitted variables are variables that influence both the dependent and independent variables. In this context, these variables must influence both the relative vote margin of mixed-gender elections as well as the number of female candidates. An example is the percentage of religious voters within the constituency since religious individuals are more likely to have traditional gender role beliefs – e.g. women should take care of the household- and are therefore less likely to vote for a female candidate, decreasing her relative vote

margin. In addition, the number of female candidates will be affected by the percentage of religious voters, since following the same reasoning, females might be less likely to run for office. Since the candidate sample is not random, a bias is likely present and therefore the estimates in this setting cannot be interpreted as a causal effect. Though, it does measure the correlation between the election of a female and the number of female candidates.

4.3 Regression Discontinuity Design

The main estimation strategy is a regression discontinuity design (RDD), which in the past decade has become a commonly used research method in electoral settings. An RDD estimates the local average treatment effect by comparing observations just before and after a threshold. Because of this, it is expected that treatment is assigned randomly within this sample. All elements that could affect women's participation are hence naturally held constant, and thus there is no need for control variables. Therefore, this approach can determine a causal effect around a threshold if its correctly implemented and its three minimal assumptions hold and is thus preferred to OLS (Hahn et al., 2001; Antonakis et al., 2010). These assumptions are more extensively discussed in subsection 4.4. The German federal election system allows for using a sharp regression discontinuity design since the first votes directly elect MPs using relative majority voting, and therefore a threshold of being elected in a close race arises. Because of this, only candidates who are directly elected to the Bundestag and their biggest competitors are considered. In close elections, the winner being a male or female is expected to be randomly distributed when the assumptions hold.

Even though global parametric RDD approaches were common, it is currently more common to use a local polynomial RDD approach. A global parametric RDD uses all observations in a sample while the local polynomial RDD only uses the observations within a specific bandwidth. One of the advantages of the local polynomial RDD is a reduction in the risk of misspecification bias (Lee & Lemieux, 2010). A misspecification bias arises for example when the form of the estimated function is incorrect, e.g. estimating a parametric function while the true function is linear. Besides, a random treatment is more likely to be true for a local polynomial RDD. Therefore, a sharp local polynomial RDD approach is used, which requires extending Equation 3 with a function of the relative vote margin $f(\delta_{c,i,t})$ and its interaction effect with the treatment effect. This function specifies the relationship form of the running variable $\delta_{c,i,t}$ and the outcome variable $Y_{c,i,t+1}$, which is also referred to as the order or degree of the polynomial. When the degree of the polynomial is one and a small bandwidth is used, the approach is called a local linear regression. As research by Gelman and Imbens (2018) shows it is not always justified to use polynomials with a polynomial of three or higher, only the first and second order polynomials are studied. They argue that using higher-order polynomials, leads to noisy estimates, sensitivity to the degree of the polynomial and poor coverage of the confidence intervals. In-general, the true function of the effect after a graphical analysis is not known, and therefore both functions are

considered, especially since there do not exist statistical methods to select the correct form. The interaction effect of this function with the treatment effect allows for different functions on both sides of the cut-off. The resulting regression equation can be seen in Equation 4 and will be performed using the STATA-package *rdrobust*.

$$(4)\ Y_{c,i,t+1} = \alpha + \beta T_{c,i,t} + f(\delta_{c,i,t}) + T_{c,i,t} f(\delta_{c,i,t}) + \epsilon_{c,t+1}$$

Two other important choices related to this design are the bandwidth and kernel choice. First, is the bandwidth selection, for which one bandwidth is presented in the main results, and several other bandwidths are considered in the robustness section. The main bandwidth is the intuitive bandwidth of ten percent. As already discussed, the bandwidths of five and one percent are also intuitive but require reducing the sample size too much for estimating a causal effect. Second, is the kernel choice. In general, all observations in the RDD sample are equally weighted, and thus the RDD is performed with a uniform kernel. Since this is the most common kernel, the uniform kernel will be considered the main kernel. In addition to this uniform kernel, the results are also performed for a triangular kernel. A triangular kernel attaches a relatively higher weight to the observations closer to the threshold. The reason for attaching a higher weight to closer observations is that these observations are more informative than observations further away from the threshold for estimating the treatment effect. This likely causes a triangular kernel to be more appropriate to use instead of a uniform kernel (Cattaneo et al., 2020). Further, the difference in results between the kernels suggests whether the effect differs for races with a smaller relative vote margin.

4.4 Endogeneity issues and robustness checks

As already mentioned, the regression discontinuity design is currently one of the preferred methods in economics since it can estimate a causal effect around the threshold if the method is correctly implemented and the three assumptions hold. Before discussing the three assumptions and related additional tests, possible endogeneity issues are discussed in this subsection.

In general, omitted variable bias (OVB) is a major concern for estimating a causal effect. When an omitted variable is not included in a model, the estimate is biased and therefore the measured effect is not causal with certainty. This concern is the main concern of OLS, which is discussed in subsection 4.1. However, for the RDD this concern is not present, since the deterministic treatment status fully depends on the running variable and since treatment is randomly assigned in a small bandwidth around the threshold. Due to these two features, the error term is not correlated with other variables, and therefore no omitted variables exist. Another endogeneity issue arises when the data suffers from non-random measurement errors. Since the data is election data, it is unlikely that non-random measurement errors arise. Of course, the vote shares can slightly be affected by wrong voting counts, but this is in general randomly distributed. Other issues with the estimation strategy arise only when at least one of the three assumptions does not hold.

First, the treatment must be assigned as good as random in a small neighbourhood around the threshold. This assumption does not hold when there is sorting around the threshold, thus when candidates can *fully and precisely* manipulate the vote shares within their constituency. It is expected that this assumption holds in a free democratic election setting – if there is no voting fraud. To test if this expectation is valid in this setting, several tests are performed. The first check for this assumption is a visual inspection of a histogram of the relative vote margin– i.e. the running variable- to observe whether there are remarkable changes in density, suggesting manipulation. To also statistically test this, a density test is performed, which examines whether there is statistical evidence for manipulation concerning the continuity of the density function. The density test that will be performed is developed by Cattaneo et al. (2018) and allows for the testing of various bandwidths. In addition to this density test, a balance test is performed for different bandwidths, in which a regression of characteristics on a dummy variable indicating the treatment and control groups is performed. If the characteristics of the candidates do have a significant effect on being in a particular treatment group, the assumption does not hold since then the treatment and control group do statistically differ. The available characteristics which will all be included in the regression are gender, year of birth and former political job. No characteristic should statistically affect the assignment to the treatment or control group for the method to be valid, this is especially important for the gender characteristic since it is the studied characteristic. Due to the sample sizes and sometimes overlapping definitions of the other job categories, no other job category dummy variables are included in the balance test. Since there are no other meaningful characteristics included in the candidates dataset, no other characteristics are included in the balance test. Even though three of the most important determinants of election candidates are included in the model, including other characteristics would enhance the balance test. Nevertheless, three of the most important characteristics are included. The above-mentioned three methods that possibly provide evidence for the assumption of random assignment around the threshold to hold, and thus provide evidence for the results to be valid, are presented in subsection 7.1.

Second, nothing else changes around the threshold except for the treatment. This assumption likely holds, since except for being elected, almost nothing else changes around the threshold in the German election setting. The only thing that does change at the threshold is that for the party of the directly elected constituency candidate, the number of seats in parliament for other candidates decreases by one. However, the total number of seats for each party stays the same, and therefore this is not expected to be a problem. Usually, a placebo test is performed to test if indeed nothing else changes around the threshold, however, no data is available to perform such a test.

Third, the relationship between the running variable, $\delta_{c.t}$, and the outcome variable Y_{ct} can be adequately modelled if the sample size requires moving further from the threshold. For this reason, the function form and the interaction effect are included. Furthermore, to test to which extent the bandwidth influences the estimates and its significance, the results are presented for various additional bandwidths,

including different bandwidths on both sides of the thresholds, in subsection 7.2. In addition to the manual bandwidths, an optimal bandwidth selector is used to determine the bandwidth. This bandwidth selector is the Mean Square Error (MSE) and is developed by Calonico et al. (2014). If the results change significantly due to different bandwidths, the results should be interpreted carefully. In addition, the Coverage Error Rate (CER) bandwidth selector is presented in section 7.3. This bandwidth selector, developed by Cattaneo et al. (2020), is designed to examine inference, and therefore the estimates do not represent the magnitude of the treatment effect accurately.

5. Results

In this section, the main results are presented. This result section starts with the presentation and discussion of descriptive statistics of the outcome variables as well as the graphical presentation of the raw data in subsection 5.1. Hereafter, the Ordinary Least Squares regressions are examined in subsection 5.2. This section ends with subsection 5.3, in which the results of the main estimation strategy, the RDD with the manually selected bandwidth of ten percent, are presented and discussed. In section 6, to investigate whether the effects are heterogeneous, the district effects are divided into the number of female candidates with and without previous political experience.

5.1 Data

For a first impression of the data, the descriptive statistics of the four outcome variables are presented for both the total and close mixed-gender election samples in Table 2. To start with the outcome variable share constituency, which measures the share of female candidates in the next election within the same constituency, a slightly higher mean is observable when a woman won the election in the total sample. Surprisingly, for the close mixed-gender election sample, the mean of this outcome variable is higher when a man won the previous election, but the difference is however modest. To be more specific, when a woman won the outcome variable is on average 27.7%, while it is 28.0% when a man won the previous election in a close mixed-gender election. Along the same line, the mean of the absolute number of female candidates within the corresponding constituency is higher when a man won the previous election for both the total and close mixed-gender election samples. This suggests that for the constituency outcome variables in the close mixed-gender election sample, the effect as aimed for in the research question is negative or insignificant.

To continue to the district variables, for all mixed-gender elections, the means of both the share and the absolute number of female candidates is higher when a man won the previous election. This is also the case for the absolute number of female candidates in the close mixed-gender election sample, but is, however, the reverse for the share variable. Interestingly, for all variables, the means of the treatment and control group within a sample are relatively close to each other, except for the absolute district variables. To be more precise, when a woman won the previous election in the total sample, the average number of female candidates within a district is 65.4, while it is 76.2 when a male won the previous

election. Since for the total sample, the election of a female or male is likely not only caused by the gender of the candidate, the means only provide an indication of the German situation but do not provide enough information to determine the effect of electing a woman. However, it is expected that the election of a female or male is random within close mixed-gender elections. Also for this sample, the absolute district variable means are not relatively close to each other for the treatment and control groups. Specifically, the mean of the number of female candidates is 57.3 for the treatment group while it is 65.7 for the control group.

Table 2: Descriptive statistics outcome variables of treatment and control groups

	All mixed- gender elections			Close mixed-gender elections		
	Full sample	Treatment	Control	Full sample	Treatment	Control
Share constituency	0.266	0.288	0.254	0.279	0.277	0.280
Absolute constituency	3.569	3.511	3.599	3.698	3.591	3.798
Share district	0.245	0.243	0.247	0.247	0.249	0.245
Absolute district	72.458	65.425	76.179	61.632	57.273	65.712
N	633	219	414	182	88	94

Remarks: this table shows the means of the four different outcome variables $Y_{c,i,t+1}$ for both all mixed-gender elections and close mixed-gender elections. A mixed-gender election is defined as close if the relative vote margin $\delta_{c,i.t}$ (difference between the vote percentage for the two candidates with the highest percentage within a constituency divided by its sum) is within the range of -0.10 and 0.10. The means are rounded and shown for the whole sample as well as the treatment and control samples. The treatment sample consists of all elections within a constituency in which a female won the election and a man lost the election. The control sample consists of all elections within a constituency in which a male won the election and a female lost the election. The N indicates the number of mixed-gender election races within the sample.

A visual presentation of the relationship between the relative vote margin and all four outcome variables within the total sample, by means of a scatterplot with all data points, can be seen in Figure 2. In this figure, the observations for which the relative vote margin is positive make up the treatment group, while the control group consists of all negative relative vote margin observations. The functional form of the relationship between the relative vote margin and the outcome variables is not identified after a visual inspection. Therefore, both the polynomial with an order of one and two are examined, since it is unclear which polynomial is more suitable in the RDD. Higher order polynomials are not examined since, as already mentioned in section 4, a polynomial of a higher form can lead to evidently nonsensical results (Gelman & Imbens, 2018). Furthermore, the figures suggest a small or insignificant effect of the gender of the winning candidate on the outcome variables, since there is no jump visible in the outcome variables. The scatterplots for the close-mixed gender election sample are presented in Appendix A.1. Regardless of the smaller bandwidth, the conclusions about the true form of the relationship and the effect are similar.

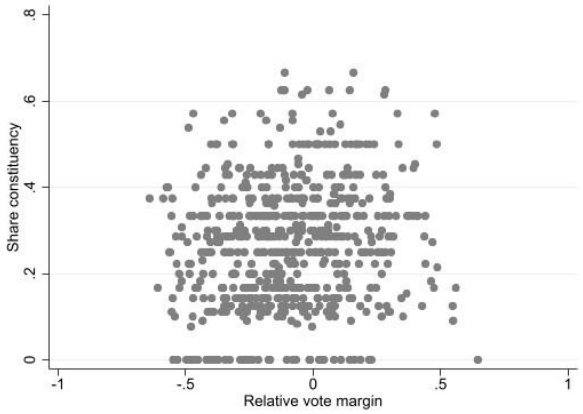

Figure 2.a: scatterplot of full sample, share constituency

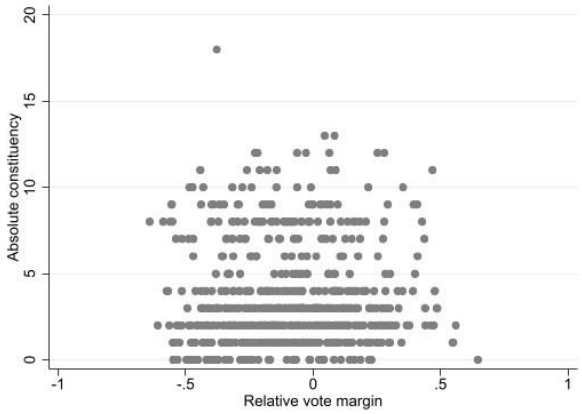

Figure 2.b: scatterplot of full sample, absolute constituency

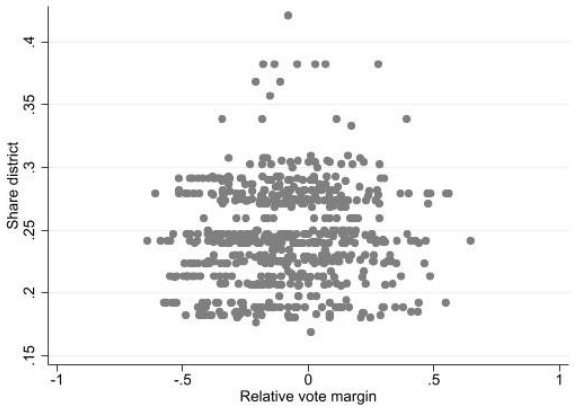

Figure 2.c: scatterplot of full sample, share district

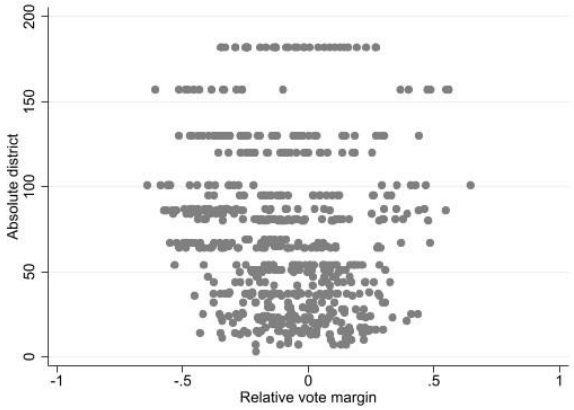

Figure 2.d: scatterplot of full sample, absolute district

5.2 Ordinary Least Squares

For comparison to the main estimation strategy, the OLS regression results can be found in Table 3. Interestingly, two of the four outcome variables in the full sample are highly significant, while none of the outcome variables is significant in the smaller samples – i.e. close mixed-gender elections. To start with the full sample, in which the results indicate a statistically significant positive effect, an elected female increases the share of female candidates in her constituency in the next election by 3.4% on average. Even though this magnitude is relatively small, this effect can be considered economically significant. In contrast, the effect on the absolute number in the constituency and on the share of female candidates within a district are both statistically insignificant. Due to the size of the robust standard errors of these estimates, it is unlikely but possible that this effect is truly zero. Moreover, the number of female candidates in a district in the next election decreases significantly when a female won a mixed-gender in this district with on average 10.8 candidates. Due to the size and significance, this effect is economically highly significant.

None of the estimates in the smaller samples are statistically significant, nevertheless, a small effect may exist. The differences in significance between the full and smaller samples imply that female candidates who were directly elected with a higher relative vote margin -i.e. won with relatively more votes- do affect the number and proportion of female candidates in the next election differently. This provides evidence for the existence of omitted variables, which is in general a major concern when applying OLS. As expected, the results of the regressions are probably biased, especially since there are no control variables included. Nevertheless, the OLS results allow for comparing other results in this thesis.

Table 3: Effects elected female in mixed-gender elections on $Y_{c,t}$ using OLS

	Full sample	**15%**	**10%**	**7.5%**	**5%**
Share constituency	0.034***	0.011	-0.003	0.004	-0.012
	(0.012)	(0.018)	(0.021)	(0.025)	(0.325)
Absolute constituency	-0.088	-0.160	-0.207	-0.249	-0.077
	(0.244)	(0.344)	(0.452)	(0.533)	(0.632)
Share district	0.004	0.002	0.004	0.006	0.007
	(0.003)	(0.005)	(0.006)	(0.007)	(0.008)
Absolute district	-10.754***	-7.750	-8.440	-4.769	-12.909
	(3.849)	(5.483)	(6.570)	(7.739)	(9.888)
N	633	275	182	133	88

Remarks: this table shows the OLS regression estimates of the effect of an elected female on several outcome variables for multiple samples. A mixed-gender election is included in the sample if the relative vote margin of the elected candidate is lower than or equal to the indicated percentage. The robust standard error is in between brackets. All estimates and robust standard errors are rounded. N indicates the number of mixed-gender election races. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

5.3 Regression Discontinuity Design

Contrary to the OLS approach, the RDD approach does not suffer from a biased estimate due to its random nature. Therefore, this is the main estimation strategy with a bandwidth of 10% - i.e. the relative vote margin of the winner is at most 10%. A graphical representation of the RDD for the outcome variable share constituency is presented in Figure 3, for both orders of polynomials and both kernels. In these figures, the grey dots represent data bins, that can be constructed to have an equal length in terms of the relative vote margin or can be constructed to include a similar number of observations. The advantage of the latter is that the means are estimated with a similar precision (Cattaneo et al, 2020), and therefore the bins represent the data's binned local averages. To be more precise, the data-driven procedure to select the number of bins is selected with a mimicking variance evenly−spaced method using spacings estimators. This is the default option in the package *rdrobust*. On the right side of the threshold, which is illustrated with the red line, the binned local averages represent the treatment group, while these on the left represent the control group.

After the graphical analysis with binned local averages, it is unclear what the functional form of the relationship is, and therefore both the polynomials will be considered. For the uniform first-order polynomial as well as the uniform second-order polynomial, a jump is visible at the threshold. These jumps indicate a positive effect of an elected female on the proportion of electable women within the same constituency. As already mentioned, the main kernel is the uniform kernel. However, the triangular

kernel could indicate whether the effect differs for the candidates who won with a smaller relative vote margin. For the share constituency outcome variable, the use of the triangular kernel somewhat changes the jump and the relationship. Specifically, the jump increases due to the triangular kernel and therefore suggests a larger effect for closer mixed-gender races. For the other outcome variables, the graphical representation of the uniform RDD approach is presented in Figures 4-6. Except for the second-order polynomial of the absolute constituency outcome variable, a jump is also visible around the threshold. Interestingly, for both district variables, the jump implies a negative effect, indicating a negative effect of the election of a female MP on the proportion and number of female candidates in the corresponding district.

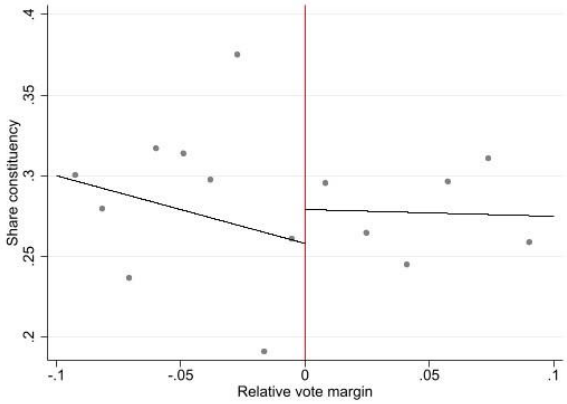

Figure 3.a: graphical representation RDD share constituency, uniform kernel, 10% bandwidth, order 1

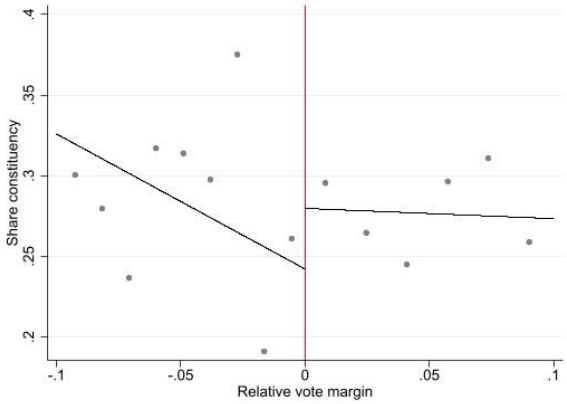

Figure 3.b: graphical representation RDD share constituency, triangular kernel, 10% bandwidth, order 1

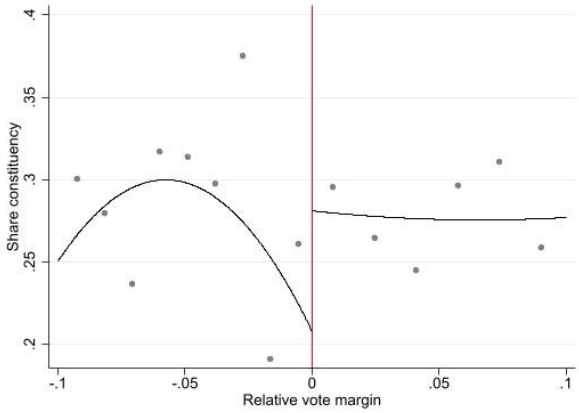

Figure 3.c: graphical representation RDD share constituency, uniform kernel, 10% bandwidth, order 2

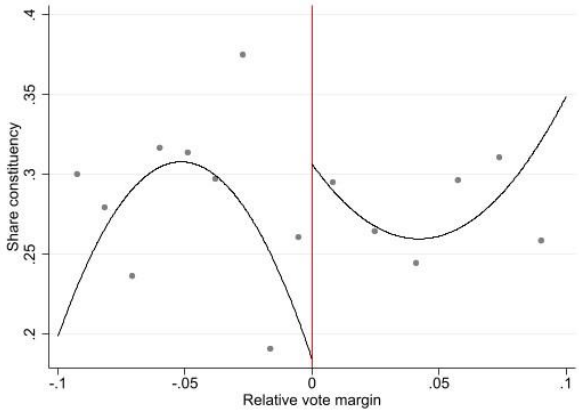

Figure 3.d: graphical representation RDD share constituency, triangular kernel, 10% bandwidth, order 2

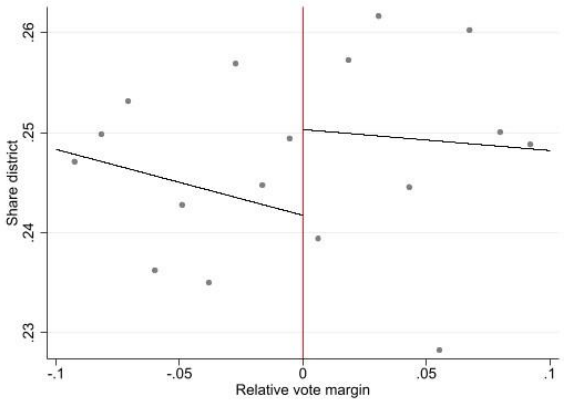

Figure 4.a: graphical representation RDD absolute constituency, uniform kernel, 10% bandwidth, order 1

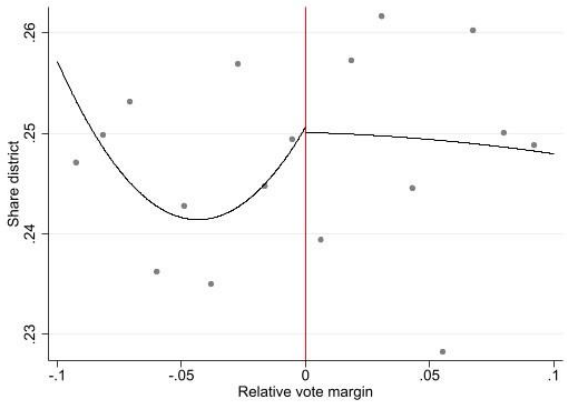

Figure 4.b: graphical representation RDD absolute constituency, uniform kernel, 10% bandwidth, order 2

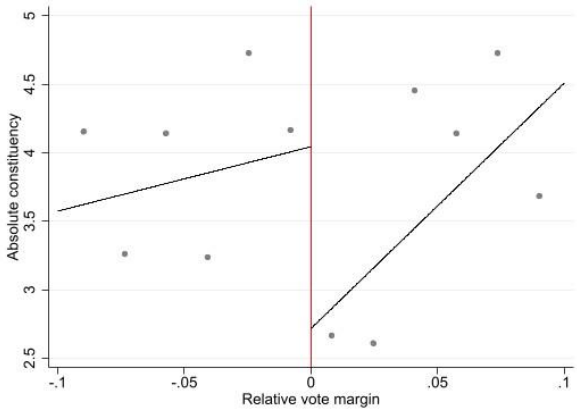

Figure 5.a: graphical representation RDD share district, uniform kernel, 10% bandwidth, order 1

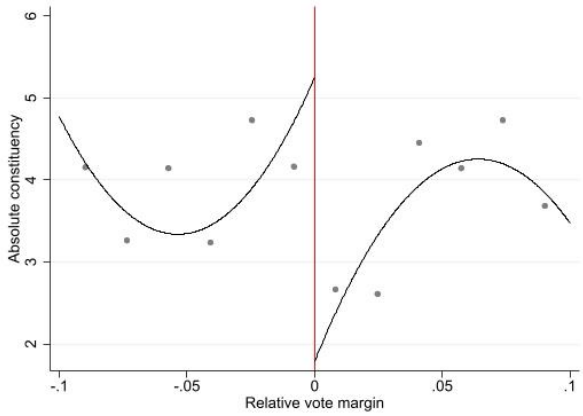

Figure 5.b: graphical representation RDD share district, uniform kernel, 10% bandwidth, order 2

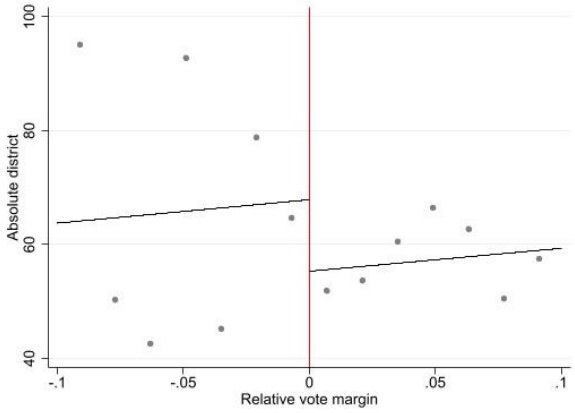

Figure 6.a: graphical representation RDD absolute district, uniform kernel, 10% bandwidth, order 1

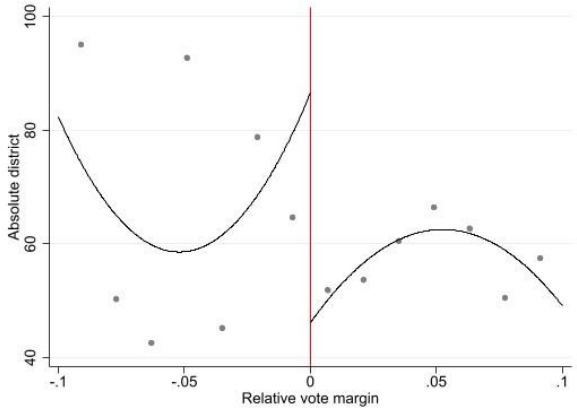

Figure 6.b: graphical representation RDD absolute district, uniform kernel, 10% bandwidth, order 2

To also examine the sign, magnitude and significance of these jumps, the results of the RDD in a close mixed-gender election, corresponding to Figures 3-6, are presented in Table 4. As already argued, the functional form of the relationship between the relative vote margin and the outcome variables is unknown and therefore both the first- and second-order polynomial are discussed. Since the estimates and their significance differ between the first and second order, the point estimates are sensitive to the choice of the order of the polynomial. Therefore, the estimates should be interpreted carefully.

First, the results of the outcome variable share constituency. The local linear regression estimates that an elected woman increases the share of female candidates in her constituency by 2,1% in the next election. However, this effect is not significant, and the size of the standard error implies that it is likely for this effect to be truly zero. In contrast, for the second order, the uniform treatment estimate is significant at the 10% level. The share of female candidates increases by on average 7.3% if a woman instead of a man is elected. However, the triangular kernel, which estimates a larger effect, suggests that this effect is insignificant for both forms. It is therefore unclear whether there is a significant positive or zero effect for this outcome variable. Since the first-order polynomial and the triangular kernels do suggest no effect, and since the second-order polynomial with a uniform kernel is only significant at the 10% level, it is most likely that the effect is zero or small and positive. Thus, the estimates of the election of a woman on the share constituency variable provide evidence for a zero or a small effect.

Second, the results of the absolute constituency variable, which measures the absolute number of female candidates within the same constituency in the next election. The results for a local linear regression with this outcome variable indicate an insignificant decrease of on average 1.3 female candidates due to the election of a woman. Whether this effect is truly zero, cannot be determined due to the size of the standard error. In contrast, for the second-order polynomial, a decrease of on average 3.5 female candidates due to the election of a female MP is estimated, with a significance level of 5%. In addition, the kernels for both polynomials are significant at the 10% level. Therefore, the election of a woman instead of a man likely negatively affect the number of female candidates within a constituency. The effect of an elected female candidate in the previous election on the number of female candidates in the same constituency is estimated to result in a decrease of 1-3.5 female candidates.

Third, the results of the share district variable. The election of a woman increases the share of women within the district by 0.9% by exploiting a local linear regression and decreases the share by 0.1% by exploiting a second-order polynomial. However, these effects are insignificant for both polynomials as well as for both kernels. Nevertheless, the standard errors do not confirm nor reject a zero-effect. Thus, the election of a female instead of a male does likely not affect the share of female candidates in the next election significantly.

Fourth, the results of the absolute district variable. For interpretation purposes, it needs to be mentioned that a district includes multiple constituencies, and therefore the district sizes are relatively larger than

the constituency sizes. The results show that the election of a woman decreases the absolute number of female candidates in the district using a first- and second-order polynomial with respectively 12.6 and 40.4 female candidates. The first mentioned effect is however not significant, but the second one is significant at the 10% level. As well as for the estimated local linear regression, the triangular kernel estimates are insignificant. Following the same reasoning as for the share constituency variable, it is most likely that the effect is zero or small and positive. Thus, the estimates of the election of a woman on the absolute district variable provide evidence for a zero or a small effect.

Table 4: RDD of elected female in mixed-gender elections on $Y_{c,t}$ with 10% bandwidth

		Uniform		Triangular	
Outcome variable	**Order**	**Treatment estimate**	**S.E.**	**Treatment estimate**	**S.E.**
Share constituency	1	0.021	0.046	0.038	0.050
Share constituency	2	0.073*	0.068	0.122	0.071
Absolute constituency	1	-1.328	0.969	-2.029*	1.094
Absolute constituency	2	-3.463**	1.670	-3.479*	1.800
Share district	1	0.009	0.011	0.006	0.120
Share district	2	-0.001	0.016	-0.004	0.018
Absolute district	1	-12.563	14.205	-21.635	15.939
Absolute district	2	-40.445*	23.619	-28.241	26.442

Remarks: this table shows the RDD regression estimates and standard errors of the effect of an elected female in a close mixed-gender election on several outcome variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most 10%. These are performed for both a uniform and triangular kernel, as well as for the regression forms with an order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The total number of mixed-gender election races is 182, with 94 observations below the cut-off and 88 after the cut-off. The stars indicate the significance level with *: $p<0.10$, **: $p<0.05$, ***: $p<0.01$.

To summarize the main results and hereby answer the research question, the election of a female instead of a male does not or modestly affect the share constituency and absolute district. Furthermore, it likely affects the absolute number of female candidates within the same constituency in the next election significantly, with an estimated negative effect between 1 and 3.5 candidates. In contrast, the election of a female candidate does not affect the share of female candidates within the corresponding district. Interestingly, only the absolute constituency variable is likely significant, and in combination with its magnitude, it is the only outcome variable with an economic significant effect. This is especially remarkable since the effect on the absolute constituency is significant while the share constituency is not or modestly affected. One possible explanation is that in the district in which a woman won a close election fewer candidates are electable, both men and women. For instance, because of the incumbency advantage of the woman, making other candidates less likely to run. Another possible explanation is that the election of a woman does not significantly affect the share of women in her constituency but does affect the gender of the candidates of parties that do not run for office anymore. In addition, it is interesting that the OLS results for the close mixed-gender elections in Table 3 also provided an insignificant result for the outcome variable district share, and provided evidence for a zero effect for the other variables. Since the share constituency and absolute district variable are likely zero or small, the estimation of OLS is somewhat similar to the RDD estimation. This indicates that the OLS

regressions are slightly biased for close mixed-gender elections. Furthermore, the triangular kernel is for the constituency variables always more extreme, while this is not always the case for the district variables. This indicates that the effect is bigger within the constituencies when the female candidate won with a smaller relative vote margin.

6. Additional analysis

Before discussing the robustness checks of the main results, the analysis will be extended with two additional outcome variables. These two outcome variables measure the share of female candidates within a district without and with a former job in politics and are respectively referred to as share new district and share former district. This allows for studying whether the election of a female MP affects potential candidates with and without previous political experience differently within her district – i.e. whether this effect is heterogeneous. For example, when a female MP is considered to be a role model, she could affect females who do not already have a job in politics relatively more. These two outcome variables are constructed using the job category classifications provided by the German government, their formal definitions can be found in section 4. The structure of this additional analysis section is similar to the structure of section 5. Subsection 6.1 reviews the descriptive statistics of the outcome variables and the scatterplots. Hereafter, in subsection 6.2, a first indication of the results using OLS is briefly discussed. The last subsection, namely 6.3, discusses the results of the RDD with a 10% bandwidth, which is the main estimation strategy.

6.1 Data

The means of the two outcome variables are presented in Table 5, for both the full and close mixed-gender election samples and the treatment and control groups. In section 5, the group with the highest mean of the share district variable differed between the two samples. Surprisingly, this is not the case for the two outcome variables in Table 5, which together make up the share district variable. The mean of the share new district variable is in both samples higher for the treatment group. In contrast, the mean of the share former district variable is higher for the control group. These two facts suggest that a female MP increases the share of female candidates without a previous job in politics, but decreases the proportion of female candidates with a former job in politics.

Table 5: Descriptive statistics additional outcome variables of treatment and control groups

	All mixed- gender elections			Close mixed-gender elections		
	Full sample	Treatment	Control	Full sample	Treatment	Control
Share new district	0.171	0.175	0.168	0.177	0.180	0.174
Share former district	0.074	0.072	0.075	0.070	0.069	0.071
N	633	219	414	182	88	94

Remarks: this table shows the means of the two additional outcome variables $Y_{c,i,t+1}$ for both all mixed-gender elections and close mixed-gender elections. A mixed-gender election is defined as close if the relative vote margin $\delta_{c,i,t}$ (difference between the vote percentage for the two candidates with the highest percentage within a constituency divided by its sum) is within the range of -0.10 and 0.10. The means are rounded and shown for the whole sample as well as the treatment and control samples. The treatment sample consists of all elections within a constituency in which a female won the election and a man lost the election. The control sample consists of all elections within a constituency in which a male won the election and a female lost the election. The N indicates the number of mixed-gender election races within the sample.

The scatterplot of the share new district variable, presented in Figure 7.a, suggests a slightly positive effect of a female MP, while the scatterplot of the share former district variable, presented in Figure 7.b, suggests a slightly negative effect of a female MP. It needs to be noted that the axis range of the outcome variables differs. However, the functional form of both variables cannot be identified using these graphs. The scatterplots for the close-mixed gender election sample are presented in Appendix A.2. Regardless of the smaller bandwidth, the conclusions about the true form of the relationship and the effect are similar. However, for close races, the graphs suggest a linear relationship, but this cannot be stated with certainty.

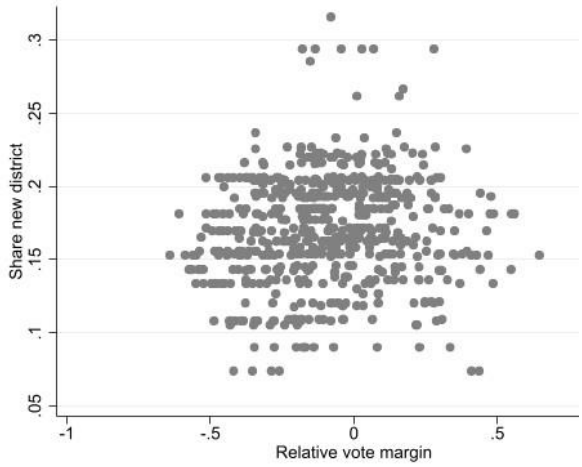

Figure 7.a: scatterplot of full sample, share new district

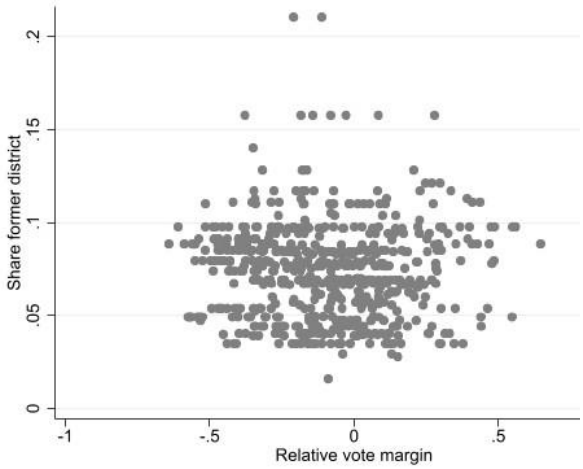

Figure 7.b: scatterplot of full sample, share former district

6.2 Ordinary Least Squares

The results of the OLS regressions are presented in Table 6, for various bandwidths. Even though these results indicate the presence of effects, these cannot be causally interpreted since it is likely that the estimates are biased due to the usage of OLS and its problem of omitted variables. While in Table 3 none of the estimates of the share district variable are significant, some of the estimates of the share new

district variable are significant. To be more precise, the estimates are significant for the full, 7.5% and 5% samples, and are insignificant for the 15% and 10% samples. According to these significant estimates, the election of a female candidate results in 0.7%, 1.1% and 1.5% more female candidates without previous political experience for respectively the full, 7.5% and 5% sample. Even though these percentages are small, one could still argue that the results are economically significant due to the large societal benefits of an increase in political female representation, as mentioned in section 3. For the share former district variable, the effects are all insignificant. The standard errors of the seven insignificant effects are, however, not small enough to certainly interpret this as a zero effect. Even though some of the estimates are insignificant, the sign of the effect is positive for all candidates in the next election without a previous job in politics, while it is negative for all candidates with a former job in politics, which is in line with subsection 6.1. This could explain the insignificant effect of a female MP on the share of female candidates in a district in the main results since this variable combines these two effects with opposite signs.

Table 6: Effects elected female in mixed-gender elections on additional $Y_{c,t}$ using OLS

	Full sample	**15%**	**10%**	**7.5%**	**5%**
Share new district	0.007**	0.004	0.006	0.011*	0.015*
	(0.003)	(0.004)	(0.005)	(0.006)	(0.008)
Share former district	-0.003	-0.002	-0.002	-0.005	-0.007
	(0.002)	(0.003)	(0.003)	(0.004)	(0.006)
N	633	275	182	133	88

Remarks: this table shows the OLS regression estimates of the effect of an elected female on two additional outcome variables for multiple samples. A mixed-gender election is included in the sample if the relative vote margin of the elected candidate is lower than or equal to the indicated percentage. The robust standard error is in between brackets. All estimates and robust standard errors are rounded. N indicates the number of mixed-gender election races. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

6.3 Regression Discontinuity Design

To start with the main approach, the graphical analysis of the RDD is presented in Figures 8-9, in which the grey points represent data bins, as in section 5. For both outcome variables and polynomials, the figures show a jump at the threshold. Whether this jump is significant, will be examined with the RDD estimate results. Moreover, the jump for the share new district variable indicates a positive effect while the jump of the share former district variable indicates a negative effect, which is in line with subsections 6.1 and 6.2. In addition, it is still unclear what the true form of the relationship is.

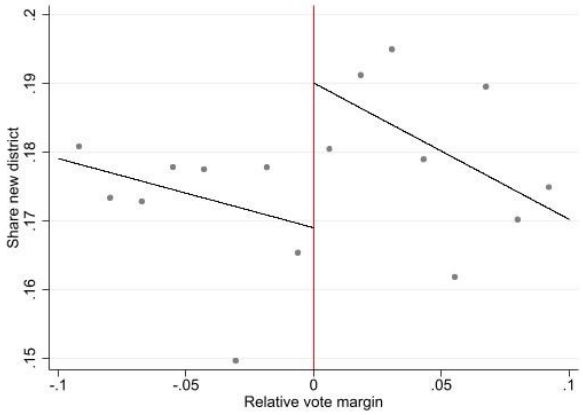

Figure 8.a: graphical representation RDD share new district, uniform kernel, 10% bandwidth, order 1

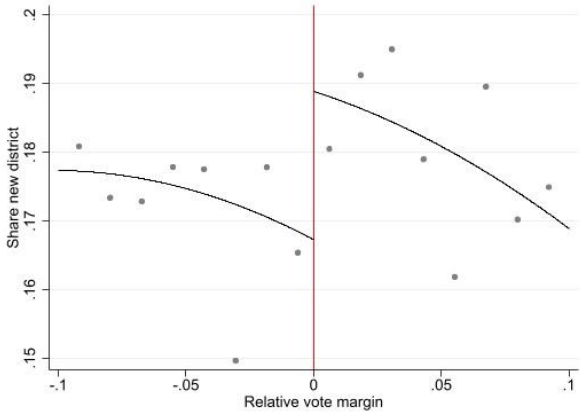

Figure 8.b: graphical representation RDD share new district, uniform kernel, 10% bandwidth, order 2

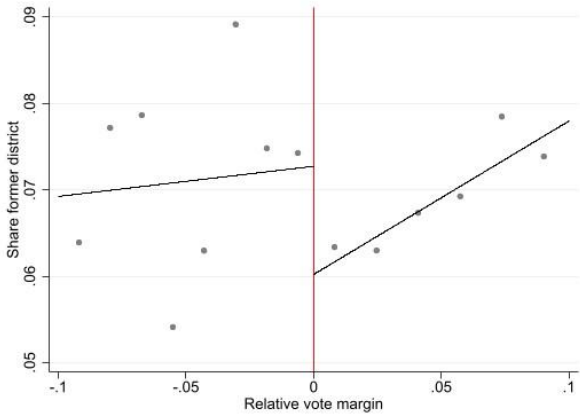

Figure 9.a: graphical representation RDD share former district, uniform kernel, 10% bandwidth, order 1

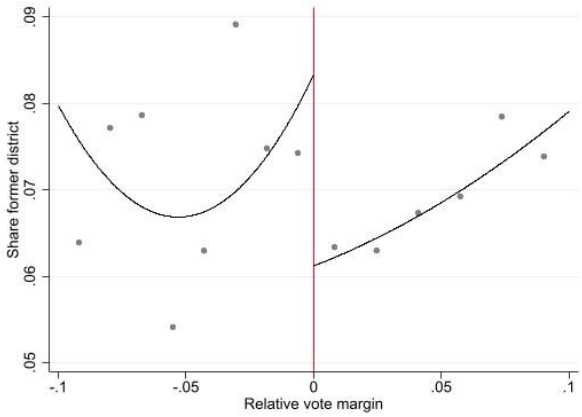

Figure 9.b: graphical representation RDD share former district, uniform kernel, 10% bandwidth, order 2

In Table 7 the RDD results are presented, for which the uniform kernel corresponds to the graphical representation in Figures 8-9. To start with the share new district variable, which estimates a significant effect, at the 10% level, for the local linear regression, in which the election of a female increases the share of female candidates without a former job in politics in her district in the next election with 1.7%. Its triangular kernel is in line with this, with a magnitude of 2.1% instead of 1.7% and a significance level of 5% instead of 10%. This effect is thus slightly larger for female candidates who won with a smaller relative vote margin. In contrast, the effects for the same variable with a second-order polynomial are for both kernels insignificant. Since the true relationship form is unknown, it cannot certainly be determined whether the effect is zero or around 2%. The magnitude of the standard errors of the second-order polynomial indicates that both a zero-effect as well as a non-zero effect is imaginable. This in combination with the results of the local linear regression indicates that a female MP affects the number of female candidates without a former job in politics in her district likely with a magnitude of 2%, but this cannot be stated with full confidence.

The other component of the share district variable is the share former district variable. The results in Table 7 indicate a significant negative effect of a female MP on the number of female candidates with a former job in politics, thus including the female MP herself, in her district in the next election. Except for the local linear regression with a uniform kernel, which is significant at the 10% level, the estimates are all significant at the 5% significance level. According to these estimates, the election of a female candidate instead of a male candidate decreases the number of female candidates with a previous job in politics in her district by 1.3% and 2.2% for respectively the uniform first- and second-order polynomial. The decrease estimated with a triangular kernel is slightly higher, respectively 1.6% and 2.4%. Therefore, the effect on share former district is between -1.3% and -2.4%. This effect is expected to be even more negative when candidates who run for re-election are excluded since it is likely that a candidate who won the election race runs again.

Thus, the effect on female candidates without a previous job in politics is likely positive, while the effect on female candidates with a job is negative. Therefore, the effect of a female MP is heterogeneous for these two subgroups. Moreover, the significance of these effects is in contrast to the significance of the share district variable in Table 4, which is made up of the two variables, since this variable is insignificant. This can be explained by the fact that the effect consists of a positive and negative component, which cancels out when combined. It is imaginable that this phenomenon is also present in the other three variables in the main analysis.

Table 7: RDD of elected female in mixed-gender elections on additional $Y_{c,t}$ with 10% bandwidth

		Uniform		Triangular	
Outcome variable	**Order**	**Treatment estimate**	**S.E.**	**Treatment estimate**	**S.E.**
Share new district	1	0.017*	0.009	0.021**	0.010
Share new district	2	0.022	0.014	0.019	0.014
Share former district	1	-0.013*	0.007	-0.016**	0.007
Share former district	2	-0.022**	0.10	-0.024**	0.010

Remarks: this table shows the RDD regression estimates and standard errors of the effect of an elected female in a close mixed-gender election on two additional outcome variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most 10%. These are performed for both a uniform and triangular kernel, as well as for the regression forms with an order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The total number of mixed-gender election races is 182, with 94 observations below the cut-off and 88 after the cut-off. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

As with the RDD in the main results, the results are only valid when the three assumptions hold. In the next section, two of these assumptions and an additional significance check are examined for the main results. Since the first assumption is on the running variable, which is the same running variable for the main and additional results, section 7.1 also holds for the additional results. In contrast, the tests for assumption 3 and the additional significance check do not hold for the additional analysis. Therefore, the results of the same tests are presented in appendices A.3 and A.4. These results do provide evidence

for the third assumption to hold for the additional analysis and are in line with the significance of the effects.

7. Robustness

In this section several robustness checks are performed and discussed, to test whether the RDD assumptions hold and thus if the RDD results are valid. Subsection 7.1 examines assumption 1, which states that there is no manipulation of the running variable- i.e. the relative vote margin. This is examined using a visual inspection of a histogram of the running variable, a density test and a balance test. Hereafter, subsection 7.2 discusses the validity of assumption 3, which states that the relationship between the running and outcome variable can be adequately modelled if the sample size requires moving further away from the threshold. Assumption 3 is reviewed by presenting the results for various bandwidths. In the last subsection, namely 7.3, a RDD with a CER-bandwidth selector, which is designed for causality, is discussed to check the validity of the significance levels.

7.1 Assumption 1: no manipulation

For the main estimation strategy to be valid, the relative vote margin may not be *precisely* manipulated. It is expected that the relative vote margin cannot be fully and precisely manipulated in the German democratic setting. Nevertheless, it is still important to provide factual evidence that this assumption holds. As a first indication, a histogram of the relative vote margin of the female candidate in the mixed-gender elections is analysed. This histogram, presented in Figure 1, does not indicate manipulation of the running variable because there is no suspicious density increase or decrease around the threshold. To also statistically test whether the running variable is likely manipulated, the density test by Cattaneo et al. (2018) is performed, for which the graphical representation can be seen in Figure 10. In this figure, the bars represent the density of the running variable, the lines represent the smoothed density estimator and the corresponding shaded areas represent the confidence intervals. This graphical representation suggests that there is no discontinuity of the running variables' density since a large part of the confidence intervals overlap at the threshold and since the smoothed density estimates are close to each other at the threshold. The p-values of this density test, which can be seen in Table 7, indicate the probability of no manipulation of the running variable. The conclusions of the visual representation of the density test are in line with the p-values of the test since none of the p-values is significant. Thus, for all mixed-gender elections and all examined smaller samples, including the main bandwidth of 10%, there is no statistical evidence for manipulation of the running variable.

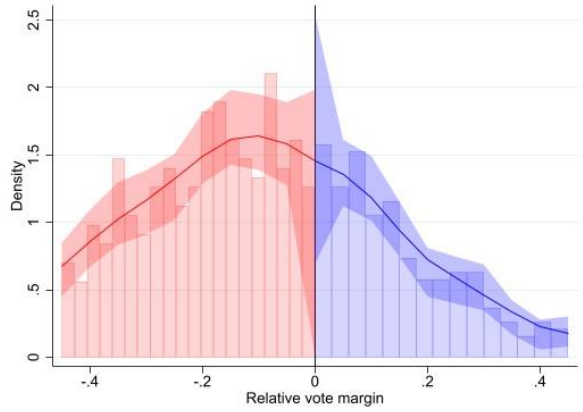

Figure 10: Density test Cattaneo et al.'s plot of running variable relative vote margin, $\delta_{c.t}$

Table 7: P-values of Cattaneo et al.'s density test

Bandwidth	Whole sample	15%	10%	7.5%	5%
P-value	0.754	0.376	0.248	0.386	0.946

Remarks: this table shows the rounded p-values for the density test for several bandwidths of the mixed-gender election sample. This density test is developed by Cattaneo et al. (2018).

Furthermore, another way to statistically test this assumption is a balance test, in which a regression is run of candidates' characteristics on a dummy indicating whether the candidate won the election. This indicates whether the predetermined characteristics of candidates are balanced before the close mixed-gender election took place, and therefore also illustrates whether the treatment and control groups are indeed randomly selected. As can be seen in Table 8, the characteristics of the winning candidates do significantly differ from the losing candidates in mixed-gender elections for the total, 20% and 15% sample. Because of this, the assumption does likely not hold for these samples, since the predetermined characteristics are not balanced and therefore the treatment and control group are non-random. Especially since the gender variable, which determines the treatment partly and is the studied subject, is not randomly distributed within the mixed-gender elections in these samples. For example, being female in the total, 20% and 15% sample decreases the probability of winning the election significantly with respectively on average 30%, 16.8% and 10%. Moreover, the characteristic of a former job in politics does statistically differ between the winners and their runners-up for the total, 20% and 15% sample, when the previous job of the candidate was in politics, the probability of winning increased respectively on average with 19.7%, 11% and 9.2%. Since statistical evidence suggests that the assumption does not hold for the samples with a bandwidth of at least 15%, the RDD design is not suitable for these bandwidths. This also decreases the validity of the estimates of the OLS regressions on the total and 15%-bandwidth sample in Table 3. Even when these characteristics are included in the regressions as a control variable, the samples are likely unbalanced with respect to other variables.

In contrast, all estimates of the smaller samples are insignificant, which implies that these characteristics do not statistically differ between the winning and runner-up groups. This is a good sign since it indicates that the winners and runners-up are randomly selected, and therefore indicates that the treatment and

control groups are randomly allocated. It also indicates that the assumption holds with respect to these three important characteristics for the smaller samples, including the sample that is used in the main estimation strategy. Therefore, this assumption holds for the main RDD strategy in Table 4, with a 10% bandwidth, and is in addition likely representative. Thus statistical evidence is in line with the expectation regarding the manipulation of the running variable.

It needs to be mentioned that this balanced test could be improved since not all characteristics are included because they were not available or did not have a statistically meaningful sample size. For example, education level, campaign budget and ethnicity could influence the probability of winning the mixed-gender election. Even though not all characteristics are included in the balance test, three of the most important characteristics are included and do provide evidence for the 10% and 7.5% samples to be balanced.

Table 8: Regression results of balance test

	Total sample	**20%**	**15%**	**10%**	**7.5%**	**5%**
Gender	-0.300***	-0.168***	-0.010**	-0.038	-0.042	-0.009
	(0.026)	(0.037)	(0.042)	(0.053)	(0.062)	(0.076)
Year of birth	0.001	0.004**	0.004*	0.003	0.004	0.006
	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Former job in politics	0.197***	0.110***	0.092**	0.049	-0.039	-0.073
	(0.027)	(0.037)	(0.043)	(0.053)	(0.062)	(0.077)
N	1266	696	550	364	266	176

Remarks: this table shows the results of a regression of available characteristics of candidates on a dummy indicating whether the candidate won or lost the mixed-gender election. The robust standard errors are in between brackets. Both the estimates and robust standard errors are rounded. Each column is on a different sample. Candidates are included in the sample if the winning candidate won with a relative vote margin of at most the percentage displayed in the column header. Gender is a dummy indicating whether the candidate is a female. The dummy former job in politics is 1 if the German government defined the job category of the candidate as 71 or 76, which includes for example a member of the Bundestag or a mayor. N indicates the number of candidates in the sample.

7.2 Assumption 3: bandwidths

The third assumption, which states that the relationship between the running and outcome variable can be adequately modelled if the sample size requires moving further away from the threshold, will be examined by presenting the RDD results for various manual and automated selected bandwidths. To start, the manually selected bandwidth estimates can be found in Table 9. Even though the samples with a 20% and 15% bandwidth are unbalanced, the estimates do provide evidence for this assumption to hold. To be more precise, the results are in line with the conclusions of the main results since the significance levels are mostly similar to the significance levels presented in Table 4. In addition, the magnitudes of the significant effects are similar to the corresponding ones in the main results.

A smaller bandwidth for which the sample is likely balanced as well as large enough is the 7.5% bandwidth. These results will therefore be discussed more extensively. First, is the discussion of the share constituency variable. For both polynomial orders, the significance does not change. However, the significance level and the estimate for the second level polynomial do change. While for the main results the estimated effect is 7.3% with a significance level of 10%, the 7.5% bandwidth results estimate an effect that is more than twice as high, namely 18.3%, with a significance level of 5%. This provides evidence for a non-zero effect, although the effect is still insignificant for the first-order polynomial. Overall, this suggests that for this variable the choice of bandwidth does slightly affect the results, and these results must therefore be interpreted carefully.

Second, is the discussion of the absolute constituency variable. The first-order polynomial is not robust to the choice of bandwidth, since the significance level differs between the 10% and 7.5% bandwidth. For the 10% bandwidth, the effect is insignificant while it is significant at the 10% level for the 7.5% bandwidth. The election of a woman who won with a relative vote margin of at most 7.5%, decreases the absolute number of female candidates within her constituency in the next election by on average 2.2. Even though the significance varies with the bandwidth in this case, the significant result is, however, in line with the main results' conclusion. Since the main results conclude that this effect is significant and since the magnitude estimated with the 7.5% bandwidth falls within the conclusions' range. In contrast, the estimate with the second-order polynomial is robust to a smaller bandwidth, even though the significance level changes, and thus provides evidence for the assumption to hold.

Third, is the discussion of the first district variable, the share of female candidates. Even though the magnitude of the effects slightly differs, the results are robust to the bandwidth choice, since all estimates are still insignificant. It is therefore likely that the effect on the proportion of female candidates is unaffected by the election of a female MP, and provides evidence for the assumption to hold.

Fourth, is the discussion of the absolute district variable. The estimates for this variable are sensitive to the choice of bandwidth. For instance, the local linear regression estimate is insignificant with a 10% bandwidth, while it is significant for the 7.5% bandwidth, at the 10% level, with a negative effect of an elected female on the number of female candidates in her district. To be more specific, a female MP decreases the number of female candidates by 30.1. Contradictory, for the second-order polynomial, the estimate is significant at the 10% level but is insignificant at the 7.5% level. Therefore, the estimates are not robust to the bandwidth length but the conclusions are robust to the bandwidth choice, just as for the share constituency variable.

Table 9: RDD of female MP on $Y_{c,t}$ with uniform kernel and manually-selected bandwidth

Outcome variable	Order	20%	15%	7.5%
Share constituency	1	-0.015	-0.014	0.012
		(0.032)	(0.037)	(0.054)
Share constituency	2	0.017	0.050	0.183**
		(0.049)	(0.057	(0.079)
Absolute constituency	1	-0.592	-0.769	-2.195*
		(0.632)	(0.707)	(1.174)
Absolute constituency	2	-1.701*	-2.138*	-3.388*
		(1.001)	(1.204)	(1.991)
Share district	1	0.004	0.008	0.006
		(0.008)	(0.009)	(0.013)
Share district	2	0.013	0.010	-0.005
		(0.012)	(0.014)	(0.020)
Absolute district	1	-15.829	-11.824	-30.079*
		(9.901)	(11.317)	(16.529)
Absolute district	2	-8.861	-17.031	-15.421
		(15.407)	(17.847)	(28.878)
N within bandwidth		348	275	133

Remarks: this table shows the RDD regression estimates and standard errors of the effect of an elected female in a close mixed-gender election on several outcome variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most the percentage indicated in the column header. The kernel is uniform. These regressions are performed for different forms with a polynomial of the order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The total number of mixed-gender election races is stated in the last row. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

In addition to testing the validity of the third assumption using manually selected bandwidths, the validity can also be tested by using automatically selected bandwidths. As discussed in section 4.3, this automatically bandwidth selector is the MSE-selector and is used to select bandwidths with equal lengths on both sides of the threshold in Table 10, as well as it is used to select bandwidths with different lengths on both sides of the threshold in Table 11. Interestingly, the automated selected bandwidths are almost all wider than the manually selected bandwidth of 10%. However, the manual selected bandwidth is likely more suitable for estimating a causal effect since the sample of 10%, since the balance test in subsection 7.2 shows that the samples with a bandwidth of 15% and 20% are unbalanced, while the 10% is balanced. Nevertheless, the significance of the estimates is in line with, and thus provides evidence for, the conclusions drawn in section 5. To be more specific, all estimates are insignificant except for some estimates of the absolute constituency variable. This variable is the only one for which it is concluded that there is a significant effect in section 5. Moreover, these significant estimates do all fall within the range of the estimated effect. Furthermore, the estimates provide more evidence for a zero effect than for a small effect, for the share constituency and absolute district variables.

To conclude, both the manual and the automatically selected bandwidths are in general in line with the results and conclusions of the RDD presented in section 5. Since most of the estimates and their significance are in line with the results or conclusion, the different bandwidths in Table 9-11 do provide evidence for the third assumption to hold. Thus, the running variable can be adequately modelled.

Table 10: RDD of elected female in mixed-gender elections on $Y_{c,t}$ with MSE-selected bandwidth

Outcome variable	Order	Uniform			Triangular		
		Bandwidth	Treatment estimate	S.E.	Bandwidth	Treatment estimate	S.E.
Share constituency	1	0.140	-0.006	0.038	0.120	0.028	0.046
Share constituency	2	0.151	0.050	0.057	0.156	0.064	0.058
Absolute constituency	1	0.160	-0.867	0.700	0.108	-1.862*	1.050
Absolute constituency	2	0.158	-2.060*	1.186	0.186	-2.168*	1.185
Share district	1	0.179	0.005	0.009	0.178	0.008	0.009
Share district	2	0.155	0.007	0.014	0.241	0.010	0.011
Absolute district	1	0.155	-11.305	11.165	0.177	-12.832	11.563
Absolute district	2	0.189	-7.778	15.368	0.257	-17.969	14.545

Remarks: this table shows the RDD regression bandwidths, estimates and standard errors of the effect of an elected female in a close mixed-gender election on several outcome variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most the percentage indicated in the bandwidth. The bandwidths are determined using the MSE bandwidth selector. These are performed for both a uniform and triangular kernel, as well as for the regression forms with an order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The number of mixed-gender election races differs for each bandwidth. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

Table 11: RDD of elected female in mixed-gender elections on $Y_{c,t}$ with different bandwidths on both sides, MSE-selected and uniform kernel

Outcome variable	Order	Bandwidth		Treatment estimate	S.E.
		Left	Right		
Share constituency	1	0.190	0.117	-0.007	0.0359
Share constituency	2	0.158	0.192	0.0331	0.055
Absolute constituency	1	0.114	0.091	-1.440	0.883
Absolute constituency	2	0.230	0.102	-3.105***	0.991
Share district	1	0.172	0.172	0.007	0.009
Share district	2	0.164	0.174	0.011	0.013
Absolute district	1	0.111	0.104	-7.997	13.133
Absolute district	2	0.218	0.196	-8.604	14.984

Remarks: this table shows the RDD regression bandwidths, estimates and standard errors of the effect of an elected female in a close mixed-gender election on several outcome variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most the percentage indicated in the right bandwidth or when a candidate lost with a relative vote margin of at most the percentage in the left bandwidth. The bandwidths are determined using the MSE bandwidth selector. These are performed for a uniform kernel and for the regression forms with an order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The number of observations differs for each bandwidth. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

7.3 Regression Discontinuity Design CER-bandwidth selector

In this last subsection, RDD estimates using the CER-bandwidth selector, developed by Cattaneo et al. (2020), are examined. The results are presented in Table 12. As already mentioned in section 4, the CER-bandwidth selector is designed for examining causal inference, and therefore, the magnitude of the estimates does not represent the magnitude of the treatment effect accurately. As with the MSE-bandwidth selectors, the bandwidths are larger than the main bandwidth of 10%. The CER-bandwidths are however smaller than the MSE-bandwidths and are therefore more likely to pass the balance test. As can be seen in Table 12, all estimates with a uniform kernel are insignificant, suggesting that there is no

effect of an elected female MP on the proportion and number of female candidates in the next election. These insignificant results do provide evidence for the effect on the share constituency and absolute district variables to be zero instead of small and do confirm the conclusion of an insignificant effect for the share district variable. In contrast, it is not in line with the conclusions about the absolute constituency variable, since the CER-bandwidth selector shows an insignificant effect and the RDD with a 10% bandwidth shows a significant effect. This result does thus question the validity of the main result with respect to the absolute constituency variable. However, the estimated significance for the local linear regression of this variable with a triangular kernel is significant at the 5% level, providing evidence for the main conclusions. The other triangular kernel estimates are in line with the uniform kernel estimates and do thus support the main results.

Table 12: RDD of elected female in mixed-gender elections on $Y_{c,t}$ with CER-selected bandwidth

		Uniform			Triangular		
Outcome variable	**Order**	**Bandwidth**	**Treatment estimate**	**S.E.**	**Bandwidth**	**Treatment estimate**	**S.E.**
Share constituency	1	0.101	0.019	0.046	0.087	0.049	0.053
Share constituency	2	0.104	0.078	0.068	0.108	0.110	0.068
Absolute constituency	1	0.116	-1.055	0.852	0.078	-2.540**	1.295
Absolute constituency	2	0.109	-2.981	1.514	0.129	-3.005	1.517
Share district	1	0.129	0.011	0.010	0.129	0.007	0.010
Share district	2	0.108	0.003	0.016	0.167	0.005	0.014
Absolute district	1	0.112	-8.678	12.833	0.128	-16.550	13.877
Absolute district	2	0.130	-32.414	19.327	0.178	-21.005	17.927

Remarks: this table shows the bandwidths, estimates and standard errors of a RDD regression on the effect of an elected female in a close mixed-gender election on several outcome variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most the percentage indicated in the bandwidth. The bandwidths are determined using the CER bandwidth selector. These are performed for both a uniform and triangular kernel, as well as for the regression forms with an order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The number of mixed-gender election races differs for each bandwidth. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

8. Discussion

This discussion section starts with a review of the results. Hereafter, the presence of mechanisms, the hypothesis and the relation to related literature are examined. Afterwards, the limitations of this study are discussed, after which there is a synopsis of the results.

In this thesis, the effects of the election of a female member of parliament (MP) on the proportion and number of female candidates in the next election, within her constituency and district, are examined for Germany. When a positive effect is present, female political representation can increase, which is highly beneficial to society. To answer this research question, data on the characteristics of the candidates and their election outcomes are matched to create a dataset of all mixed-gender election races in the period 2002-2017. Moreover, the main estimation strategy is a regression discontinuity design (RDD) and is due to its random nature able to estimate causal effects. To start with the effects within the constituencies, the election of a female instead of a male does not or only modestly affect the share of

female candidates, but likely affects the absolute number of female candidates in her constituency in the next election. To be more precise, a female MP decreases the number of female candidates within her constituency by 1 to 3.5 candidates. Interestingly, these effects are both larger for female MPs who won with a relatively smaller vote margin. To continue with the effects within districts, which consist of multiple constituencies, a female MP does not affect the share of female candidates and does not or only modestly affect the number of female candidates. Thus, the effect on districts is economically insignificant. However, when separating the proportion into the proportion of female candidates with and without a former job in politics in the district, the results do become significant. More specifically, the election of a female MP increases the share of female candidates without a former job in politics by around 2% and decreases the share of female candidates with a former job in politics between 1.3% and 2.4%. Further, most of the results are in line with the OLS estimation, but not all are, providing evidence for the existence of omitted variables.

The significant effects can be explained by one or more of the four mechanisms mentioned in section 3.1, but it cannot be determined what their individual influence is. However, if the incumbency effect were present, the negative effects would be even larger when the first-vote incumbents are excluded from the number of candidates, except in the event of an incumbency disadvantage. Furthermore, it is probable that mechanisms are actually present for the insignificant effects, but work in opposite directions and are therefore cancelling each other out. The difference in the sign of these effects implies three things. First, it explains the insignificance of the share within districts due to the effects cancelling each other out. Second, it implies a heterogenous effect for these two subgroups within districts. Third, it suggests that the effect could be heterogeneous for the effects within a constituency as well, which could therefore also change its significance. Since in a district, the effect is positive for females without a former job in politics, and negative for females with a female job in politics, one could argue that this provides evidence for the role model effect since one would expect women outside of politics to be affected relatively stronger by a female MP.

These results are surprising since the insignificant and negative effects are not in line with the hypothesis, which expected a total positive effect. In contrast, the positive significant result in the subgroup of female candidates without a previous job in politics in districts is in line with the hypothesis. In comparison to the related literature, all insignificant effects, implying a zero-effect, are similar to the results for the United States and the United Kingdom (Broockman, 2014; Valentim, 2021), but not to the positive effects measured in India and Poland (Bhalotra et al., 2017; Jankowski et al., 2019). Moreover, the significant negative results on the absolute number of female candidates are not in line with any of the four studies and are thus surprising. The differences between the results of this paper and the hypothesis and related literature can be explained by the differences in study and setting. The most important explanation is the country of study, which is also the main difference between this thesis and the related literature. One difference is the German political system with, for example, first and

second votes, and the constituency and district. Another difference is the data period, in which the German chancellor was in the studied period a woman for four out of five elections. She may be considered a role model, decreasing the impact of the female MP. Also the other discussed differences, in section 3.2, could explain the differences. However, further research is needed to explain the differences between the studies precisely.

Increasing female political representation is beneficial for society, especially since half of the society is female. As literature shows, increasing female political representation decreases corruption and positively affects the exact policy choices and parliamentary deliberations (Hessami & da Fonseca, 2020), and thus shows the importance to increase female political representation for policymakers. As the results indicate a total negative effect of the election of a female MP on the absolute number of female candidates within her constituency and a zero or small effect on the other outcome variables, policies should be adjusted or created to minimalize or even positively shift these effects. In addition, policymakers could consider using the available resources for future research or policies in other settings where women are underrepresented, to allocate resources efficiently. For example, policies in other settings in Germany could be in line with the suggested policies by Baskaran and Hessami (2018) since female representation increases in their local government setting. Further, the additional analysis implied that a female MP in districts significantly decreases the share of female candidates with a former job in politics. It is important to study what happens to these women – i.e. why do they not run for first vote re-election – since this information could be used to shape policy aimed at increasing female representation. For example, when females with a previous job in politics do not run for re-election through the first votes but do run for the list elections, which is a likely scenario, the effect of a female MP could be in fact positive and should be exploited instead of minimalized by policymakers. For this exploitation, it is important to also study the mechanisms, to allocate the available resources as efficiently as possible. Studying the negative effect on females with a former job in politics is especially important for policymaking since the additional results do also imply a positive effect on female candidates without a former job in politics, which does increase female political representation. Moreover, as already mentioned, it is likely that the heterogenous effect is also present for the constituencies, but this cannot be stated with full confidence. If further research does confirm this statement, the policy advice for constituencies is similar to the above-mentioned district policy advice. To conclude, policymakers should adjust their policies to minimalize and positively affect the negative and zero effects and should study the heterogeneous effects and mechanisms more extensively to adjust their policies and hereby allocate their resources most efficiently.

Unlike most other estimation methods, the results from the RDD presented in this thesis can be interpreted causally for close mixed-gender election races. However, for the RDD to be valid, three assumptions must hold. Section 7 provided evidence for the first and third assumptions to hold, which respectively state that there is no manipulation of the running variable -i.e. the relative vote margin –

and that the running variable can be adequately modelled when moving further away from the threshold when this is required due to the sample size. However, due to data limitations, no statistical evidence could be presented for the second assumption, which states that nothing else changes at the threshold except for the treatment. This assumption likely holds, as already argued in section 4.4, the treatment only changes the winning and therefore job status of the candidate. Ideally, a placebo test would be provided for evidence. Thus, the assumptions of the RDD do likely hold, and the results can therefore be interpreted causally.

To summarize, this thesis estimated the effects of a female MP on the proportion and number of female candidates in the next election, in both her constituency and district, using a RDD. No significant effect is found, except for a negative decrease in the number of female candidates within her constituency. However, for the insignificant effect on the proportion of female candidates within her district, a heterogenous significant effect on the subgroups of female candidates with and without a previous job in politics is found, which affects these groups respectively negatively and positively.

9. Conclusion

In this thesis, the effects of a female member of parliament (MP) in Germany on the proportion and number of female candidates in the next election within her constituency and district are examined. This research question is answered using a regression discontinuity design (RDD) as the main estimation strategy. An expansion of female political representation occurs when a positive effect would be estimated, which is highly beneficial to society. For example, public policies will be positively adjusted and corruption will decrease due to more female representation in politics. This positive effect can be caused due to four mechanisms; the empowerment effect, the role model effect, the contagion effect and the incumbency effect. In contrast to the expected positive effect mentioned in the hypothesis, no significant effect is found, except for a negative decrease in the number of female candidates within her constituency. First, the effects within the constituencies, a female instead of a male MP does not or only modestly affect the share of female candidates, but this effect is negative for the absolute number of female candidates in her constituency in the next election. Specifically, a female MP decreases the number of female candidates by 1 to 3.5 candidates. To continue with the effects within districts, the election of a female MP does also not affect the share and absolute number of female candidates in her district in the next election. For the absolute number, this could also be a modest effect. Thus, the effect of a female MP on the number of female candidates in her district is not economically significant. Interestingly, when the insignificant proportion is separated into the proportion of female candidates with and without a former job in politics in the district, the effects are significant. More specifically, the election of a female MP increases the share of female candidates without a former job in politics by around 2% and decreases the share of female candidates with a former job in politics between 1.3% and

2.4%. This implies that the effects within constituencies could also be significant when estimating the effect for the same subgroups.

From these findings follow some interesting topics for future research. First, future studies ought to further explore the mechanisms underlying these zero findings, for instance, whether the results differ in and outside the party of the female member of parliament. This research should especially pay attention to the heterogenous effect on women with and without a previous job in politics. In line with this, future research could study the future career path of women with a former job in politics, since the results indicated a negative effect on this subgroup. In this way, policymakers would be better informed and could allocate their resources more efficiently. In addition, future research should examine whether the effect is indeed heterogeneous for the same subgroups in the constituencies, to specifically adjust policies on the constituency level. As a last suggestion for future research, researchers should consider examining the effect of a female MP on the rank and number of female candidates in the second votes instead of first votes – i.e. list elections within districts – to have more information on the effect of a female MP on the entire German federal elections.

10. References

Allik, M. (2015). Who stands in the way of women? Open vs. closed lists and candidate gender in Estonia. *East European Politics*, *31*(4), 429–451. https://doi.org/10.1080/21599165.2015.1084924

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Atkeson, L. R. (2003). Not All Cues Are Created Equal: The Conditional Impact of Female Candidates on Political Engagement. *The Journal of Politics*, *65*(4), 1040–1061. https://doi.org/10.1111/1468-2508.t01-1-00124

Baskaran, T., & Hessami, Z. (2018). Does the Election of a Female Leader Clear the Way for More Women in Politics? *American Economic Journal: Economic Policy*, *10*(3), 95–121. https://doi.org/10.1257/pol.20170045

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful Women: Does Exposure Reduce Bias?*. *Quarterly Journal of Economics*, *124*(4), 1497–1540. https://doi.org/10.1162/qjec.2009.124.4.1497

Bhalotra, S., Clots-Figueras, I., & Iyer, L. (2017). Pathbreakers? Women's Electoral Success and Future Political Participation. *The Economic Journal*, *128*(613), 1844–1878. https://doi.org/10.1111/ecoj.12492

Bhuller, M., Dahl, G. B., Løken, K. V., & Mogstad, M. (2020). Incarceration, Recidivism, and Employment. *Journal of Political Economy*, *128*(4), 1269–1324. https://doi.org/10.1086/705330

Bloemraad, I., & Schönwälder, K. (2013). Immigrant and Ethnic Minority Representation in Europe: Conceptual Challenges and Theoretical Approaches. *West European Politics*, *36*(3), 564–579. https://doi.org/10.1080/01402382.2013.773724

Broockman, D. E. (2014). Do female politicians empower women to vote or run for office? A regression discontinuity approach. *Electoral Studies*, *34*, 190–204. https://doi.org/10.1016/j.electstud.2013.10.002

Bühlmann, M., & Schädel, L. (2012). REPRESENTATION MATTERS: THE IMPACT OF DESCRIPTIVE WOMEN'S REPRESENTATION ON THE POLITICAL INVOLVEMENT OF WOMEN. *Representation*, *48*(1), 101–114. https://doi.org/10.1080/00344893.2012.653246

Cambridge Dictionary. (2022, June 22). *incumbent definition*. Retrieved 28 June 2022, from https://dictionary.cambridge.org/dictionary/english/incumbent

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, *82*(6), 2295–2326. https://doi.org/10.3982/ecta11757

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2020). *A Practical Introduction to Regression Discontinuity Designs*. Cambridge University Press.

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation Testing Based on Density Discontinuity. *The Stata Journal: Promoting Communications on Statistics and Stata*, *18*(1), 234–261. https://doi.org/10.1177/1536867x1801800115

Cheng, C., & Tavits, M. (2009). Informal Influences in Selecting Female Political Candidates. *Political Research Quarterly*, *64*(2), 460–471. https://doi.org/10.1177/1065912909349631

Childs, S., & Hughes, M. (2018). "Which Men?" How an Intersectional Perspective on Men and Masculinities Helps Explain Women's Political Underrepresentation. *Politics & Gender*, *14*(2), 282–287. https://doi.org/10.1017/s1743923x1800017x

Die Welt. (2022, February 10). *Umfrage: Frauenquote, Gendersprache? Führungskräfte sehen es kritisch*. Retrieved 4 May 2022, from https://www.welt.de/wirtschaft/article236816007/Umfrage-Frauenquote-Gendersprache-Fuehrungskraefte-sehen-es-kritisch.html

European Institute for Gender Equality. (2021). *Gender Equality Index | 2021 | Germany*. Retrieved 4 May 2022, from https://eige.europa.eu/gender-equality-index/2021/country/DE

Federal Ministry of the Interior and Community. (2017, October 5). *How Bundestag elections work*. Retrieved 2 May 2022, from https://www.bmi.bund.de/EN/topics/constitution/electoral-law/bundestag-elections/bundestag-elections-node.html;jsessionid=1B801A09D12D3B5DDF3E528185E217A6.1_cid295

Federal Ministry of the Interior and Community. (2018, June 6). *The voting system*. Retrieved 2 May 2022, from https://www.bmi.bund.de/EN/topics/constitution/electoral-law/voting-system/voting-system-node.html

Fiva, J. H., & Røhr, H. L. (2018). Climbing the ranks: incumbency effects in party-list systems. *European Economic Review*, *101*, 142–156. https://doi.org/10.1016/j.euroecorev.2017.09.011

Foos, F., & Gilardi, F. (2019). Does Exposure to Gender Role Models Increase Women's Political Ambition? A Field Experiment with Politicians. *Journal of Experimental Political Science*, *7*(3), 157–166. https://doi.org/10.1017/xps.2019.21

Fox, R. L., & Lawless, J. L. (2010). Gendered Perceptions and Political Candidacies: A Central Barrier to Women's Equality in Electoral Politics. *American Journal of Political Science*, *55*(1), 59–73. https://doi.org/10.1111/j.1540-5907.2010.00484.x

Fridkin, K. L., & Kenney, P. J. (2014). How the Gender of U.S. Senators Influences People's Understanding and Engagement in Politics. *The Journal of Politics*, *76*(4), 1017–1031. https://doi.org/10.1017/s0022381614000589

Gelman, A., & Imbens, G. (2018). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, *37*(3), 447–456. https://doi.org/10.1080/07350015.2017.1366909

German Federal Foreign Office. (2022, January 21). *Women and gender equality*. Retrieved 4 May 2022, from https://www.auswaertiges-amt.de/en/aussenpolitik/themen/menschenrechte/05-frauen/-/227616

Gilardi, F. (2015). The Temporary Importance of Role Models for Women's Political Representation. *American Journal of Political Science*, *59*(4), 957–970. https://doi.org/10.1111/ajps.12155

Heathcote, J., Storesletten, K., & Violante, G. L. (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, *118*(4), 681–722. https://doi.org/10.1086/656632

Hahn, J., Todd, P., & Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209. https://doi.org/10.1111/1468-0262.00183

Hessami, Z., & da Fonseca, M. L. (2020). Female political representation and substantive effects on policies: A literature review. *European Journal of Political Economy*, *63*, 101896. https://doi.org/10.1016/j.ejpoleco.2020.101896

Hopkins, D. J., & McCabe, K. T. (2012). After It's Too Late. *American Politics Research*, *40*(4), 665–700. https://doi.org/10.1177/1532673x11432469

IPU Parline. (2022). *Data on women in national parliament*. Parline: The IPU's Open Data Platform. Retrieved 4 May 2022, from https://data.ipu.org/node/65/data-on-women?chamber_id=13316

Jankowski, M., Marcinkiewicz, K., & Gwiazda, A. (2019). The Effect of Electing Women on Future Female Candidate Selection Patterns: Findings from a Regression Discontinuity Design. *Politics & Gender*, *15*(2), 182–210. https://doi.org/10.1017/s1743923x19000096

Kantola, J., & Lombardo, E. (2019). Populism and feminist politics: The cases of Finland and Spain. *European Journal of Political Research*, *58*(4), 1108–1128. https://doi.org/10.1111/1475-6765.12333

Karp, J. A., & Banducci, S. A. (2008). When politics is not just a man's game: Women's representation and political engagement. *Electoral Studies*, *27*(1), 105–115. https://doi.org/10.1016/j.electstud.2007.11.009

Kittilson, M. C. (2010). Women, parties and platforms in post-industrial democracies. *Party Politics*, *17*(1), 66–92. https://doi.org/10.1177/1354068809361012

Ladam, C., Harden, J. J., & Windett, J. H. (2018). Prominent Role Models: High-Profile Female Politicians and the Emergence of Women as Candidates for Public Office. *American Journal of Political Science*, *62*(2), 369–381. https://doi.org/10.1111/ajps.12351

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281–355. https://doi.org/10.1257/jel.48.2.281

Matland, R. E., & Studlar, D. T. (1996). The Contagion of Women Candidates in Single-Member District and Proportional Representation Electoral Systems: Canada and Norway. *The Journal of Politics*, *58*(3), 707–733. https://doi.org/10.2307/2960439

Mattozzi, A., & Merlo, A. (2008). Political careers or career politicians? *Journal of Public Economics*, *92*(3–4), 597–608. https://doi.org/10.1016/j.jpubeco.2007.10.006

McGregor, R. M., Moore, A., Jackson, S., Bird, K., & Stephenson, L. B. (2017). Why So Few Women and Minorities in Local Politics?: Incumbency and Affinity Voting in Low Information Elections. *Representation*, *53*(2), 135–152. https://doi.org/10.1080/00344893.2017.1354909

Morais Maceira, H. (2017). Economic Benefits of Gender Equality in the EU. *Intereconomics*, *52*(3), 178–183. https://doi.org/10.1007/s10272-017-0669-4

Nye, J. V. C., Rainer, I., & Stratmann, T. (2014). Do Black Mayors Improve Black Relative to White Employment Outcomes? Evidence from Large US Cities. *Journal of Law, Economics, and Organization*, *31*(2), 383–430. https://doi.org/10.1093/jleo/ewu008

Palmer, B., & Simon, D. M. (2005). When Women Run Against Women: The Hidden Influence of Female Incumbents in Elections to the U.S. House of Representatives, 1956–2002. *Politics & Gender*, *1*(01). https://doi.org/10.1017/s1743923x05050026

The Federal Returning Officer. (2021). *Publications - The Federal Returning Officer* [Dataset]. https://www.bundeswahlleiter.de/en/bundestagswahlen/2021/publikationen.html

The World Bank. (n.d.). *Population, female (% of total population) - Germany | Data*. The World Bank Data. Retrieved 4 May 2022, from https://data.worldbank.org/indicator/SP.POP.TOTL.FE.ZS?locations=DE

The World Bank. (2021). *Proportion of seats held by women in national parliaments (%) | Data*. Retrieved 16 May 2022, from https://data.worldbank.org/indicator/SG.GEN.PARL.ZS

Valentim, A. (2021). The political effects of female representation: Evidence from a regression-discontinuity design. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3992109

Vogl, T. S. (2014). Race and the politics of close elections. *Journal of Public Economics*, *109*, 101–113. https://doi.org/10.1016/j.jpubeco.2013.11.004

Wolbrecht, C., & Campbell, D. E. (2007). Leading by Example: Female Members of Parliament as Political Role Models. *American Journal of Political Science*, *51*(4), 921–939. https://doi.org/10.1111/j.1540-5907.2007.00289.x

11. Appendix

A.1 Scatterplots main results

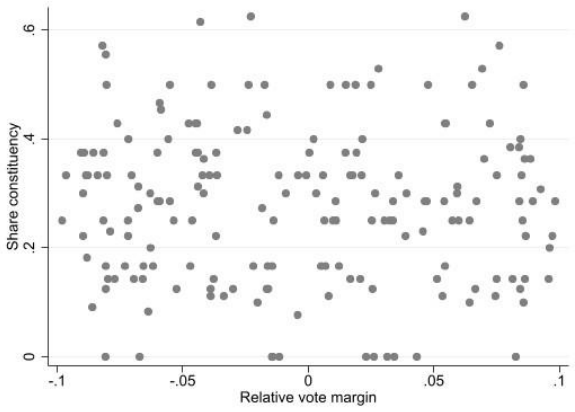

Figure A1.a: scatterplot of 10% sample, share constituency

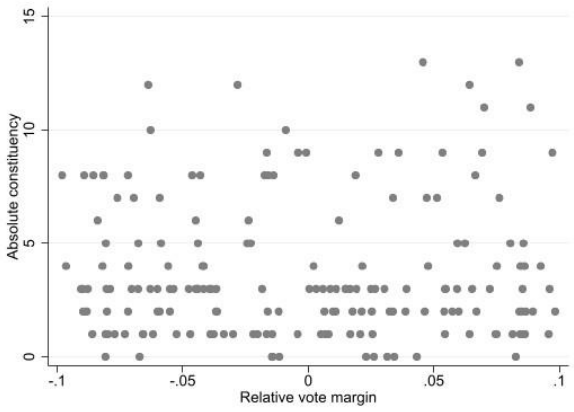

Figure A1.b: scatterplot of 10% sample, absolute constituency

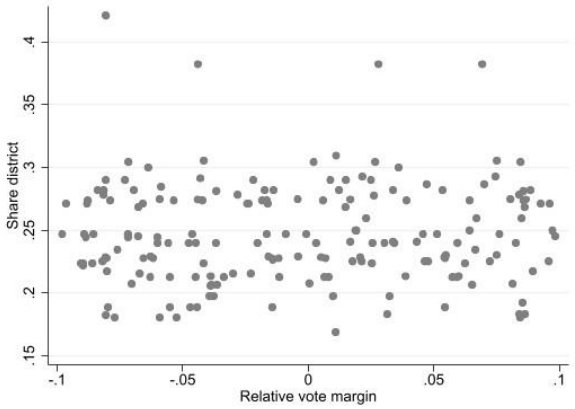

Figure A1.c: scatterplot of 10% sample, share district

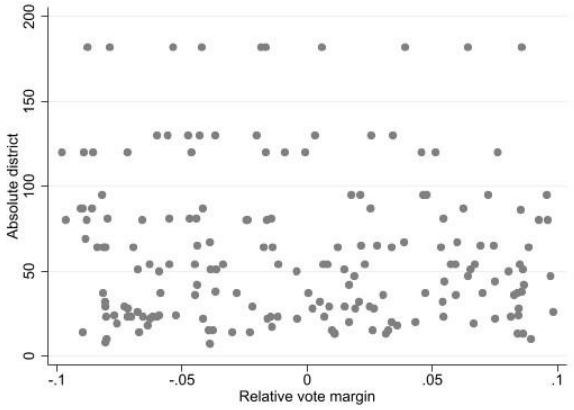

Figure A1.d: scatterplot of 10% sample, absolute district

A.2 Scatterplots additional results

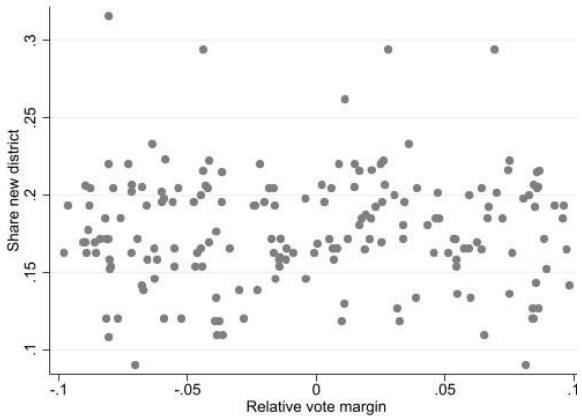

Figure A2.a: scatterplot of 10% sample, share new district

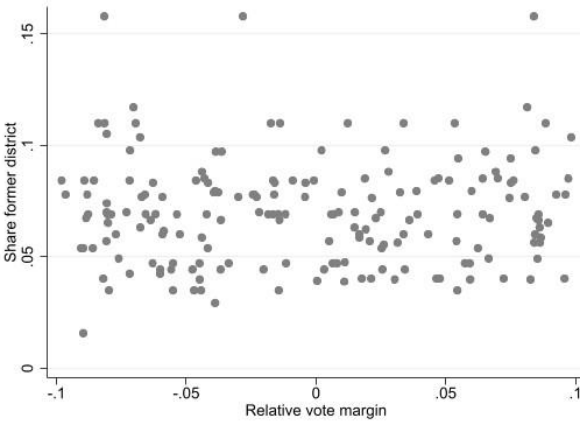

Figure A2.b: scatterplot of 10% sample, share former district

A.3 Robustness checks assumption 3 of additional analysis

Table A1: RDD of elected female in mixed-gender elections on additional $Y_{c,t}$ with uniform kernel and manually-selected bandwidth

Outcome variable	Order	20%	15%	7.5%
Share new district	1	0.007	0.015*	0.020*
		(0.007)	(0.008)	(0.011)
Share new district	2	0.027***	0.024**	0.026
		(0.010)	(0.012)	(0.017)
Share former district	1	-0.002	-0.008	-0.013*
		(0.005)	(0.005)	(0.008)
Share former district	2	-0.014*	-0.014*	-0.034***
		(0.007)	(0.008)	(0.012)
N within bandwidth		348	275	133

Remarks: this table shows the RDD regression estimates and standard errors of the effect of an elected female in a close mixed-gender election on the two outcome variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most the percentage indicated in the column header. The kernel is uniform. These regressions are performed for different forms with an polynomial of the order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The total number of mixed-gender election races is stated in the last row. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

Table A2: RDD of elected female in mixed-gender elections on additional $Y_{c,t}$ with MSE-selected bandwidth

		Uniform			Triangular		
Outcome variable	**Order**	**Bandwidth**	**Treatment estimate**	**S.E.**	**Bandwidth**	**Treatment estimate**	**S.E.**
Share new district	1	0.101	0.020**	0.009	0.172	0.017**	0.008
Share new district	2	0.134	0.022*	0.012	0.167	0.023**	0.011
Share former district	1	0.102	-0.011*	0.007	0.152	-0.010*	0.006
Share former district	2	0.130	-0.022**	0.009	0.152	-0.019**	0.009

Remarks: this table shows the RDD regression bandwidths, estimates and standard errors of the effect of an elected female in a close mixed-gender election on the two additional outcome variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most the percentage indicated in the bandwidth. The bandwidths are determined using the MSE bandwidth selector. These are performed for both a uniform and triangular kernel, as well as for the regression forms with an order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The number of mixed-gender election races differs for each bandwidth. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

Table A3: RDD of elected female in mixed-gender elections on additional $Y_{c,t}$ with different bandwidths on both sides, MSE-selected and uniform kernel

		Bandwidth			
Outcome variable	**Order**	**Left**	**Right**	**Treatment estimate**	**S.E.**
Share new district	1	0.089	0.139	0.016*	0.009
Share new district	2	0.140	0.123	0.020	0.013
Share former district	1	0.103	0.135	-0.006	0.006
Share former district	2	0.156	0.137	-0.016*	0.009

Remarks: this table shows the RDD regression bandwidths, estimates and standard errors of the effect of an elected female in a close mixed-gender election on the two additional variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most the percentage indicated in the right bandwidth or when a candidate lost with a relative vote margin of at most the percentage in the left bandwidth. The bandwidths are determined using the MSE bandwidth selector. These are performed for a uniform kernel and for the regression forms with an order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The number of observations differs for each bandwidth. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.

A.4 CER-bandwidth selector additional results

Table A4: RDD of elected female in mixed-gender elections on additional $Y_{c,t}$ with CER-selected bandwidth

		Uniform			Triangular		
Outcome variable	**Order**	**Bandwidth**	**Treatment estimate**	**S.E.**	**Bandwidth**	**Treatment estimate**	**S.E.**
Share new district	1	0.073	0.021*	0.011	0.125	0.020**	0.009
Share new district	2	0.092	0.021	0.014	0.116	0.021	0.013
Share former district	1	0.074	-0.013*	0.008	0.110	-0.015**	0.007
Share former district	2	0.090	-0.023**	0.010	0.105	-0.024**	0.010

Remarks: this table shows the bandwidths, estimates and standard errors of a RDD regression on the effect of an elected female in a close mixed-gender election on the two additional variables. A mixed-gender election is close when the elected candidate won with a relative vote margin of at most the percentage indicated in the bandwidth. The bandwidths are determined using the CER bandwidth selector. These are performed for both a uniform and triangular kernel, as well as for the regression forms with an order of 1 or 2. An order of 1 is the same as a local linear regression. The treatment estimates and standard errors are rounded. The number of mixed-gender election races differs for each bandwidth. The stars indicate the significance level with *: p<0.10, **: p<0.05, ***: p<0.01.