

STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2019–2020

Female representation and the composition of municipal budgets: Evidence from Czech municipalities

Martina Kaplanová (41446)

Abstract. The share of women in political positions has been steadily rising for decades, yet it remains unclear whether this increase has a significant effect on policy outcomes. This thesis investigates whether higher female representation in local councils has implications for policy decisions, evidenced by the composition of public expenditures in Czech municipalities. Using survey data on satisfaction with selected areas of public life, I classify local public expenditures into male, female and neutral. In a research context of open-list, competitive elections with proportional representation, I utilize publicly available data from the 2010 and 2014 municipal council elections, covering 460,841 candidates from 6,254 municipalities. Exploiting the unpredictability of the vote margin needed to win a seat, I use Regression Discontinuity Design to compare the share of female expenditures in municipalities with marginally elected women to those with marginally elected men. This thesis fails to find an effect of an additional female on the budget composition and the findings are robust across a variety of specifications.

Keywords: female representation, gender, local governance, political economy, regression discontinuity

JEL: D72, H72, J16

Supervisor:	Pamela Campa
Date submitted:	May 15, 2020
Date discussed:	May 26, 2020
Discussant:	Mikaela Knutson
Examiner	Andreea Enache

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Pamela Campa, for her continuous support, patience and motivation. Her guidance was immensely helpful throughout this project. Additionally, I would like to thank my family and friends for their support and encouragement during my studies. The advice and support was greatly appreciated. All remaining errors are mine.

Contents

List of tables	3
List of figures	4
1 Introduction	5
2 Literature review	10
3 Background	13
4 Data	18
5 Methodology	21
6 Results	28
7 Discussion	37
8 Conclusion	39
References	41
A Additional output	44
B Mandates allocation and margin calculation	55
C Power calculations	58
D Data sources	61

List of Tables

1	Survey results: Satisfaction with selected areas of public life	17
2	Summary statistics: Spending data by gender and area	19
3	Summary statistics: Municipal elections	20
4	Sharp RD: Female victory and the average share of female spending	30
5	Fuzzy RD: Female victory and the share of female councillors (first-stage results)	32
6	Fuzzy RD: Share of female councillors and the average share of female spending	33
7	Extensions: Female victory and the average share of education spending	34
8	Extensions: Dependent variable in differences	35
A.1	Summary statistics: RD sample by number of mixed contests	44
A.2	OLS: Female spending pooled across election period	44
A.3	Validity: Vote margin by the gender of the winner	45
A.4	Validity: Covariate balance	45
A.5	Validity: Alternative cutoffs	46
A.6	Validity: Donut hole approach	46
A.7	Validity: Alternative bandwidths	47
A.8	Sharp RD: Female victory and the yearly share of female spending	47
A.9	Robustness: Estimation with annual change in share of female spending	48
A.10	Robustness: Estimation excluding election years	48
A.11	Robustness: Estimation in municipalities with only one mixed contest	49
A.12	Robustness: Estimation restricted to small municipalities	49
B.1	Slate order before and after election	55
B.2	Post-election vote distribution	56
C.1	Power calculations	58
C.2	Minimum detectable effect size (MDES)	59

List of Figures

1	Sharp RD: Global RD plot of female victory and the share of female spending	29
2	Sharp RD: Local RD plot of female victory and the share of female spending	29
3	Fuzzy RD: Global RD plot of female victory and the share of female councillors . . .	31
4	Fuzzy RD: First stage estimation plot	32
A.1	Validity: Candidate-level predetermined covariates	50
A.2	Validity: Municipality-level predetermined covariates	51
A.3	Validity: 2011 Census predetermined covariates	52
A.4	Validity: Density plots	53
A.5	Validity: Alternative bandwidth test	53
A.6	Extensions: Local RD plot of the average share of education spending	54
A.7	Extensions: Local RD plot of the difference in average share of female spending . . .	54
C.1	Power calculations	60

1 Introduction

The participation of women in many areas of public life has been steadily increasing for decades. Currently, in most developed countries, women and men are equal in terms of their rights and freedoms. The political participation of women, enabled by universal enfranchisement, has had a strong effect on reshaping public policy, particularly by inducing the expansion of the welfare state: the increasing share of female voters is associated with growing expenditures with redistributive purposes and social welfare spending [Abrams and Settle, 1999; Aidt and Dallal, 2008]. However, the representation of females in political offices still lags behind that of males. In every EU country, the parliament has more male than female members: the EU average of female members of parliament is 32.2%, ranging from 12% in Lichtenstein to 47.6% in Sweden.¹ In the Czech Republic, women remain comparably underrepresented in political functions: the Czech Republic has never had a female president or prime minister, the current share of women in the parliament and the government is 21% and 27%, respectively, and in the local governments, women occupy on average only 20% of seats on the regional councils and 27% of municipal council seats [Šprincová and Mottlová, 2015].

Numerous survey- and experiment-based studies have shown that women tend to have different policy preferences than men: women tend to be less selfish and more social welfare oriented and thus, more concerned with policy areas such as health care, education, and social security [Bagues and Campa, 2017; Croson and Gneezy, 2009; Dollar et al., 2001; Ranehill and Weber, 2017; Svaleryd, 2009]. Additionally, as women continue to be the primary care givers in many countries, issues such as child care are often more important to them as accessible child care enables them to enter the labour force more easily [Baskaran and Hessami, 2018]. This difference in policy preferences, combined with the unequal share of women in office, suggests that the interests of women might be currently underrepresented in the policy-making bodies. An increase in political representation of women could therefore lead to more female-oriented decision-making and different policy outcomes.

It is not clear, however, that an increased share of females holding a political office is either sufficient or necessary to represent the preferences of female voters. Two main theories explain the potential effect of personal characteristics such as gender, education or ethnic background on the decision-making of an elected representatives. On the one hand, the median voter theorem suggests that the individual preferences should not influence decision-making, as the elected officials' decisions converge on the preferences of the median voter [Downs, 1957]. According to this theory, the main objective of a person in office is re-election, and therefore, they are incentivized to represent the preferences of the largest possible share of their electorate. On the other hand, the citizen-candidate model suggests that personal characteristics are relevant: they offer the candidates strategic advantages as they allow them to credibly commit to a particular policy standpoint [Besley and Coate, 1997; Osborne and Slivinski, 1996]. With respect to gender representation, this relies on the assumption that members of the same gender share a set of preferences and perspectives acquired through a common life experience [Mansbridge, 1999].

¹Statistics retrieved from European Institute for Gender Equality [2020].

This paper analyzes the effects of female representation in Czech municipal councils on the composition of municipal budgets. It relates to existing studies which estimate the impact of female representation on policy outcomes.² Based on a survey of male and female policy preferences, the municipal expenditures are classified into ‘female’, ‘male’ and ‘neutral’. Then, I use the local polynomial Regression Discontinuity (RD) method to investigate whether the share of female expenditures differs in the municipalities with a marginally elected female from those with a marginally elected male. My analysis shows that municipalities with an additional female do not demonstrate any statistically or economically significantly different spending patterns than those with additional males.

To shed light on the differences in the policy preferences between males and females in the Czech Republic, I use survey data on satisfaction with selected aspects of public life of Czech residents. Comparing the means of the two genders,³ I find that women are significantly less satisfied with infrastructure, science and research, and culture. On the other hand, men appear to be more concerned with the issues of healthcare, crime, safety and employment policy. I use the differences in satisfaction as a proxy for a difference in preference for policy change in a particular area. Hence, the expenditures in the areas where women are less satisfied are classified as female expenditures, while male expenditures are those areas with which are men less satisfied. All remaining spending is labeled as neutral.

Using this classification, I first run a simple OLS regression to find any associations between the share of female councillors and female expenditures. As the councillors are in office for four years, I use the average share of female spending in the four year office term as my dependent variable. In contrast to previous literature,⁴ I find no statistically significant relationship, suggesting that the share of females on the council is not correlated with the share of female spending. Comparable findings are revealed when including a control for lagged female expenditures from the previous legislature.

In the main estimation, I use the local polynomial method with optimal bandwidth for sharp RD.⁵ My RD sample includes the marginal *gender-mixed contests*: the competitions for the last seat allocated to an election party where the marginal winner and loser are of *opposite genders*. I calculate the vote margin between the female and the male candidate, creating a running variable with the cut-off at zero: a positive margin translates into a female winner and thus a marginal female councillor, while a negative margin indicates a male winner. In the sharp RD, the vote margin becomes the independent variable, while the dependent variable is the average share of female spending. The estimation yields a slightly negative effect which is statistically insignificant. My estimation thus fails to find an effect of an additional woman on the budget composition compared to an additional man. The results are robust to various

²See for instance Bagues and Campa [2017]; Baskaran and Hessami [2018]; Chattopadhyay and Duflo [2004]; Funk and Gathmann [2015]; Kuliomina [2019]; Rigon and Tanzi [2012].

³Note that in this thesis, I restrict ‘gender’ to a binary form, discriminating only between men and women. I refrain from using the (by definition) binary term ‘sex’ instead, firstly as the term ‘gender’ is entrenched in the existing literature investigating female representation in politics, and secondly, as the term ‘sex’ suggests biological nature in the differences between men and women, rather than their social roles.

⁴See for example Bagues and Campa [2017] or Kuliomina [2019].

⁵Following the framework set up in Cattaneo et al. [2019a].

alternative bandwidths, a change to a quadratic polynomial, and a number of robustness checks. Across the different estimations, the point estimate is always close to zero and its direction is unstable, fluctuating between negative and positive.

I present a number of extensions to the baseline model. Firstly, I re-run the estimation using a fuzzy RD method: I study the effect of the share of female councillors, instrumented by the vote margin, on the average share of female expenditures. Although the first stage is statistically significant – one additional female councillor increases the total share of females on the council on average by 7.6 p.p. – the main results confirm the sharp RD findings: I fail to reject the null hypothesis that an increased share of females has no statistically significant effect on the average share of female spending. Secondly, I repeat the estimation using only education expenditures as the dependent variable. Education was not indicated as either male or female expenditure in my survey results. However, the existing literature often highlights education as one of the main female policy areas.⁶ Therefore, I use the average share of education expenditures as a dependent variable in a sharp RD specification. The results remain largely unchanged: I cannot reject the null hypothesis that an additional female has no influence on the education budget. Finally, I run the estimation using the difference between the share of female spending of the 2014 and 2010 legislatures. Again, the results stay statistically insignificant, failing to find evidence that electing an additional female significantly increases the share of female spending compared to the previous council’s spending.

The existing empirical evidence of female representation on policy outcomes has so far been inconclusive. Although some studies suggest that female representation has a statistically significant effect on public good provision and policy making in areas of education, healthcare, child care, or infrastructure [Baskaran et al., 2018; Baskaran and Hessami, 2018; Chattopadhyay and Duflo, 2004; Svaleryd, 2009], many others fail to find such significant impact [Bagues and Campa, 2017; Ferreira and Gyourko, 2014; Kuliomina, 2019; Rigon and Tanzi, 2012]. These studies cover different levels of governance (from village councils and mayors and to state legislatures) and distinct outcome variables (economic performance, fiscal responsibility, budget composition, provision of services). Therefore, it is valuable to investigate this topic further in a new research context, as it might help to identify the channels through which female representation affects policy outcomes and the conditions required for this effect to take place.

The existing literature investigates this relationship mainly in two geographic regions: developing countries and Western democracies. The Czech Republic, formerly a part of the Eastern bloc, is arguably a member of neither and thus offers a useful extension to the existing literature in these two groups of countries. Due to its history as a socialist country, some issues which generally concern women such as childcare provision or maternity benefits are less pronounced there than in many Western countries.⁷ Although women are underrepresented in the governing bodies in the Czech Republic, the generous welfare policies might ensure that the interests of

⁶See for instance Chattopadhyay and Duflo [2004], Funk and Gathmann [2015], Bagues and Campa [2017], or Svaleryd [2009].

⁷As an example, mothers are entitled to maternity leave with 100% guaranteed pay. Subsequent parental leave can last up to 3 years, during which the mother gets up to 70% of her previous salary and a guarantee of her previous workplace. After that, childcare is provided for children from 2-3 years of age and heavily subsidised (at state pre-schools).

women are better taken care of and thus, the effect of additional female politicians might not be as pronounced.⁸ Since similarly generous social welfare policies are still in place in many other Eastern European countries,⁹ my results might be easier to generalize across this pool of countries.

My thesis further expands the existing literature by focusing on a setting with competitive elections. The majority of the previous studies look at the effect of female representation in countries with mandated gender quota (either on the ballot or as seat reservation policies). The quota can, however, have a direct effect on both the female representation and policy outcomes. Firstly, the quota's impacts on the share of female councillors might be limited due to both party bias (women are placed on lower positions on the ballot) or voter bias (voters cast less preferences for women or for lists that due to the quota have to increase their share of women to a larger extent). Secondly, the quota could lead to electing female candidates with different characteristics and motivations. For instance, they might have weaker election incentives (as re-election might be easier for them compared to a setting without a quota) and thus might carry out more policy decisions that they personally prefer. On the other hand, the elected women could be less influential in the council as they might be less experienced. Thirdly, the quota by definition change the share of female candidates, which in itself could have an impact on policy, particularly if policy choices are made at party rather than council level. Finally, the quota could also affect female turnout, inducing policy change through the median voter theorem. Due to these peculiarities of the quota-based systems, the results of the previous studies might not generalize well to a context with competitive elections. Hence, it is desirable to explore the relationship between the gender of the legislators and policy outcomes in a competitive setting without the quota induced distortions to the candidate pool and election or policy outcomes.

Lastly, this paper contributes to the literature in that it studies the impacts of female representation in the local legislative bodies. With the exception of [Baskaran and Hessami \[2018\]](#); [Kuliomina \[2019\]](#), previous studies in a competitive election setting look at the representation in executive positions, mainly the mayor's office.¹⁰ In many countries, mayors are rather powerful and thus, a mayor can possibly exert much higher levels of influence on the municipality than the local legislators. Therefore, it is not clear whether female representation in legislative bodies can have an effect on policy that is similar to that of females in the executive offices. On the other hand, in the Czech Republic, the powers of the mayor are much more limited and a majority of the municipal decisions are taken by the council using consensus-based voting. It might thus be the case that the influence of women in the legislature is comparable to that of women in the mayor position. Nevertheless, as [Baskaran and Hessami \[2018\]](#) point out, more expansive policy changes require legislative action beyond the authority of the leader of the executive. Therefore, it is valuable to gain more insight into the policy effects of female representation in

⁸On the other hand, it could be the case that some interests of women which are usually taken care of in the Western countries are not taken care of in the Czech Republic. These might then be affected by the female policy makers.

⁹Some of the countries who also guarantee 3-year parental leave include Croatia, Lithuania, Hungary, Poland or Slovakia.

¹⁰See for instance [Ferreira and Gyourko \[2014\]](#); [Gagliarducci and Paserman \[2012\]](#); [Holman \[2014\]](#) who all look at the effects of female mayors.

the legislative branch.

The context of Czech municipalities is favourable to my analysis for a number of reasons. Firstly, municipal election and spending data is very rich. The Czech Republic has more than 6,200 municipal councils and more than 220,000 candidates on municipal council seats in each election, which takes place every four years. This provides a large number of races where the marginally elected and not-elected candidates from a party slate are of different genders, thereby creating a large dataset. The spending data are also very rich: the Czech Statistical Office and the Czech Ministry of Finance periodically provides granular, publicly accessible data with accounting and budgetary information about the municipal administrations. The Czech context thus offers an extensive dataset, but it has almost never been used for a research study of this sort.¹¹

Secondly, the structure of the elections is open-list, meaning that the voters do not simply cast a vote for one party as a whole, but can also cast individual votes for particular candidates. The voters can cast as many votes as there are seats on the council, and can vote for candidates from different party slates. Using these preferential votes, the voters have the ability to change the rank of the candidates on the ballot, and thus to increase the likelihood of particular candidates winning a seat. This feature makes it impossible for both the voters and the candidates to predict the cut-off number of votes needed to win a seat. That in turn enables me to use a quasi-experimental RD method and estimate the effect of marginally elected candidates on the spending patterns.

The final benefit of the Czech context is that the absence of a quota allows me to study all municipalities including the ones with a very small population. In most studies that use quotas as a basis for the RD,¹² the very small municipalities are exempted from the implementation of the quota. Additionally, as quotas are usually mandated at a particular population threshold, the nature of RD restricts the sample only to the municipalities with a population close to the cut-off. This makes it difficult to generalize the results to municipalities with much larger or much smaller population than the threshold. The size of the councils mostly depends on the size of the municipality, meaning that an additional female councillor in a small municipality might have much more influence than one in a larger municipality (since their respective impact on the overall share of female councillors differ). Hence, using the Czech context is useful as it allows to estimate the relationship between female representation and policy outcomes in municipalities which are usually excluded from quota-based studies.

The rest of this thesis is structured as follows. Section 2 presents the relevant literature. Section 3 explains the context of Czech municipalities. Section 4 describes the collection and consolidation of the data. Section 5 illustrates the methodology. Section 6 reports the results. Sections 7 discusses the implications of the results and the limitations of the study. Finally, Section 8 concludes the paper.

¹¹To my knowledge, only Kuliomina [2019] studies the effect of an increase of women in office in the Czech municipalities. Her research, however, focuses on *general* financial indicators such as overall spending, deficit and debt, and not on the *specific* composition of the spending.

¹²See for instance Bagues and Campa [2017] or Rigon and Tanzi [2012].

2 Literature review

2.1 Gender and policy preferences

A crucial concern in analyzing the effects of gender on policy decisions is whether males and females have different policy preferences. This topic has been covered in many studies, finding that women tend to differ from men both in terms of social and policy preferences. [Croson and Gneezy \[2009\]](#) present a literature review showing fundamental differences between men and women in their risk aversion and competitiveness (women are more risk averse, while men are more competitive). Similarly, [Ranehill and Weber \[2017\]](#) use a laboratory experiment to show that women prefer more egalitarian redistributive policy. For specific policy areas, studies indicate that females tend to be more concerned with childcare, elderly care, health, education, and gender related policy issues [[Bagues and Campa, 2017](#); [Funk and Gathmann, 2015](#); [Svaleryd, 2009](#)]. [Bagues and Campa \[2017\]](#) show this to be the case using survey data on Spanish residents, while [Svaleryd \[2009\]](#) presents similar results from a survey of Swedish local policy makers. Exploiting the direct democracy system in Switzerland, [Funk and Gathmann \[2015\]](#) come to a similar conclusion using information from local referenda on a variety of policy proposals. Males, on the other hand, seem to be more concerned with problems related to housing, work conditions, corruption, and immigration [[Bagues and Campa, 2017](#)]. This evidence suggest that if personal preferences matter in decision-making of elected representatives, the increase in female representation could have an effect on policy matters.

2.2 Effect of female representation on decision-making

Previous studies have mainly examined whether the gender of the representatives has an effect in two areas: behaviour in office and policy outcomes. The studies which focus on behaviour in office show evidence suggesting that female representatives tend to be more accountable and less prone to criminal and corrupt conduct. At the country level, survey-based studies by [Dollar et al. \[2001\]](#) and [Swamy et al. \[2001\]](#) show that a higher share of women in the government reduces corruption, as women tend to exhibit less opportunistic behaviour. Using micro-level data, [Swamy et al. \[2001\]](#) also show that women are less inclined to accept and offer bribes. [Baskaran et al. \[2018\]](#) further add that female legislators are less likely to accumulate assets and be subject to pending criminal charges during their time in office. An RD study of mayors in Brazil by [Brollo and Troiano \[2016\]](#) found that women tend to employ fewer temporary workers (which is a common way to engage in patronage) and collect less funds from campaign contributions often used to exert political influence. Similarly, [Andersen et al. \[2008\]](#) use a simple experiment to compare matriarchal and patriarchal societies and provide suggestive evidence that the matriarchal societies provide more public goods and have fewer agents with a tendency to free-ride.

On the other hand, more adverse behavioural outcomes of female representation are presented by [Gagliarducci and Paserman \[2012\]](#), who indicate that Italian female mayors are more prone to resign before the end of their term, despite no significant difference in their performance in office. Another possible hindrance to effective female representation is the argument that women representatives can only have substantial influence on policy if they can gain critical mass in

the decision-making body. An experimental study by [Born et al. \[2018\]](#) suggest that women are less inclined to take a leadership position in male dominated environments. Using empirical data and an RD method, [Baskaran and Hessami \[2018\]](#) show this to be the case in Bavarian municipalities: they conclude that the increase in female representation is most influential if it markedly reduces the minority position of the women.¹³

The literature that looks at the impact of increased representation on policy outcomes contains studies from various levels of decision-making, geographically diverse settings, and different types of electoral structure. These research contexts offer inconclusive evidence in terms of the effect of female representation.

A portion of the literature which mostly finds statistically significant effects come from India. On a state level, [Baskaran et al. \[2018\]](#) show that economic performance increases in women-led constituencies. More ambiguous results are presented by [Clots-Figueras \[2011\]](#), who studies the influence of female legislators on policy outcomes in 16 Indian state legislatures, showing that only females from lower social standing have a statistically significant effect on passing health, education and gender equality policies. In a later study, she also finds a significant effect of increasing female political presence on educational attainment in urban areas, but not in rural areas [[Clots-Figueras, 2012](#)]. Finally, focusing on municipal councils, [Chattopadhyay and Duflo \[2004\]](#) exploit the seat reservation under mandated representation policy in Indian village councils and find that the seat reservation does affect policy choices particularly in the areas of infrastructure. Due to the particular context of these studies, it is not clear whether the results could be generalized to a developed country in the EU.

Large scale studies in Western countries provide evidence that female representation can have an effect on budgetary and policy decisions. For instance [Braendle and Colombier \[2016\]](#) search for the drivers of increased health expenditures in Swiss cantons since the 1970s using dynamic panel estimation and find the share of female legislators to be a statistically significant determinant. Similarly, using a cross country comparison, [Clayton and Zetterberg \[2018\]](#) show that only a large increase in the proportion of women legislators induced by gender quotas increases healthcare spending and shrinks military spending. It is important to note that both of the studies utilize a rather sizeable increase in female representation. [Clayton and Zetterberg \[2018\]](#) use the introduction of gender quotas, which produces a large increase in female elected officials, while [Braendle and Colombier \[2016\]](#) use the Swiss setting where women were first allowed to run for office in 1971, effectively having the share of females legislators equal to 0 at the beginning of the data set. In many countries, such conditions are rather impossible to replicate, and the effects may not generalize to settings with smaller variations. Finally, using the setting of mandated quota in the French Parliament, [Lippmann \[2019\]](#) finds that female legislators are more likely to support women-related legislation and author women-related amendments to legislation.

The micro-evidence from local studies shows a much more mixed picture of the effects of increasing female representation, though the portion of studies which fail to find a substantial influence is somewhat larger than that of studies which find an effect. However, there seem to

¹³[Baskaran and Hessami \[2018\]](#) show that the effect of an additional female councillor on the provision of child care is the largest in councils which already have two other female members.

be no clear patterns in the electoral and geographic settings, levels of policy-making, or even methodological approach of the studies that could explain what are the conditions in which female representation matters.

The studies which found statistically significant influence of female representation include for instance [Holman \[2014\]](#), who finds a strong positive effects of female mayors on funding and participation in social welfare programs in the context of US cities. However, she also presents evidence that the share of females on the city council does not affect social welfare spending by itself but is relevant in interaction with the gender of the mayor. [Svaleryd \[2009\]](#), who studies female representation in Swedish municipal councils reports that a higher proportion of female councillors is associated with increased spending on childcare and elderly care. Finally, a paper by [Baskaran and Hessami \[2018\]](#) studies Bavarian municipalities and the effect of female councillors on the provision of childcare, finding that in municipalities with marginally elected females, the public child care provision increases by more than 40%. This study is particularly noteworthy as its context closely resembles the setting in the Czech Republic - the electoral and municipal set-up is comparable with that of my study, which allows me to adopt a similar methodological approach.

On the other hand, numerous studies show no statistically significant influence of increased female representation. In contrast to [Holman \[2014\]](#), [Ferreira and Gyourko \[2014\]](#) find no statistically significant effects of female mayors on spending patterns, criminality, or unemployment in US cities.¹⁴ In the context of the EU, [Carozzi and Gago \[2017\]](#) study the impact of female mayors in Spain on passing gender-specific policies. Finding no statistically significant effect, they argue that a much larger difference occurs along party lines. Focusing instead on Spanish municipal councils, [Bagues and Campa \[2017\]](#) suggest that the implementation of a gender quota affects neither the size nor the composition of the municipal budget. Finally, [Rigon and Tanzi \[2012\]](#) use gender quotas in Italian municipalities as a source of exogenous variation, showing no statistically significant effect on spending patterns in the treated municipalities.

Due to this inconclusive evidence, it is unclear whether the female representatives in the Czech Republic are likely to have an influence on the composition of the municipal spending. The literature clearly lacks a focus on countries in Eastern Europe, which are in many ways different from both Western democracies and developing countries. In the only previous paper focusing on the Czech Republic, [Kuliomina \[2019\]](#) estimates the effect of gender, education and entrepreneurial background of council members on municipality income, expenditure and deficit, and finds that an additional female candidate has little effect on the budget indicators. Although the study somewhat differs from this thesis in terms of method, it suggests that this paper will likely fail to find an effect as well.

¹⁴However, they find that females tend to have a higher incumbency effects, suggesting that they outperform men in terms of their unobserved political skills.

3 Background

3.1 Local governments in the Czech Republic

The Czech system of public governance is divided into three levels: municipal, regional and national. The country's fourteen regions contain 6,254 municipalities ranging in size from small villages with less than a hundred residents to the capital city Prague with a population of over one million. There are five types of municipalities based on their population and historical status: capital city, statutory city,¹⁵ city, market town and municipality.¹⁶

3.1.1 Municipal functions, powers and responsibilities

All the municipalities, regardless of their type, have some autonomous powers, but also powers delegated from the central state government. The extent of the delegated powers depends on the size and the population of the municipality. The autonomous powers mainly cover the areas of education, waste management and environmental protection, local infrastructure, health services, and local police. The primary responsibilities of the municipality are to provide waste removal services and ensure that pupils are able to comply with the compulsory schooling rules.¹⁷ Additionally, the municipalities can implement other projects and programs. Some large municipalities might have additional responsibilities delegated from the state, such as the provision of elderly care, and social welfare or housing services [Balík, 2009].

The municipality administration consists of multiple authorities. The main governing body is the municipal council (or simply, council), which is directly elected in the municipal elections. The council size ranges between 5 and 70 and depends on the population of the municipality.¹⁸ In 2010, the most common sizes of councils were 7, 9, and 15 members, covering 43%, 23% and 15% of the municipalities, respectively. Most councils have an odd number of councillors – a practical feature as all decisions of the council have to pass by a simple majority. However, most decisions taken by the municipal councils tend to be consensus-based, and the occurrence of unanimous voting is not exceptional [Balík, 2009]. The main tasks of the council are to prepare and pass the municipality budget, approve the development program of the municipality, establish or dissolve public-benefit organizations (such as schools, museums or healthcare centers), pass local ordinances, and approve local referenda.¹⁹

Once a municipality council is elected, the councillors elect one council member to the position of mayor. The main powers of the mayor include calling a municipality meeting and setting its agenda. The mayor also reviews the municipality budget and regularly informs the residents

¹⁵Seven of the statutory cities and the capital are further divided into city-parts, which then can have their own autonomous authorities and budgets, essentially creating a sub-municipal level of local governance. These city-parts do not publish their public finance information in the same manner as the municipalities, and thus are not considered in this study. The statutory cities have both a municipal council and city-part councils. The municipal councils of these seven cities are included in my sample.

¹⁶Municipality is both the overall term and the name of the basic type of municipality with no special status.

¹⁷Children are subject to nine years of compulsory education, provided by elementary schools.

¹⁸Note that the council size in Prague fluctuates between the election rounds. Although the maximum *allowed* size of the Prague municipal council is 70 members, the *actual* size of the council was 63 and 65 in the 2010 and 2014 election rounds, respectively.

¹⁹From Act No. 128/2000 Coll., on Municipalities (on Municipal Order), originally, Zákon č. 128/2000 Sb., Zákon o obcích (o obecním zřízení), see Parlament České Republiky [2000].

about the work done by the municipality authorities. Additionally, the council also appoints councillors to the municipal committee which serves as the executive authority of the municipality and is led by the mayor. The committee's main responsibilities include implementing the budget approved by the council, controlling the functioning of the municipality office and commissions, and issuing fines to persons or legal entities disobeying local ordinances. In small municipalities without a committee, the responsibilities are taken over by the mayor alone.²⁰

3.1.2 Municipal budgets

Although the council has the power to propose and approve the budget, the budget contains many items which are mandatory (such as pensions or parental contributions). Moreover, there are some differences in the size of the budget depending on the size of the municipality and the extent of the delegated powers.

The income part of the municipal budget consists of tax revenues, non-tax revenues, capital revenues, and subsidies and transfers. The majority of the income is determined by the state legislation (such as the tax rates). However, the municipalities can collect payments from the population based on their own local ordinances (for instance, by implementing a fee on dog ownership, a fee on using public spaces for events etc.). Municipalities can also generate revenue from their own economic activity (e.g. revenue from advertisements at the municipality property).

The expenditures of the municipalities are divided into current spending and capital spending. Current spending, which takes up about three quarters of the total spending, covers expenses that tend to be repetitive and is financed through the tax and non-tax revenues, subsidies, and transfers. Ideally, the current section of the budget should be balanced. Current spending is further divided into six categories: agriculture and forestry, industry and economy, services to inhabitants, social affairs and employment policy, security and public administration. The largest category on the budget is usually public administration (on average 39% of spending), followed by the services to inhabitants (on average 37% of spending). The services to inhabitants cover a variety of areas such as education, waste management, health services, environmental protection, and cultural and religious activities. The capital spending, which accounts for the remaining quarter, is financed through the capital revenue, and can run a deficit as it covers investments which exceed the period of the one-year budgets.

3.1.3 Municipal elections

The municipal councillors are elected every four years and the elections occur in all municipalities at the same time. Eligible voters are all permanent residents in the municipality above the age of 18 with a Czech or EU citizenship, but the right to run for a mandate is restricted to Czech citizens. There is no term limit for the council members: they can be re-elected indefinitely. All candidates are grouped on the ballot within slates, which are sub-lists of the candidates based on their election parties. An election party can consist of individual political parties or their

²⁰From *Act No. 128/2000 Coll., on Municipalities (on Municipal Order)*, originally, *Zákon č. 128/2000 Sb., Zákon o obcích (o obecním zřízení)*, see [Parlament České Republiky \[2000\]](#).

association, individual independent candidates or their association, or an association of political parties and independent candidates. Each candidate can be listed on one slate only, and each slate can have a maximum number of candidates that is equal to the number of the available council seats.²¹

The rank of the slates on the ballot is picked randomly within each municipality,²² while the election parties themselves determine the order of the candidates within their slates. Each candidate is listed on the slate with the information about their age and occupation and academic titles, based on which the voter can find out whether the candidate has completed any tertiary education. Although gender is not explicitly listed, the Czech language differentiates between female and male surnames and job titles, which enables the voters to learn the gender of the candidate with high precision.

The initial rank of the candidates on the slates is not binding for the success of the candidates, as voters have the ability to split their votes: instead of casting one vote for a slate, they can cast multiple votes for individual people on one or more slates. Each voter can cast as many votes as there are seats in the municipality council. A voter can choose one of the following strategies: (i) cast one vote to one slate, whereby each candidate on the slate is given one vote, (ii) cast as many votes to individual candidates across slates as there are seats on the council, or (iii) combine the previous two strategies, cast one vote to a slate and then some additional votes for particular candidates from different slates. Using the third method, the votes are first distributed among the individual candidates; the remainder is then given to the candidates from the chosen slate in their rank on the slate.²³ Votes, however, cannot be accumulated, i.e. a voter cannot cast multiple votes to one candidate [Balík, 2009].

Once the votes are counted, the candidates' final rank on a slate is determined: for each slate, an average number of votes per candidate is calculated and any candidate who received more than 10% of the average moves up on the slate. The council mandates are allocated using a proportional distribution based on d'Hondt's method, which determines the number of mandates for each election party receiving more than 5% of the total votes (see Appendix B.1 for further details on mandate allocation, including an example). The appropriate number of mandates is then distributed within each election party to the candidates in their final rank on the slate. The combination of the d'Hondt's method with the preference votes makes it impossible to predict how many and which candidates from a slate will be successful in winning a mandate. This allows for the use of a RD design, as it is impossible to anticipate which candidates on a slate will be the last successful and first unsuccessful.

Bernard [2012] studies the importance of the characteristics listed on the ballots, both in terms of the initial rank of candidates on the slates and their election outcomes. In the Czech context, the

²¹From Act no. 491/2001 Coll., on Elections to Municipality Councils and on Amendments to Selected Acts, originally, *Zákon č. 491/2001 Sb. Zákon o volbách do zastupitelstev obcí a o změně některých zákonů*, see Parliament České Republiky [2001].

²²The same random order is used for all ballots in the municipality, that is, each voter in a municipality obtains an envelope with slates ordered the same way.

²³Note that the individual votes of a voter do not carry any rank by themselves. For instance, it is not possible for a voter to vote individually for each candidate on a slate but in a different rank than that on the slate itself. The votes are all equal in their weight.

incumbency is the strongest determinant of both high position on a slate and subsequent electoral success. Gender of the candidates seem to also have an effect: women tend to be placed lower on the slate. This placement on positions which are not as ‘electable’ decreases their probability of becoming a councillor [Jurajda and München, 2015]. Finally, most municipalities in the Czech Republic are rather small and the candidates are often well known in the municipality. Therefore, the characteristics on the ballot might not matter as much as the characteristics observed in real life [Balík, 2009]. In the few large cities, the information about education, age, occupation and political affiliation likely becomes relatively more important. This is further supported by the fact that the candidates in the cities are much more likely to be affiliated with a political party than those from the small municipalities.

3.2 Policy preferences of males and females

To examine the differences in policy preferences between men and women, I use data from the Czech Social Science Data Archive (*Český Sociálněvědní Datový Archiv*) collected by the Public Opinion Research Centre of the Institute of Sociology of the Czech Academy of Sciences (*Centrum pro Výzkum Veřejného Mnění Sociologického Ústavu Akademie Věd České Republiky*).²⁴ The Public Opinion Research Centre collects and publishes monthly data as part of a project called ‘Our Society’ (*Náše Společnost*), which tracks public opinion across a range of areas. One of the recurring areas of inquiry in this project is the ‘Satisfaction with chosen aspects of public life’ (*Spokojenost se stavem vybraných oblastí veřejného života*), where the respondents score their satisfaction with public policy areas such as education, health care, social security, agriculture, immigration, unemployment, or crime.

The ‘Satisfaction with chosen aspects of public life’ survey was carried out on eight instances between January 2012 and February 2019.²⁵ The surveys included socio-demographic questions in which the respondents provided information on their gender, age, highest achieved education level, political preferences (on a left-right scale), municipality size, and living conditions. On each occasion, the size of the respondent pool was around 1,000 people which in total sums up to a sample size of more than 8,000 respondents. The sample is by construction representative of the population of the Czech Republic in terms of gender, age and geographic location. In each round of the project, a new group of respondents is approached and the interviews are conducted in person. The data from the surveys reveal that some respondents do not answer all questions, and thus the final sample size is somewhat smaller, and differs for each of the relevant aspects of public life.

To estimate the differences in satisfaction between men and women, I compared the means of responses between males and females.²⁶ In the survey, the satisfaction was reported on a 5-

²⁴Data retrieved from <http://nesstar.soc.cas.cz/webview/>, see Appendix D for details.

²⁵These instances were 01/2012, 06/2012, 01/2013, 06/2013, 09/2015, 02/2016, 01/2017 and 02/2019. Prior to 2012, the survey had a different format and structure, incomparable with the latter versions.

²⁶Notably, I use the raw comparison of means instead of OLS framework with additional controls. This decision was made deliberately to preserve the tendencies of groups based on their characteristics. For instance, if women on average earn less income, and low-income people are less satisfied with the situation in social services, controlling for income would skew the results by comparing low-earning women with only low-earning men. Without the control, on the other hand, the results account for the difference in characteristics between men and women in the population. It would be desirable to include the controls only if there were issues with the sample

level scale (very satisfied, somewhat satisfied, neutral, somewhat dissatisfied, very dissatisfied). To account for the fact that different respondents might have different willingness to report extreme responses, I convert the 5-level scale into a 3-level one (satisfied, neutral, dissatisfied). However, I also use the original 5-level scale for a sensitivity check, revealing that the outcomes remain largely comparable. Table 1 shows the results of the t-test. Using the 3-level scale, the results suggest that women are statistically significantly less satisfied with the areas of culture, infrastructure, and science and research, while men report less satisfaction with crime, healthcare, safety, and unemployment. These results are largely similar to those of previous studies: women tend to care more about areas related to culture and infrastructure, while men are more concerned with protection and labor issues [Bagues and Campa, 2017; Chattopadhyay and Duflo, 2004].

Table 1 Survey results: Satisfaction with selected areas of public life

	Satisfaction (3-level)			Satisfaction (5-level)		
	Male	Female	Difference	Male	Female	Difference
Agriculture	0.113	0.104	0.010	0.210	0.205	0.005
Crime	0.180	0.226	-0.046**	0.298	0.350	-0.052*
Culture	-0.555	-0.598	0.042**	-0.691	-0.759	0.068***
Economy	0.234	0.235	-0.001	0.386	0.374	0.012
Education	-0.125	-0.149	0.024	-0.116	-0.154	0.037
Environment	-0.284	-0.266	-0.018	-0.299	-0.276	-0.022
Goods and services	-0.687	-0.707	0.019	-0.926	-0.970	0.044*
Healthcare	-0.032	0.019	-0.051**	0.016	0.078	-0.062**
Infrastructure	-0.280	-0.322	0.042*	-0.320	-0.378	0.058**
Living conditions	0.063	0.072	-0.009	0.143	0.155	-0.012
Safety	-0.133	-0.092	-0.041*	-0.108	-0.053	-0.055*
Science and research	-0.433	-0.468	0.035*	-0.500	-0.538	0.037
Social security	0.416	0.438	-0.022	0.674	0.697	-0.023
Unemployment	0.412	0.452	-0.040*	0.673	0.747	-0.074**

Note: The table reports the results of a t-test of satisfaction with selected areas of public life between males and females. The areas are listed in the left-most column. The first block of three columns reports the satisfaction on a 3-level scale (satisfied = 1, neutral = 0, dissatisfied = -1) while the second block uses a 5-level scale (very satisfied = 2, somewhat satisfied = 1, neutral = 0, somewhat dissatisfied = -1, very dissatisfied = -2). The difference columns use males as a reference category. A negative (positive) difference is therefore associated with lower male (female) satisfaction.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Although the survey does not directly ask about the policy preferences in particular areas, it is reasonable to assume that the dissatisfaction with a particular area can be used as a proxy for the preference for a policy change in that area: if a person is dissatisfied with a particular area of public life, they are more likely to prefer policy action in the area. Therefore, I use these survey-based results on satisfaction to construct my dependent variable (described in greater detail in section 4) for the main estimation. Using the areas where males and females significantly differ, I classify the public expenditures of each municipality into three groups: male, female and neutral. Female expenditures include all spending items associated with culture, science and research, and infrastructure. Male budget items are those intended for protection, safety, crime representativeness (for instance, educated women are oversampled compared to educated men).

prevention and public order, and employment measures. All the remaining items in the budget are classified as neutral. This approach is comparable to that of [Chattopadhyay and Duflo \[2004\]](#) and [Bagues and Campa \[2017\]](#). [Chattopadhyay and Duflo \[2004\]](#) use complaints raised by female residents as a basis for defining the areas in which the desire for policy differs along gender lines. Similarly, [Bagues and Campa \[2017\]](#), utilize survey data for Spanish residents, asking them to list three main public policy problems that affect the respondents the most, estimate the differences between men and women, and classify the expenditures accordingly.

A major difference in the survey-based results above from previous literature is the absence of a statistically significant difference between males and females in the satisfaction with education. The area of education policy, along with childcare, is commonly reported as an area particularly important to women [[Bagues and Campa, 2017](#); [Baskaran and Hessami, 2018](#); [Funk and Gathmann, 2015](#); [Svaleryd, 2009](#)]. As women tend to be the primary care givers, the availability of childcare centres and provision of education directly affects the ability of women to participate in the workforce. In the Czech Republic, the educational facilities also include pre-school and nursery centres, and hence they fall into the education section of the public finances. Hence, I follow the existing literature for an alternative specification in which I use only the share of expenditure on education as a dependent variable instead of the survey-based share of female expenditures. The *literature-based* results, reported in section 6.2.2, do not largely differ from the *survey-based* baseline estimation with the share of female expenditures. I fail to find an effect of increased female representation using either of the approaches.

4 Data

4.1 Spending data

The relevant spending data was obtained from the Monitor – the information portal of the Ministry of Finance of the Czech Republic.²⁷ The Monitor portal collects and publishes budget and accounting information from all levels of the public administration. The budget data for municipalities in this thesis cover the period between 2011 and 2018. The data includes a regional description of the municipality (identification number, district and region) where the ID matches the municipal ID used in the election data. The data also contains a granular description of each spending entry in terms of the type of transaction and the area of spending. This enables me to select only the expenditure part of the budget and classify the expenditures into female, male and neutral as described in section 3.2.

Table 2 reports the summary statistics of the spending data for all municipalities between 2011 and 2018. The average amount spent equals 159 million Czech Crowns (approx. 6 million Euro) for the whole municipality. However, the mean is driven by the large cities with sizeable budgets – the median municipality spent less than 15 million Czech Crowns (approx. 500 thousand Euro). Average annual municipal spending per capita is equal just above 45 thousand Czech crowns (approx. 1700 Euro). A municipality spends on average 13% of its budget on female expenditures, out of which infrastructure and communication is the largest item (accounting for

²⁷The Monitor portal can be found at <https://monitor.statnipokladna.cz/>; see details in Appendix D.

Table 2 Summary statistics: Spending data by gender and area

	Mean	St.dev.	Median
Total spending (in mil. CZK)	159.34	3509.60	14.13
Spending per capita (in thous. CZK)	45.10	62.66	33.20
Share female spending	0.13	0.13	0.09
Culture, religion and media	0.04	0.06	0.02
Infrastructure and communication	0.09	0.11	0.04
Research and development	0.01	0.02	0.00
Share of male spending	0.03	0.05	0.01
Civic security and protection	0.02	0.05	0.01
Employment	0.04	0.04	0.03
Healthcare	0.00	0.02	0.00
Share of neutral spending	0.84	0.13	0.88
Administration	0.39	0.18	0.36
Agriculture, forestry and fishing	0.03	0.06	0.01
Economic matters	0.00	0.01	0.00
Education	0.09	0.10	0.06
Environmental protection	0.10	0.08	0.08
Government organs	0.00	0.00	0.00
Housing and urban services	0.12	0.11	0.08
Industry, energy and trade	0.02	0.04	0.00
Other community and social services	0.03	0.06	0.01
Social welfare	0.01	0.04	0.00
Water management	0.10	0.16	0.04

Note: The sample contains 49,980 observations from two legislatures (2011-2014 and 2015-2018).

about 9% of the total budget). Male expenditures are on average 3% of the budget, and neutral spending accounts for the remaining 84%.²⁸ The largest spending area is administration with on average nearly 40% of the total municipality budget. Distant but still sizeable are then the areas of housing and urban services, environmental protection, water management and education with 12%, 10%, 10% and 9%, respectively.

Municipality elections occur every four years, therefore, each council approves four annual budgets during its term. Since I am interested in the differences in female oriented spending when the councils change in the same municipality, I create four-year averages of the share of female expenditures in a municipality as my main dependent variable.²⁹ Note that I do not exclude the election years in my baseline estimation, as some previous studies do.³⁰ Since the elections only occur in October, and the rules for adjustments to the existing budget are strict, the newly elected council has little power to make amendments to the budget in the election year. How-

²⁸Note that the female-oriented section of the budget is larger than the male-oriented section. Though this might seem to indicate that females are actually better represented, I argue that it is likely the contrary. The survey results which are used to construct the male and female sections suggest *dissatisfaction* with the particular policy areas. Therefore, the fact that female expenditures account for a larger portion of the budget indicates that women are dissatisfied with larger portion of the budgetary and policy decisions: the larger share likely means that female concerns are represented to a lesser extent in the existing budget.

²⁹For 2010 and 2014 elections, I create average spending in the periods 2011-2014, and 2015-2018, respectively

³⁰See for instance, [Bagues and Campa \[2017\]](#).

ever, I do repeat the estimation excluding the election years as a robustness check in section 6.3 and show that the results are comparable.

4.2 Election data

The election data was retrieved from the official Czech election portal ‘Volby.cz’, managed by the Czech Statistical Office.³¹ The data covers two election cycles: the 2010 and 2014 municipal elections. For each candidate, the data shows the name, their initial and final rank, the number of votes the candidate obtained, and whether they won a mandate. Additionally, the candidate characteristics provided are age, profession, political affiliation, and education titles.³² Although the data does not include the gender of the candidates, I am able to impute it based on the candidate’s first and last name, and occupation.³³ Furthermore, on a slate level, the data provides information on the total number of votes and mandates the slate received, and on a municipality level the data shows the total available number of seats, a total number of slates and candidates, and the municipality population in the election year.

Table 3 presents municipality and candidate level summary statistics. The average council has less than 10 members with an average of 30 candidates from 4 slates competing in an election. The table further reports that the average share of female candidates equals 30% while only 27% of councillors are female, indicating that the female success rate is worse than that of men. Furthermore, the voters tend to prefer educated candidates: although only 18% of the candidates hold a university degree, 25% of the councillors do. As mentioned in section 3, a large majority of candidates has no political affiliation: more than 50% of the municipal elections feature *only* non-partisan candidates.

Table 3 Summary statistics: Municipal elections

	Mean	St.dev.	Median
Council size	9.56	4.48	7.00
Number of candidates	33.63	53.70	17.00
Number of candidates per seat	2.82	2.09	2.00
Number of slates	4.31	3.65	3.00
Share female candidates	0.30	0.13	0.30
Share female councillors	0.27	0.16	0.29
Share educated candidates	0.18	0.14	0.16
Share educated councillors	0.25	0.22	0.22
Share party-affiliated candidates	0.08	0.11	0.00
Share party-affiliated councillors	0.09	0.15	0.00

Note: The sample contains the data from two elections, 2010 and 2014. The sample size is 12,483 which includes 6,248 and 6,235 municipalities in 2010 and 2014, respectively.

³¹Retreived from <https://volby.cz/opendata/opendata.htm>; see details in Appendix D.

³²Although it is not compulsory for candidates to disclose their academic titles, completed university education is widely perceived as a desirable characteristic in a candidate. Therefore, it is safe to assume that anyone who has a university degree would disclose their title on the ballot.

³³In the Czech language, women traditionally have different endigns to their surnames (e.g. the male surname Novák becomes Nováková for a female). Similarly, many occupations have ‘feminine’ suffixes as well (e.g. a male teacher is ‘učitel’ while a female teacher is ‘učitelka’). This allows me to determine the gender of each candidate with high precision.

I create the RD sample by selecting only the contests where the last winning and first losing candidates within a slate were of opposite gender, called the *gender-mixed contests*. For these pairs of candidates in the gender-mixed contests, I calculate the vote margin of the female to the male candidate. The method for margin calculation is described in detail Appendix B.2. Given the focus on marginal contests *within slates*, some municipalities may have multiple marginal races. Table A.1 displays the summary statistics of the municipal elections by the number of gender-mixed contests. Across the two election cycles, 6,640 municipalities had *at least* one mixed contest with the majority of the municipalities having *only* one mixed contest. The total number of mixed contest in the two elections is equal to 10,686. More mixed contests can be found in elections for larger councils with more parties running, since a municipality can have as many mixed contests as there are seats or parties running (whichever is lower).³⁴ A roughly similar pattern can be found for municipality population, as the council size is determined by the municipality population. The average share of females on the council does not seem to differ across the municipalities with different number of mixed contests. However, with an increasing size of the council, the effect of an additional woman on the overall share of women declines.

4.3 Demographic data

Municipality-level demographic data was collected from the Czech Statistical Office in order to test the validity of the RD design. Czech Statistical Office provides yearly data on municipality population as part of the project ‘The Small Lexicon of the Municipalities’ (*Malý Lexikon Obcí*).³⁵ Furthermore, I retrieved the 2011 Census (*Sčítání lidu, domů a bytů 2011*) results for further demographic information, such as nationality, religious affiliation, employment, and education of the municipality residents.³⁶

5 Methodology

In an ideal world, estimating the effect of female representation on municipal spending would be possible using the share of female councillors as an independent variable, where the share of female councillors would be randomly assigned (and thus exogenous). The relationship between the share of female representatives and municipal spending could then be estimated using the following specification:

$$y_{i,t} = \alpha + \beta \cdot \text{share_females}_{i,t} + \gamma \cdot \mathbf{X}_{i,t} + \epsilon_{i,t} \quad (5.1)$$

where $y_{i,t}$ is the outcome variable, that is, the share of female spending in the municipality i in year t , $\text{share_females}_{i,t}$ represents the share of females on municipal council i in time t and $\mathbf{X}_{i,t}$

³⁴For instance, a 5-seat council with 5 parties running can have maximum 5 mixed contests, while a 7-seat council with 7-parties running can have 7-contests. Notably though, a 7-seat council with 6 parties running cannot have more than 6 contests. Similarly, the same 7-seat council with 8 parties running can have only 7 mixed contests. The likelihood of having a gender-mixed contests thus increases with the number of mandates and number of parties running, whichever is lower.

³⁵See Appendix D for the yearly Lexicon data sources.

³⁶Retrieved from <https://www.czso.cz/csu/sldb>, see details in Appendix D.

stands for a set of municipal-level covariates. In this setting, the coefficient β would identify the effect. However, in a real world setting, where the assignment of women to the council is not random, estimating the relationship from Equation 5.1 using simple OLS would likely suffer from endogeneity issues. In reality, the share of females on the council is determined by the election parties who choose the candidates and their rank on the slate, and the electorate who casts votes for the candidates. The underlying characteristics of municipalities and their preferences for female councillors thus play a role in selecting male or female candidates. Additionally, the same characteristics can affect the spending of the municipality regardless of the gender representation on the council. For instance, municipalities with more educated and progressive constituents might be more keen to elect more female candidates. At the same time, the educated constituents may demand higher spending on the areas of research or culture, which belong to the female spending. This spending would therefore be driven by the characteristics of the municipality and not by the additional females in the council; the municipality would likely spend more on culture and research even in the absence of women on the municipal council. Failing to include this characteristic in the OLS estimation would therefore cause omitted variable bias. In this example, the underlying progressivism is positively correlated with both the share of females and female spending, hence the bias would be upwards.

5.1 RD Framework

To solve the endogeneity issue, I use an RD design, exploiting the unpredictability of the minimum number of votes needed to gain a seat on the council. The number of mandates for each slate is determined by the overall number of votes the slate gained. Therefore, within each slate, there will be a candidate who wins the last mandate and barely gets into the council, and a candidate whose final rank is one position lower and ends up without a mandate. Of these marginal contests, I restrict my focus to the *gender-mixed contests*, which are defined as the contests between the marginal winner and loser within a slate that are of *opposite gender*.³⁷ These mixed-gender contests ensure that the last mandate is allocated in a quasi-random way: the vote difference between the marginally winning and losing candidate is usually quite small, thus some municipalities with gender-mixed contests will virtually by chance end up with more women on the council while others will have a stronger male presence.

The five core ingredients for any RD design are the following: the unit of analysis, running variable, treatment, cut-off, and outcome. In my specification, the *unit of analysis* is the mixed-gender contest within a slate for the last seat on the municipal council. The *running variable* is the margin of victory of a female candidate, that is the difference between the vote percentage obtained by the female candidate to the vote percentage obtained by the male candidate (for more details see Appendix B.2). The margin of victory can hence be either positive or negative, where a positive margin indicates that the female candidate was elected, while the negative margin indicates a male winner. The *cut-off* for winning a seat is therefore 0, by definition. The *treatment* is defined as obtaining a seat on the municipal council, while the control is not winning a seat. The *outcome* is the average share of female spending across the four years of a

³⁷Notice that those competitions are for the last seat within a *slate*. Since there are normally multiple slates within one election, there are as many *contests* as there are *slates*, but not all of the contests are gender-mixed.

municipality council in office (described in section 4).

I can implement a sharp RD design for the baseline specification as the election outcomes are the sole determinant of who wins a council seat. In the sharp RD, the probability of being treated (i.e. receiving a seat on the council) is equal to 1 for all vote margins that are higher than the cut-off, while all candidates with negative vote margins are automatically assigned to the control group (where treatment is equal to 0). The sharp RD specification is the following:

$$y_{i,t} = \alpha + \beta \cdot \text{winner}_{i,t} + f(\text{vote_margin})_{i,t} + \text{winner}_{i,t} \times g(\text{vote_margin})_{i,t} + \gamma \cdot \text{year} + \epsilon_{i,t} \quad (5.2)$$

where $y_{i,t}$ is the outcome variable, in the main specification equal to the average share of female spending of municipality i in election cycle t , winner is a dummy indicator equal to one if a female wins the gender-mixed contest and zero otherwise, and the vote_margin represents the vote margin of the female to the male candidate. The coefficient of interest is identified by β . Functions $f()$ and $g()$ are polynomial functions of the vote_margin variable, which are allowed to vary at the cut-off. Since the data spans two election cycles, I control for election year to account for changes in spending patterns over time. Finally, standard errors are clustered at the municipality level as my data can include more than one mixed contests in each municipality.

The baseline model is estimated using the local polynomial method with an optimal bandwidth and first order polynomial (linear fit). Consecutive estimations with different bandwidths and different order of the polynomial fit are used to test the sensitivity of the results.³⁸ It is noteworthy that the RD sample can contain municipalities which have multiple mixed contests in the same election round. This means that some municipalities might have multiple marginal winners. The marginal winners from the different mixed contests can either have the same gender, which could amplify the results, or opposite genders which could cancel out some of the effects. To mitigate this issue, I use the approach presented in [Baskaran and Hessami \[2018\]](#): I include all mixed contests in a municipality-year in the sample, where each has an individual vote margin but all of them have the same value of the outcome variable. To correct for the amplification or cancellation in the estimation, I weigh each of the observations by the inverse of the number of mixed contests in the municipality.³⁹ Furthermore, as a robustness test, I re-run the specification restricting my sample to the municipalities with only one gender-mixed contest. The findings remain largely unchanged.

The main drawback of using the sharp RD design comes with the nature of the effect it estimates – the effect estimated using sharp RD does not differentiate between one additional woman on a small and a large council, which might influence the ability to find an effect. One additional female on a small council could have a much larger effect on the outcome variable than one additional woman in a large council. To solve this issue, I carry out an additional estimation where I restrict my sample only to small municipalities and show the results in section 6.3. In

³⁸All RD analysis is conducted using *Stata*, in particular, the packages *rdrobust*, *rddensity*, and *rdpower* developed by Calonico, Cattaneo, Farrell, Jansson, Ma, Titiunik, Vazquez-Bare. See [\[Calonico et al., 2017, 2014a, 2015\]](#).

³⁹For instance, an observation in a municipality with 4 mixed contests would be weighed by 0.25.

five estimations, I allow the council to be the maximum size of 7, 9, 11, 15 and 21 councillors, and compare the results with the baseline estimation. The restricted estimation produces results similar to the main estimation. I use the full, unrestricted sample for the baseline model, as it provides the largest number of observations.

5.1.1 Extensions to the baseline model

In addition to the sharp RD, I perform additional estimation using fuzzy RD. In the fuzzy RD framework, the probability of receiving treatment does not change from 0 to 1 at the cut-off, creating imperfect compliance with the treatment assignment. I use this framework to estimate the effect of the share of female councillors on the share of female spending, using the vote margin as an instrument for the share of female councillors. The idea is that the share of female councillors jumps discontinuously at the cut-off of the vote margin: if a female wins a mixed contest, the share of females on the council substantially increases. The share of females, however, does not strictly change from 0 to 1, creating the imperfect compliance required by the fuzzy RD method. The share of female councillors is then used as an explanatory variable for the average share of female spending.

The main estimation is done using sharp RD rather than fuzzy RD due to a number of reasons. Firstly, the sharp RD results have considerably higher power than the fuzzy RD (for more details see the power calculations in Appendix C). Secondly, the sharp RD estimation is clearer and simpler to implement and interpret. Finally, it is better to use sharp RD for the sake of methodological clarity, as the local polynomial method is easier explained and illustrated for sharp RD than fuzzy RD.

As a second extension, I repeat the analysis using the share of female spending in differences between the legislatures, following the method used by [Bagues and Campa \[2017\]](#). This is done to eliminate the time-invariant, municipality-specific differences in spending which could affect the consistency of the RD estimates. In particular, I run the specification from Equation 5.2 using the change in the average share of female spending between the legislatures elected in 2010 and 2014, or Δy_i . Since I only have data on two legislatures, I restrict my sample to the gender-mixed contests from 2014. I then take the difference between average female spending of the current legislature (2015-2018) and the previous legislature (2011-2014) which was elected in 2010, thereby creating the new dependent variable Δy_i .

Lastly, as described in section 3.2, the last extension of the model uses the education spending as a dependent variable. Analogous to the baseline estimation, I create four-year averages of the education spending for the whole term of a legislature and use this average share of education spending in the estimation from Equation 5.2.

5.2 Local polynomial approach with optimal bandwidths

I estimate the model from Equation 5.2 using the local polynomial (or continuity-based) approach to RD described in [Cattaneo et al. \[2019a\]](#). This method is *local* as it only uses the observations close to the cut-off, and *polynomial* because it fits a polynomial function at both sides of the cut-off. Due to its local nature, this method is non-parametric as opposed to a

global fit using all observations, which would be parametric. The main benefit of the non-parametric approach is the flexibility and the desirable properties in estimating the treatment effect, mainly the robustness of the estimates and lower sensitivity to outliers. The parametric (global) approach can deliver poor and misleading approximation of the effect at the boundary point as the observations far from the cut-off can have disproportionate influence on estimating the underlying regression functions [Cattaneo et al., 2019a].

This approach involves a specifications of three core factors: the kernel, the polynomial order, and the bandwidth. The kernel determines the weight assigned to the observations on either side of the cut-off. One can either use a uniform fit that weighs all the observations equally, or a triangular or Epanechnikov fit, which uses linear and quadratic weights, respectively, giving the observations closer to the cut-off relatively more importance. Based on the suggestion from Cattaneo et al. [2019a], I use the triangular kernel, which in combination with the optimal bandwidth choice delivers a point estimate with desirable properties: it is both asymptotically optimal and consistent. The local polynomial order decides the fit of the regression line on both sides of the cut-off. An order that is too high often overfits to the data, while a polynomial of order 0 (constant) is not sensitive enough to the data. Therefore, I use a linear fit for the baseline specification, and a quadratic fit as a sensitivity check.

The most important, however, is the choice of the bandwidth, denoted h , that determines which observations are included in the local polynomial estimation. The bandwidth is the maximum distance of the running variable from the cut-off that an observation can have in order to be included in the estimation. The observations that are further away, or outside of the bandwidth, are not used for estimation (they are assigned a zero weight). Following the proposed technique in Cattaneo et al. [2019a], I retrieve the point estimate from Equation 5.2, using the bandwidth h_{MSE} , which minimizes the mean square error (MSE) of the treatment effect. This involves a trade-off between variance and misspecification bias: as the bandwidth gets smaller, the point estimate gets more precise, but the variance increases, while as the bandwidth extends the variance drops at the expense of an error in the point estimate.⁴⁰

However, the MSE optimal bandwidth is not suitable for valid causal inference because the choice of the bandwidth depend on the bias-variance trade-off. Proceeding with the conventional confidence intervals (which treat the local polynomial approximation as parametric) for inference as if the misspecification bias did not exist would then produce incorrect inferences. A methodologically sound approach, recommended by Cattaneo et al. [2019a], is to use a robust bias correction approach which constructs a new confidence interval, correcting for the misspecification bias. The new robust bias-corrected confidence interval changes both in terms of its centre and length: it is centered around the bias-corrected point estimate, and it is longer as the variance terms accounts for the increased variability induced by the bias-correction (if using the same bandwidth). In practice, the robust bias correction is carried out by using a new bandwidth, h_{CER} , that minimizes the coverage error which arises when the confidence interval does not capture the true value of interest [Calonico et al., 2018, 2014b]. Therefore, I

⁴⁰As MSE is defined as $MSE(\hat{\beta}) = bias^2(\hat{\beta}) + variance(\hat{\beta})$, choosing h that minimizes MSE gets the the best trade-off.

use the MSE-optimal bandwidth, h_{MSE} , for point estimation, while constructing the confidence intervals using the CER-optimal bandwidth, h_{CER} , which yields valid inference.

5.3 Validity of RD design

The initial condition for the use of RD is that the probability of being assigned to the treatment group is discontinuous at the cut-off. In my specification, it means that the probability of winning a seat must jump discontinuously at the zero margin threshold. I assess whether my data fulfills this condition using simple summary statistics and find that this condition is satisfied: all observations with a positive vote margin report a female win, while those with negative margin show a female loss. Table A.3 presents summary statistics of the vote margin by the gender of the winner, revealing that all female winners indeed had a positive margin, while all male winners report a negative margin. The narrowest margin for a female winner equals 0.005 percentage points above the male candidate, while the closest margin for a male winner is -0.003.⁴¹ The table further indicates that a larger portion of the marginal contests were won by men (5,562 vs 5,123 won by women). The absolute value of the median vote margin is very similar for both genders.

The identifying assumption for the use of the RD design requires that the only determinant of the outcome which varies discontinuously at the cut-off is the (probability of) treatment assignment; no other observed and unobserved characteristics should change discontinuously at the cut-off. In the context of my research, this assumption demands that the municipalities which marginally elected a male or a female are do not differ in other aspects apart from the gender of the marginally elected councillor. Additionally, the marginally winning and losing candidates should only differ in their council seat allocation (one gets a council seat, the other one does not) and their gender (by construction of my sample); they should be similar in all other characteristics.

The identifying assumption is at least partially verifiable using a set of validation and falsification tests, as suggested by Cattaneo et al. [2019a]. The testable implications of the identifying assumption are the following:

1. The units close to the cut-off do not differ in their predetermined covariates.
2. The units lack the ability to manipulate the score and thus increase their probability of being treated.
3. The regression function for the treatment and control units would be continuous in the absence of the treatment.

I translate each of the testable implications into my particular research context and explain the tests I use to validate my research design.

The first implication demands that the municipalities with a marginal female and marginal male councillor are similar in their characteristics, that is, no other possible determinant of the

⁴¹The -0.003 vote margin indicates that the female candidate received a share of votes 0.003 percentage points *lower* compared to the male candidate.

outcome should jump discontinuously at the cut-off. Furthermore, it is also necessary that the candidates who barely got into the council are not distinct from the candidates who barely lost the race for the last seat. This is tested by running a local polynomial regression from the main specification (Equation 5.2), with predetermined covariates as the dependent variable. These covariates should not be affected by the elections in any way (as they are *predetermined*) and thus, the results and the accompanying RD plots should show no discontinuity.

Table A.4 presents the results of the predetermined covariate tests. Panel A shows the selected candidate-level covariates age and education of the male and female candidates, and confirms that the candidates just above and just below the cut-off do not statistically significantly differ in those characteristics. Panel B reports the municipality level covariates (log population in the election year, council size, number of slates in an election and the number of mixed contests in an election), which all show no sign of discontinuity at the cut-off. Finally, using the 2011 Census data, I analyse the balance of further socio-demographic covariates focusing on age composition of the municipality and share of residents based on their citizenship, religious affiliation, education, and employment, and report the results in Panel C. The covariates are again not showing any evidence of discontinuity. Note that as the Census data were collected only in 2011, the variables cannot serve as a *predetermined* covariate for the 2010 election round, therefore I only use the 2014 election data in Panel C. The RD plots for the covariate balance tests can be found in Figures A.1, A.2 and A.3.

Demonstrated by the second testable implication, it is crucial for the RD design that the votes of the candidates cannot be manipulated such that a candidate with a negative margin ends up winning a seat. The ability to influence the vote margin is unlikely due to stringent rules and close monitoring of the vote-counting process, and the strictly determined number of mandates each slate obtains. Nonetheless, formally testing the second implication adds further credibility to the claim that manipulation can be ruled out.

The second implication can be assessed by the RD density test, which tests whether the density of the observations is continuous around the cut-off. If score manipulation is possible, the density of cases would be lower just below the cut-off and higher just above, and hence, the test would reveal that the density is not continuous. In my case, the test reports the p-value of 0.65, which means that I cannot reject the null hypothesis of continuous density of observations at the cut-off. Figure A.4 shows the histogram of the observations, and the RD density plot, which further confirms the absence of manipulation of the score around the cut-off.

The final implication of continuity of the regression functions around the cut-off in the absence of the treatment cannot be empirically tested. However, Cattaneo et al. [2019a] suggest at least a partial verification, by testing whether the regression functions are continuous at other points away from the cut-off using a set of arbitrary placebo cut-offs. In this test, I run the specification from Equation 5.2, but redefine the cut-off to vote margin values above or below the initial zero cut-off. If the new cut-off is positive, I restrict my observations only to the treated units (where vote margin > 0), while if the new cut-off is negative, I include observations only from the control group (where vote margin < 0). I repeat this specification for a set of cut-offs between -0.5 and 0.5 with increments of 0.1. The results of the test, reported in Table A.5, are all statistically

insignificant, indicating that the share of female spending does not jump discontinuously at any of the alternative cut-offs.

[Cattaneo et al. \[2019a\]](#) also suggest using sensitivity tests to ensure that a particular set of observations does not drive the results. The two main sensitivity checks proposed are the ‘Donut hole’ approach and the alternative bandwidth choice. The ‘Donut hole’ approach tests the sensitivity of the estimation to the observations near the cut-off. Using this approach, I run the RD specification from Equation 5.2 but omit a group of observations in the close neighborhood of the cut-off. I repeat this process for multiple values of the radius around the cut-off and show the results in section 6.3, along with various robustness checks. Conversely, the use of alternative bandwidths tests the sensitivity when observations are added or removed at the outer boundary of the bandwidth. Although I report the regular half- and double-bandwidth estimation with the main results, I also present the results of the estimation for a larger variety of other bandwidths in section 6.3, revealing that the conclusions are largely robust to the bandwidth choice.

6 Results

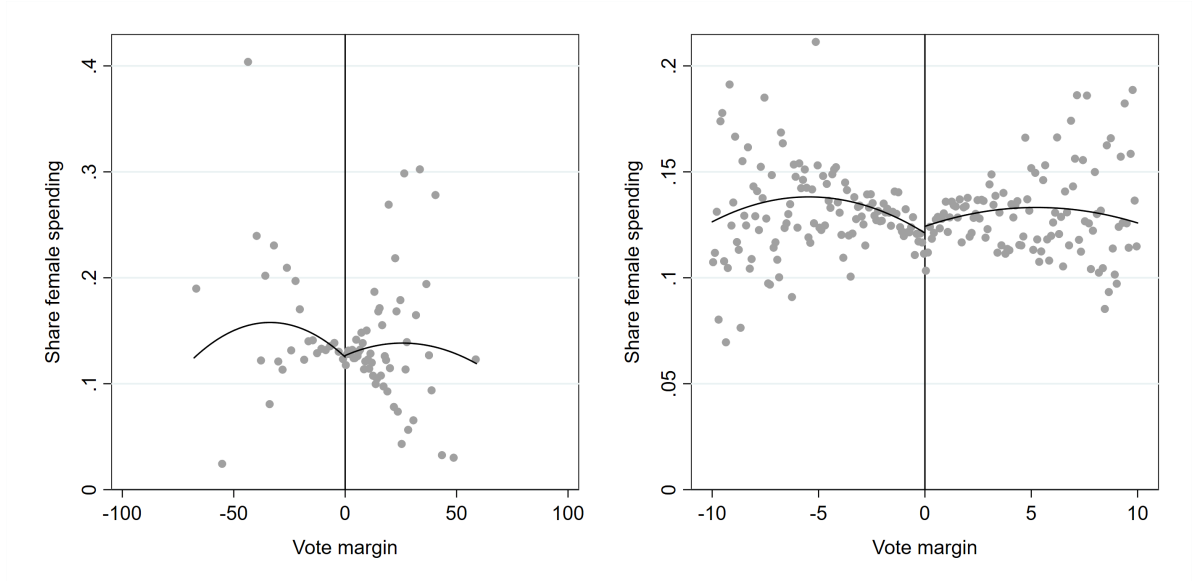
Before I proceed with the RD analysis, I estimate the associations between the share of female councillors and the share of female spending. I use OLS to retrieve these naive estimates and report the results in Table A.2 in Appendix A. Model 1 reports the estimation using the full RD sample while controlling for the election year. In Model 2, I restrict my sample only to the 2014 election round and control for a lagged value of the share of female spending in the 2010 election round. Neither of the estimations show a statistically significant relationship between the share of female councillors and the share of female spending. These results differ from many of the previous studies looking at the relationship between the female representation and the budget indicators which find a positive associations with OLS, but cannot find a causal relationship using the RD framework (for instance [Bagues and Campa \[2017\]](#) or [Kuliomina \[2019\]](#)).

6.1 Baseline estimation

For my baseline model of the effect of marginal female councillors on the share of female spending in the municipal budgets, I report three kinds of results: RD plots, baseline specification with a selection of different bandwidths and polynomials, and finally, the effects for individual years of the councils’ term.

First, I present RD plots which allow for visual inspection of the RD estimation. Using global RD plots, I show the second order polynomial fit of the share of female spending on the vote margin of the female candidate, where the polynomial is separately fitted above and below the zero cut-off, and the local means of the outcome variable are displayed for numerous vote margin bins. Figure 1 shows the global RD plots with a quadratic fit, where on the left I use the full support of the data (vote margin between -68 p.p. and 59 p.p.), whereas the right graph uses a smaller vote margin bandwidth of 10 p.p. Figure 2 displays a local RD plot with a linear fit using the optimal bandwidth, depicting the results of the estimation shown in Table 4. Inspecting

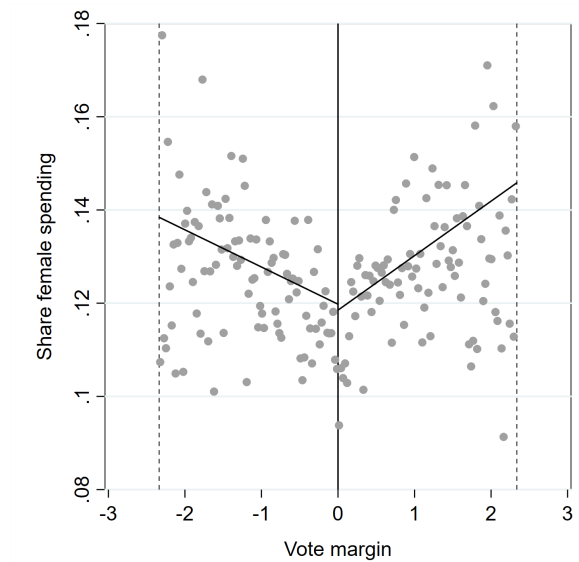
Figure 1 Sharp RD: Global RD plot of female victory and the share of female spending



Note: Global RD plot of the vote margin against the average share of female spending. The left figure shows the full global plot with the vote margin ranging from -68 to 59, while the right plot restricts the vote margin to -10 to 10.

the graphs suggest that the the existence of a discontinuity in the outcome variable at the the cut-off is rather unlikely.

Figure 2 Sharp RD: Local RD plot of female victory and the share of female spending



Note: Local RD plot of the vote margin against the average share of female spending. The solid vertical line represents the zero vote margin and the dashed vertical lines depict the optimal bandwidth.

Second, I present the baseline specification in Model 1 of Table 4, using the MSE optimal bandwidth (described in section 5.2) and a linear fit on both sides of the cut-off. The outcome variable, as described in section 4 is the average share of female spending across the 4-year term period. The bandwidth is equal to 2.337, meaning that the maximum vote margin between the

female and male candidate is 2.337 percentage points. This optimal bandwidth choice leaves about 3,500 efficient observations on each side of the cut-off. The results suggest that the additional elected female had no statistically significant effect on the share of female spending compared to the control group average of 12.4%: the point estimate is equal to -0.13 p.p., and based on the robust-adjusted 95% CI, I can reject an increase of more than 0.6 p.p. (8% st. dev)⁴² or a decrease of more than 1 p.p. (14% st. dev). In Models 2 and 3 of Table 4, I repeat the estimation using half and double of the optimal bandwidth, respectively, and Model 4 presents a quadratic instead of linear fit; both the mean of the control group and the point estimate remain largely unchanged in all three models, and all the results remain statistically insignificant.

Table 4 Sharp RD: Female victory and the average share of female spending

	(1)	(2)	(3)	(4)
Female victory	-0.0013 (0.0036)	-0.0062 (0.0044)	0.0011 (0.0030)	-0.0023 (0.0040)
Polynomial	Linear	Linear	Linear	Quadratic
Bandwidth type	Optimal	Half	Double	Optimal
Bandwidth size	2.337	1.168	4.674	4.394
Mean control group	0.124	0.120	0.126	0.125
Robust CI	[-.01 ; .006]	[-.022 ; .002]	[-.01 ; .005]	[-.011 ; .005]
Observations	10685	10685	10685	10685

Note: The dependent variable is the average share of female spending across the 4-year legislative term. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of gender-mixed contests})$ to correct for municipalities with multiple mixed contests. Model (1) and (4) use optimal bandwidth (based on Calonico et al. [2014b]), while Models (2) and (3) use half and double of the optimal bandwidth, respectively. Models (1)-(3) use linear polynomial while the polynomial in Model (4) is quadratic. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The third type of results is the effect of marginally elected female candidate on the share of female spending in the individual years of the council's term.⁴³ Models 1 to 4 of Table A.8 show the effect in the individual years. For Model 5, I use the observations from all individual years in a single estimation, meaning that each election outcome appears four times in the sample, once for each year of the council's term. In each of the years apart from year 1 the point estimate is negative, suggesting that an additional female in the council has a negative effect on the share of female spending. This effect is, however, statistically insignificant and therefore, I cannot reject the null hypothesis that the share of female spending does not differ above and below the cut-off. The estimation for all years from Model 5 is largely similar to the results from Model 1 from the baseline estimation in Table 4. The point estimate of -0.13 p.p. is not statistically significant and therefore I can rule out that an additional female increases or decreases the outcome by

⁴²The boundary values of the 95% CI are calculated into % st. dev. using the standard deviation of the control group.

⁴³For instance, year 1 is the first year of the term, that is 2011 and 2015 for the 2010 and 2014 elections, respectively.

more than 0.6 p.p. (5% st. dev) or 1 p.p. (9% st. dev), respectively.

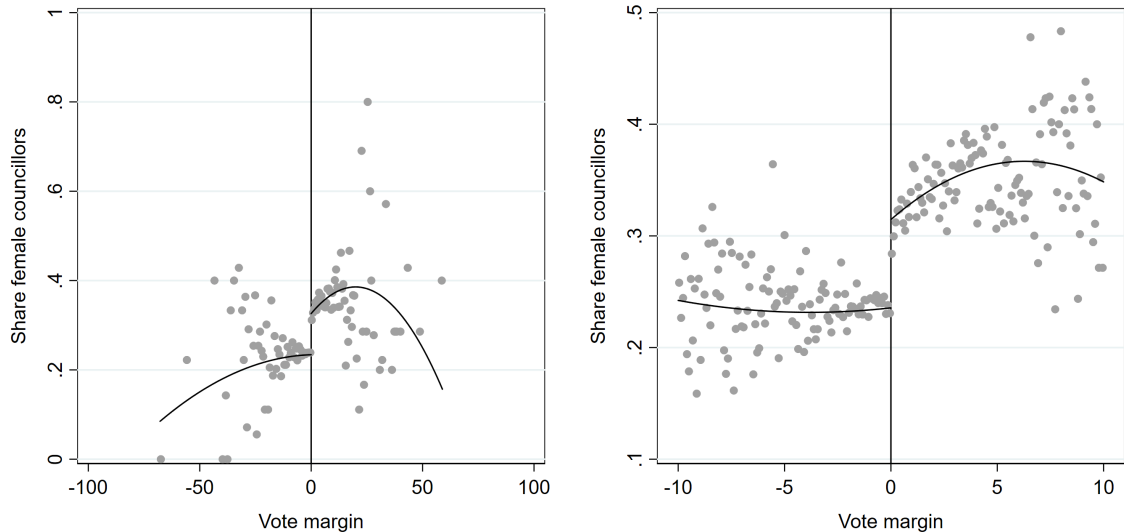
6.2 Extensions

In this section, I present the results of a number of further estimations. Firstly, I perform fuzzy RD using the share of females as the independent variable. Secondly, I use education spending as a dependent variable in the baseline sharp RD estimation. Finally, I carry out an estimation with the dependent variable in differences, where I use the difference in the average share of female spending between the 2014- and 2010-elected legislatures.

6.2.1 Fuzzy RD

This section summarizes the results of the fuzzy RD estimation. I use the vote margin as an instrument for the share of females on the council to estimate the structural relationship between the share of female councillors and share of female spending. The global RD plots in Figure 3 indicate that indeed a female winner has a positive effect on the share of female councillors. Table 5 confirms these results, showing a statistically significant increase in the share of female councillors of about 7.6 p.p. in the councils with at least one marginally elected woman compared to the mean of 23.8% in the control group, which translates into a 31.9% increase in female representation. Model 1 reports the results for the first stage using the optimal bandwidth from the full fuzzy RD estimation, while Models 2 to 5 are estimated as a separate sharp RD with the share of female councillors as the outcome variable. The results are consistent across the five models. The local RD in Figure 4 offers visual representation of the results from Model 1.

Figure 3 Fuzzy RD: Global RD plot of female victory and the share of female councillors



Note: Global RD plot depicting the relationship between the vote margin and the share of female councillors. the left figure shows the full global plot with the vote margin ranging from -68 to 59, while the right plot restricts the vote margin to -10 to 10.

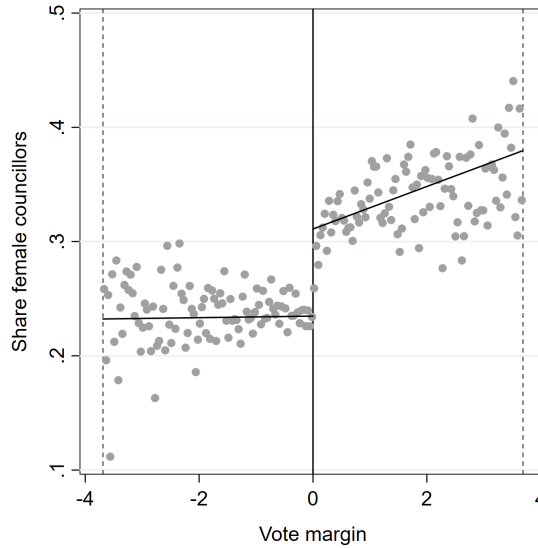
The complete results of the fuzzy RD estimation are reported in Table 6. The second stage

Table 5 Fuzzy RD: Female victory and the share of female councillors (first-stage results)

	(1)	(2)	(3)	(4)	(5)
Female victory	0.076*** (0.006)	0.071*** (0.007)	0.067*** (0.008)	0.077*** (0.006)	0.069*** (0.007)
Polynomial	Linear	Linear	Linear	Linear	Quadratic
Bandwidth type	Fuzzy optimal	Optimal	Half	Double	Optimal
Bandwidth size	3.688	2.172	1.086	4.344	4.055
Mean control group	0.238	0.239	0.239	0.238	0.238
Robust CI	[.061 ; .086]	[.055 ; .083]	[.032 ; .079]	[.056 ; .084]	[.053 ; .084]
Observations	10685	10685	10685	10685	10685

Note: The table reports the results of the first stage of the fuzzy RD estimation. The dependent variable is the share of females on the councils. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of gender-mixed contests})$ to correct for municipalities with multiple mixed contests. Model (1) uses the optimal bandwidth (based on Calonico et al. [2014b]) if ran separately as independent sharp RD design, while Models (2) and (5) use the optimal bandwidth from the second stage. Models (3) and (4) use half and double of the optimal bandwidth from the second stage, respectively. Models (1)-(3) use linear fit while the polynomial in Model (4) is quadratic. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 4 Fuzzy RD: First stage estimation plot

Note: The figure shows the local RD plot of the first stage, with the dashed vertical line depicting the optimal bandwidth and the solid line marking the zero margin.

estimation yields statistically insignificant results in all the estimated models. Model 1 reveals a point estimate of 0.3 p.p. which is in magnitude similar to that from the sharp RD, although the direction is opposite. This estimate allows me to reject an increase larger than 8.9 p.p. (124% st. dev) or a decrease larger than 8.9 p.p. (124% st. dev) in the share of female spending with an additional winning female.

The fuzzy RD is easier to interpret and generalize, as the effect of an additional female is directly

Table 6 Fuzzy RD: Share of female councillors and the average share of female spending

	(1)	(2)	(3)	(4)
Share female councillors	0.003 (0.042)	-0.035 (0.055)	0.029 (0.032)	-0.027 (0.052)
First stage estimate	0.076	0.069	0.083	0.071
Polynomial	Linear	Linear	Linear	Quadratic
Bandwidth type	Optimal	Half	Double	Optimal
Bandwidth size	3.688	1.844	7.376	5.094
Mean control group	0.125	0.123	0.126	0.126
Robust CI	[-.089 ; .089]	[-.267 ; .023]	[-.077 ; .082]	[-.145 ; .07]
Observations	10685	10685	10685	10685

Note: The table reports the results of the second stage of the fuzzy RD estimation. The dependent variable is the average share of female spending across the four years of a council's legislative term. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of gender-mixed contests})$ to correct for municipalities with multiple mixed contests. Model (1) and (4) use optimal bandwidth (based on Calonico et al. [2014b]), while Models (2) and (3) use half and double of the optimal bandwidth, respectively. Models (1)-(3) use linear fit while the polynomial in Model (4) is quadratic. The third row presents the first-stage estimates, that is the average effect of an additional female on the share of female councillors. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

scaled by the share of female councillors. However, the point estimates of the models are noisier than in the baseline sharp RD: adjusting the bandwidth and polynomial changes the point estimate by orders of magnitude. Moreover, the standard errors are also an order of magnitude larger than the baseline sharp RD results from Table 4. As a consequence, the power of the fuzzy RD is drastically lower compared to the sharp RD estimation (see Appendix C).

6.2.2 Education as a dependent variable

All the previous estimation were done using the share of female spending, constructed based on my own results from the estimation of policy preferences among males and females. However, my survey analysis did not highlight education as one of the areas which concern women more than men as was previously found in other studies (see section 3.2).⁴⁴ Therefore, I estimate the baseline model from Equation 5.2, using the education spending as a share of the total budget as a dependent variable. The results in Table 7 indicate no statistically significant difference in the share of education spending above and below the cut-off. The point estimate indicates on average a 0.4 p.p. decrease in education spending if a female is elected compared to the control mean of 9.8%, but the effect is not statistically significant. The 95% confidence interval suggests that the additional elected female does not cause a decrease of more than 1.2 p.p. (16% st. dev.) or an increase of more than 0.3 p.p. (4% st. dev.) in the share of spending on education. The results are robust across the different bandwidths and polynomials.

⁴⁴See for instance Bagues and Campa [2017]; Chattopadhyay and Duflo [2004]; Funk and Gathmann [2015]; Svaleryd [2009].

Table 7 Extensions: Female victory and the average share of education spending

	(1)	(2)	(3)	(4)
Female victory	-0.004 (0.003)	-0.004 (0.005)	-0.001 (0.003)	-0.003 (0.004)
Polynomial	Linear	Linear	Linear	Quadratic
Bandwidth type	Optimal	Half	Double	Optimal
Bandwidth size	3.251	1.626	6.503	6.326
Mean control group	0.098	0.100	0.096	0.096
Robust CI	[-.012 ; .003]	[-.019 ; .008]	[-.011 ; .005]	[-.011 ; .006]
Observations	10685	10685	10685	10685

Note: The dependent variable is the average share of education spending across the four years of a council's legislative term. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of gender-mixed contests})$ to correct for municipalities with multiple mixed contests. Model (1) and (4) use optimal bandwidth (based on [Calonico et al. \[2014b\]](#)), while Models (2) and (3) use half and double of the optimal bandwidth, respectively. Models (1)-(3) use linear polynomial while the polynomial in Model (4) is quadratic. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.2.3 Dependent variable in differences

This section presents the results of an estimation where the dependent variable is the change in share of female spending between the 2010 and the 2014 legislatures. Similarly to [Bagues and Campa \[2017\]](#), I use this approach to eliminate the time-invariant spending differences across municipalities. Table 8 reveals the results from the estimation. The point estimates are statistically insignificant as in all the previous estimations: an additional female councillor is associated with a 0.01 p.p. higher change in the average share of female spending. The 95% CI indicates that the additional female did not decrease the change in spending by more than 0.9 p.p. (12% st. dev.) or increase it by 0.8 p.p. (11% st. dev.). Similar results are found across the different bandwidths and polynomials. Notably though, the mean of the control group suggests that on average all the municipalities with a gender-mixed contest in 2014 increased significantly their share of female spending, regardless of the winner of the mixed-gender contest: for the municipalities included in the sample, the average share of female spending between 2015 and 2018 was 13.1% compared to 12.2% in the 2011-2014 legislature. This 0.89 p.p. difference between the two legislatures is statistically significant at the 1% level.⁴⁵

6.3 Sensitivity tests and robustness checks

In this section, I report the results of the sensitivity checks described in section 5.3, along with numerous robustness checks.

⁴⁵I performed a comparison of means for average share of female spending between the two legislatures in the municipalities which had a gender-mixed contest in 2014. The p-value of 0.0001 and the 95% CI of (0.007; 0.011) suggest a statistically significant difference.

Table 8 Extensions: Dependent variable in differences

	(1)	(2)	(3)	(4)
Female victory	0.0001 (0.004)	-0.0054 (0.006)	0.0003 (0.004)	-0.0015 (0.006)
Polynomial	Linear	Linear	Linear	Quadratic
Bandwidth type	Optimal	Half	Double	Optimal
Bandwidth size	2.767	1.384	5.534	4.830
Mean control group	0.009	0.009	0.009	0.009
Robust CI	[-.009 ; .008]	[-.02 ; .011]	[-.012 ; .01]	[-.013 ; .01]
Observations	5300	5300	5300	5300

Note: The dependent variable is the difference in the average share of education spending across between the 2010 and the 2014 legislative terms. The sample includes only the municipalities which had gender-mixed contests in 2014; these observations are matched with the spending of the municipality from the 2010 council term. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of gender-mixed contests})$ to correct for municipalities with multiple mixed contests. Model (1) and (4) use optimal bandwidth (based on [Calonico et al. \[2014b\]](#)), while Models (2) and (3) use half and double of the optimal bandwidth, respectively. Models (1)-(3) use linear polynomial while the polynomial in Model (4) is quadratic. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.3.1 Sensitivity checks

The first test is the ‘Donut hole’ test, which excludes a number of observations within a small radius of vote margins around the cut-off. I use nine sizes of the radius from 0.01 p.p. to 0.09 p.p. with steps of 0.01 p.p., thus excluding between 5 and 220 observations on each side of the cut-off. The results, which are reported in [Table A.6](#), do not change the conclusion: the effect of the marginal female councillor is statistically insignificant in all cases. However, with the increasing radius, the direction of the estimate changes from slightly negative to slightly positive, suggesting that the observations close to the cut-off have some effect on the results.

To perform the alternative bandwidth test, I use bandwidths of the vote margin from 1 p.p. to 5 p.p. with 0.2 p.p. steps. The results, reported in [Table A.7](#) and visualized in [Figure A.5](#), are compatible with those from the baseline specification. None of the results indicate a statistically significant effect of increased female representation on the share of female spending. Similar to the ‘Donut hole’ test, the point estimate changes from negative to positive as the bandwidth expands, and the size of the point estimate changes by an order of magnitude. This does not, however, change the statistical insignificance of the results.

6.3.2 Robustness checks

The first robustness check focuses on the definition of the dependent variable. In contrast to the main estimation, which uses the *level* of the share of female spending as the outcome, I rerun the estimation using the annual *change* in the dependent variable and the female spending per person. It would not be particularly informative to average the annual change in share of female

spending across the four-year terms, hence I use the annual change in female spending only for the estimation focusing on individual years. Table A.9 reports the results of this estimation. Once again, in all the models, I am unable to reject the null hypothesis that the annual change in share of female spending is the same above and below the cut-off. Interestingly, in year 1 the average share of female spending decreases compared to year 4 of the previous council. This result is in line with the output from Table A.8, which indicated that the average share of female spending in the control group is the highest in the last year in office. In years 2, 3, and 4, the share of female spending display an average annual increase.

The second concern stems from the inclusion of the election years in the data set, as the baseline dependent variable is the average share of female *across the four years* of the council's term. Although the newly elected council has limited power to change the budget for their first months in office, such manipulation cannot be completely ruled out. Therefore, I re-run the main specification, excluding the election years from the data set and creating the spending averages only for the first three years in office. The relevant output in Table A.10 is consistent with the baseline results. The point estimate reveals a statistically insignificant effect of -0.1 p.p. and based on the 95% CI, I can reject a decrease or increase in the share of female spending larger than 0.9 p.p. (12% st. dev.) or 0.6 p.p. (8% st. dev.), respectively. The results are robust across the half and double bandwidths and the quadratic polynomial.

Another robustness check relates to the number of mixed contests in municipalities. As described in section 4, the RD sample used for the baseline model include multiple mixed contests for some municipalities. I tackle this issue in the baseline estimation using weights for municipalities with multiple mixed races, weighting by the inverse of the number of mixed contests. To ensure that the results are robust to this approach, I estimate the baseline model from Equation 5.2 using a sample restricted to municipalities with only one mixed contest. Excluding municipalities with multiple contests leaves a sample of 3,678 observations, almost equally split between the 2010 and 2014 election rounds. Table A.11 reports the results of the estimation. With an statistically insignificant point estimate of -0.2 p.p. compared to the control mean of 13%, the results are comparable with those from the baseline estimation. The results suggest that the effect is neither larger than 0.9 p.p. (12% st. dev.) in the positive direction, nor larger than 1.5 p.p. (19% st. dev.) in the negative direction.

The final robustness test estimates the effect in a restricted sample of smaller municipalities. The effect of a single additional female largely depends on the size of the council: one additional woman increases the share of females by 20% in the smallest councils of 5 but only by 1.5% in the largest possible council of 65. The council size depends on the municipality population, and therefore the smaller municipalities with smaller councils might experience much larger effects on female spending when an additional female wins a council seat. Thus, in Table A.12, I present the results of a set of estimations with the maximum size of the council equal to 7, 9, 11, 15 and 21 in Models 1 to 5, respectively. The results, however, remain similar to the main estimation. Only when the maximum council size is restricted to 9 and 11 councillors are the effects somewhat larger (0.56 p.p. and 0.49 p.p., respectively). However, even these point estimates mark a statistically insignificant change in the average share of female spending.

7 Discussion

All of the results from the previous section failed to find a statistically significant effect of an additional female in a municipal council on the composition of spending, namely the average share of female expenditures. These results are comparable with those from existing papers which find no statistically significant effects of female representation on municipal expenditures, such as [Bagues and Campa \[2017\]](#) or [Rigon and Tanzi \[2012\]](#). My outcomes are also consistent with the study by [Kuliomina \[2019\]](#), the only previous paper written in the context of the Czech Republic, which finds no effect of female representation on the *overall size* of the municipal budget. Somewhat strikingly, the findings of this study largely contrast with [Baskaran and Hessami \[2018\]](#), which I closely follow in terms of the method. Using a comparable set-up of competitive, open-list municipal elections with preferential voting, and focusing on mixed-gender contests, [Baskaran and Hessami \[2018\]](#) find a statistically and economically significant positive effect of an additional female on the provision of child care in Bavaria. Although my paper uses the budget composition as an outcome variable, the child care spending is included in the education spending, on which the effect of increased female representation was fairly underwhelming. The following paragraphs present a set of possible explanations to why the effects I find are both statistically and economically insignificant.

Firstly, there is a crucial difference between the Bavarian municipalities from [Baskaran and Hessami \[2018\]](#) and the Czech context that could account for the contrasting results – the existing social welfare and policy environment in the two countries. As suggested in the introduction, Czech Republic, as a formerly socialist country differs from many Western democracies in terms of the generosity of various welfare benefits. [Baskaran and Hessami \[2018\]](#) find the effect in the child care policy area, which in the Czech Republic falls under the municipal education budget. In the Czech Republic, child care spots are essentially free, readily available for children from the age of 2 and guaranteed from the age of 4, and parental leave regularly lasts up to 3 years. This could mean that the need for additional childcare spending is simply not as urgent in the Czech Republic as it might be in Bavaria. This argument is further supported by the survey results where women did not seem more concerned with education (which includes child care) compared to men. The same reasoning can extend to many other areas of policy which are normally considered a primarily female issue, both in the existing literature and in my study.⁴⁶ If policy in the ‘female-oriented’ areas is already well-developed and reasonably functional, it might be much harder to push for more spending in the area, rather than increasing expenditures on some other, likely more urgent issues.

There are three additional caveats which are specific to my study, which could also explain the lack of statistically significant effect in my estimation. The first is the use of the survey results for classification of the spending items. The survey questions are rather general, giving

⁴⁶For instance, public transportation networks both within and between municipalities are rather well-developed and affordable in the Czech Republic compared to many Western countries, which makes further spending in the area relatively less important and thus less likely. Comparable story is applicable for the area of healthcare: in many Western countries, women health and access to healthcare tends to be worse than that of men, making healthcare more of a ‘female’ issue [[European Commission, 2009](#)]. In the Czech Republic and some other post-socialist countries on the other hand, the situation is reversed. Women have on average better access to healthcare and as my survey results indicated, it is men that are more concerned with healthcare.

the respondents the space for interpreting on which *level* of governance are they satisfied with a particular area of public life. A respondent who was asked about their satisfaction with infrastructure can translate the question to the local terms (e.g. satisfaction with local public transport) or national-level terms (e.g. satisfaction with the state of highways). Given that the national level issues are prevalent on the media, it is likely that the respondents were inclined to answer about their satisfaction with the areas at the *national* level, thinking about issues which lie outside the scope of the municipality's power. In an ideal study, I would use survey data that *explicitly* ask the respondents about their policy preferences on the *local* level; however, in an absence of such data, the survey I used is the best approximation of the gender-specific preferences for policy change.

The second caveat stems both from the structure of the survey and from the existing generous welfare state. A part of the Czech population believes that the welfare state is too generous and would like to see cuts in some of the benefits.⁴⁷ If the survey respondents have such preferences, they will likely express dissatisfaction with a particular area, but not because the expenditures there are *too low* but on the contrary because they are *too high*. These preferences could in some cases even cross the gender lines: in a scenario with generous child care policy, some free-market oriented men could have a preference for cutting some of the child benefits. Hence, an ideal survey would not only be tailored to the local-level issues, but also inquire about the nature of the preferences for spending and policy change.

Lastly, the Czech municipalities are sometimes engaged in inter-municipality cooperation. As the municipalities are on average quite small, many do not have the capacity and funds to individually run organizations such as schools, or healthcare and elderly care centers. Municipalities, usually in geographic proximity to one another, can enter an agreement to create a 'municipality association' which has a form of a statutory-corporation that can then collaborate on such common issues.⁴⁸ Each of these associations have their own budget.⁴⁹ Therefore, in a hypothetical case where a municipality marginally elects a female, starts to care more about local public transportation and manages to arrange an expansion of the transportation system within its municipality association, the additional expenditures would be likely picked up by not only the instigating municipality, but the other members of the association as well. If any of these other members marginally elected men, my estimation would associate the extra spending with the elected male although they technically had little to no influence on the budget change.

7.1 Further research

The limitations, outlined above, suggest possible extensions to my study that could expose the more nuanced effects of female representations, and the possible obstacles which females in office face. My first proposition is to assess the behaviour of women in office, similar to the work

⁴⁷A recurring theme is for instance the unsustainability of the current healthcare pension systems (with ageing population, the costs to the state will become unbearable) or overly generous social welfare payments. For more details (in Czech), see Holanová [2014] or Havligerová [2018].

⁴⁸The most common areas of collaboration are education, healthcare, environmental protection, infrastructure, safety and security, tourism and waste management.

⁴⁹There is no publicly available centralized register of the municipality associations. Although the budgets are published, it would be highly challenging to investigate the municipalities which belong to each of the associations.

done by [Baskaran and Hessami \[2018\]](#). As stated in the introduction, one of the hindrances to effective female representation could be the minority status of the women in most of the municipal councils. Therefore, it would be worthwhile to analyze the more subtle effects of women in office by looking at outcomes such as presence at council meetings, speaking time at meetings, share of policy proposals by females and female voting patterns. It could be the case that as a minority, the women councillors are less likely to speak during the council meetings, or to make policy propositions. Conversely, they might be making a large number of propositions but not get backed by the remaining council members. Additionally, it would be valuable to investigate whether these patterns change with the share of females on the council, or with the rank in the election when winning a mandate. It is possible that the council has an informal hierarchy given by the election success of the councillors. Since women on average tend to be listed lower on the ballot and thus likely have a lower rank during the mandate allocation process, their power to affect policy could be lower.

The second extension would include analysis of the impact of the mayor's gender. Although the mayor does not possess extensive executive powers, their influence can still be high in the council meetings. The mayor is usually the speaker in the council meetings and can thus affect the direction of the policy and budget related discussions. The mayor also heads the municipality committee, the executive body of the municipality. Previous literature suggest that female representation on a local council can be influential, conditional on having a female mayor [[Holman, 2014](#)]. For the female councillors in a minority on the council, a woman in a leadership position could be encouraging to make more female-oriented propositions.⁵⁰ On the other hand, such extension might be problematic as the mayors in the Czech Republic are elected by the council members and not by the public. Councils with more females might therefore be more likely to elect a female mayor.⁵¹

My final suggestion is to carry out a survey which would follow more closely the policy preferences on a *local* scale, similar to the approach of for instance [Chattopadhyay and Duflo \[2004\]](#). Using such surveys for classifying the expenditures into female, male and neutral would lead to estimating the effect of local female representation with more precision. Additionally, as I suggested in the previous section, such survey should more closely follow the particular preferences of the respondents.

8 Conclusion

This study investigates the impact of increased female representation in municipal councils on the public finance composition in the context of Czech municipalities. The spending on the areas that are 'female-oriented' is of particular interest. Based on survey results, women in the Czech Republic seem to have distinct policy preferences from men. However, using RD approach, my analysis suggests that the presence of an additional woman on the council fails to affect the

⁵⁰It is crucial to note though, that in such cases, the female mayor could be the *sole* or *dominant* source of the impact on policy. However, as mayor elections are not public in the Czech Republic, it would be difficult to estimate the effect of only the gender of the mayor on policy-making.

⁵¹Therefore, such study could rather be carried out in a similar country where both the council and the mayor are elected, such as the Slovak Republic.

composition of the budget in a statistically significant way. The results are consistent across a variety of specifications and robustness tests.

My analysis is subject to a number of specific conditions. The Czech Republic, as a post-socialist country, retains a generous system of welfare policies (such as affordable pre-school or long maternity leave). This could affect the ability to pass further legislation in the area, as it is likely not regarded as an urgent matter. Additionally, the majority of Czech municipalities and their councils are rather small and powers are relatively limited. The results will likely be substantially different in countries where local governance includes a larger scope of powers and covers a geographically and demographically more sizeable area.⁵²

The fact that my study cannot find a statistically or economically significant impact of female representation does not mean that the lack thereof is not an issue. Firstly, given the fact that women account for a half of the population and are on average better educated [[Czech Statistical Office, 2017](#)], there is little doubt enough women are qualified and competent to serve in a public office. Expanding the candidate pool by including a larger portion of women could thus yield more efficient selection process and improve the quality of the candidates and councillors. The simplest and fastest way to do so would be to follow many other EU countries and implement quota for gender balance on the ballot, including rules for the initial ranking of the candidates.⁵³

Finally, this study provides insight only about one of the multiple mechanisms through which female representation might have an influence on the general public. I cannot capture more subtle effects such as creating role models for future generations, encouraging female participation in different areas of public life, or preventing policy-making which would harm, disadvantage, or burden women.⁵⁴ Further studies should investigate the channels in which female representation might affect society. Additionally, the impact of female representation is likely influenced by the existing gender stereotypes, electorate systems, legal environment and even cultural and social norms. More research should thus be devoted to the consequences of female representation across various geographic areas, levels of governance and branches of power.

⁵²Sweden, for example, has a similar population to the Czech Republic, yet it has only 290 municipalities with much larger set of responsibilities and powers.

⁵³For instance, in Spain, the quota policy mandates not only 40% female candidates, but it also requires the parties to have at least 2 women in each rank of five. That means that among the candidates ranked 1-5, two have to be women; among those ranked 6-10, two have to be women, and so on.

⁵⁴A woman councillor, even in a minority, can point out details of proposed policies problematic for women, which the male colleagues might not realize.

References

- Abrams, B. A. and Settle, R. F. (1999). Women’s suffrage and the growth of the welfare state. *Public Choice*, 100(3-4):289–300.
- Aidt, T. S. and Dallal, B. (2008). Female voting power: The contribution of women’s suffrage to the growth of social spending in Western Europe (1869-1960). *Public Choice*, 134(3-4):391–417.
- Andersen, S., Bulte, E., Gneezy, U., and List, J. A. (2008). Do women supply more public goods than men? Preliminary experimental evidence from matrilineal and patriarchal societies. *American Economic Review*, 98(2):376–381.
- Bagues, M. and Campa, P. (2017). Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design. Unpublished manuscript.
- Balík, S. (2009). *Local politics: Municipalities, actors and goals of local politics [Komunální politika: Obce, aktéři a cíle místní politiky]*. Grada Publishing as.
- Baskaran, T., Bhalotra, S., Min, B., and Uppal, Y. (2018). Women legislators and economic performance. *IZA Discussion Papers*.
- Baskaran, T. and Hessami, Z. (2018). Does the entry of a woman into political office affect policy choices? Unpublished manuscript.
- Bernard, J. (2012). The personal characteristics of candidates in municipal elections and their effect on election results [Individuální charakteristiky kandidátů ve volbách do zastupitelstev obcí a jejich vliv na volební výsledky]. *Sociologický Časopis - Czech Sociological Review*, 48(4):613 – 640.
- Besley, T. and Coate, S. (1997). An economic model of representative democracy. *Quarterly Journal of Economics*, 112(1):85–114.
- Born, A., Ranehill, E., and Sandberg, A. (2018). A man’s world? The impact of a male dominated environment on female leadership. Unpublished manuscript.
- Braendle, T. and Colombier, C. (2016). What drives public health care expenditure growth? Evidence from Swiss cantons, 1970–2012. *Health Policy*, 120(9):1051–1060.
- Brollo, F. and Troiano, U. (2016). What happens when a woman wins an election? Evidence from close races in Brazil. *Journal of Development Economics*, 122:28–45.
- Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2018). Coverage error optimal confidence intervals for local polynomial regression. Unpublished manuscript.
- Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). Rdrobust: Software for regression-discontinuity designs. *Stata Journal*, 17(2):372–404.
- Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4):909–946.

- Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014b). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.
- Calonico, S., Cattaneo, M. D., and Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769.
- Carozzi, F. and Gago, A. (2017). Female mayors and gender policies in a developed country. Unpublished manuscript.
- Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019a). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.
- Cattaneo, M. D., Jansson, M., and Ma, X. (2019b). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 0(0):1–11.
- Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2019c). Power calculations for regression-discontinuity designs. *The Stata Journal*, 19(1):210–245.
- Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, 72(5):1409–1443.
- Clayton, A. and Zetterberg, P. (2018). Quota shocks: Electoral gender quotas and government spending priorities worldwide. *Journal of Politics*, 80(3):916–932.
- Clots-Figueras, I. (2011). Women in politics: Evidence from the Indian states. *Journal of Public Economics*, 95(7-8):664–690.
- Clots-Figueras, I. (2012). Are female leaders good for education? Evidence from India. *American Economic Journal: Applied Economics*, 4(1):212–244.
- Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.
- Czech Statistical Office (2017). Women and men in the data - 2017 [Ženy a muži v datech - 2017]. Online statistics. <https://www.czso.cz/csu/czso/zeny-a-muzi-v-datech-2017>.
- Dollar, D., Raymond, F., and Gatti, R. (2001). Are women really the “fairer” sex? Corruption and women in government. *Journal of Economic Behavior and Organization*, 46(4):423–429.
- Downs, A. (1957). An economic theory of political action in a democracy. *The Journal of Political Economy*, 65(2):135–150.
- European Commission (2009). Access to healthcare and long-term care: Equal for women and men? Report.
- European Institute for Gender Equality (2020). Gender statistics database, National parliaments: presidents and members. Online statistics. <https://eige.europa.eu/gender-statistics/dgs/indicator/>.
- Ferreira, F. and Gyourko, J. (2014). Does gender matter for political leadership? The case of U.S. mayors. *Journal of Public Economics*, 112:24–39.

- Funk, P. and Gathmann, C. (2015). Gender gaps in policy making: Evidence from direct democracy in Switzerland. *Economic Policy*, 30(81):141–181.
- Gagliarducci, S. and Paserman, M. D. (2012). Gender interactions within hierarchies: Evidence from the political arena. *Review of Economic Studies*, 79:1021–1052.
- Havlígerová, J. (2018). Generous year of Andrej Babiš: The prime minister and his cabinet are building the social state [Štědrý rok Andreje Babiše: premiér a jeho kabinet budují sociální stát]. *E15.cz*. Newspaper article.
- Holanová, T. (2014). The Czech Republic is burdened by the social state, agreed the economists [Česko tíží sociální stát, shodli se ekonomové]. *Aktuálně.cz*. Newspaper article.
- Holman, M. R. (2014). Sex and the city: Female leaders and spending on social welfare programs in U.S. municipalities. *Journal of Urban Affairs*, 36(4):701–715.
- Jurajda, S. and München, D. (2015). Candidate ballot information and election outcomes: the Czech case. *Post-Soviet Affairs*, 31(5):448–469.
- Kuliomina, J. (2019). *Essays on the Study of Local Politics*. PhD thesis. CERGE-EI.
- Lippmann, Q. (2019). Gender and lawmaking in times of quotas. Unpublished manuscript.
- Mansbridge, J. (1999). Should blacks represent blacks and women represent women? A contingent "yes". *The Journal of Politics*, 61(3):628–657.
- Osborne, M. J. and Slivinski, A. (1996). A model of political competition with citizen-candidates. *Quarterly Journal of Economics*, 111:65–96.
- Parlament České Republiky (2000). Act No. 128/2000 Coll., on municipalities (on Municipal Order) [Zákon č. 128/2000 Sb., Zákon o obcích (o obecním zřízení)]. Legal document. <https://www.zakonyprolidi.cz/cs/2000-128>.
- Parlament České Republiky (2001). Act no. 491/2001 Coll., on elections to municipality councils and on amendments to selected acts [Zákon č. 491/2001 Sb., o volbách do zastupitelstev obcí a o změně některých zákonů]. Legal document. <https://www.zakonyprolidi.cz/cs/2001-491>.
- Ranehill, E. and Weber, R. A. (2017). Do gender preference gaps impact policy outcomes? Unpublished manuscript.
- Rigon, M. and Tanzi, G. M. (2012). Does gender matter for public spending? Empirical evidence from Italian municipalities. *Bank of Italy Temi di Discussione*, 35(442):1–28.
- Svaleryd, H. (2009). Women's representation and public spending. *European Journal of Political Economy*, 25(2):186–198.
- Swamy, A., Knack, S., and Lee, Y. (2001). Gender and corruption. *Journal of Development Economics*, 64:25–55.
- Šprincová, V. and Mottlová, M. (2015). Quota and other policies for higher female representations in politics [Kvóty a další opatření pro vyšší zastoupení žen v politice]. *Fórum 50%*. Report.

A Additional output

Table A.1 Summary statistics: RD sample by number of mixed contests

Contests	Council size	Parties	% Women	Δ % Women	Population	# Elections
1	9.35	3.10	28.08	12.07	1166.30	3679
2	11.49	4.38	28.60	10.14	3262.21	1870
3	14.49	6.07	29.29	8.05	5120.48	668
4	17.40	7.54	28.05	6.65	9680.27	210
5	20.24	8.83	29.35	5.51	12941.76	59
6	19.67	10.00	25.95	5.33	9239.17	12
7	21.00	11.33	25.71	5.04	17516.33	3
8	21.00	10.00	28.57	4.76	7591.00	1
13	63.00	20.00	22.22	1.59	1249026.00	1
14	45.00	13.00	17.78	2.22	100362.00	1

Note: This table includes 6,640 local elections with at least one gender-mixed contest for the last seats within a slate, out of the 12,483 elections in the full sample. The first column reports the number of gender-mixed contests, the second column the average council size, the third shows the average number of slates in the election, the fourth column shows the average share of women in the council, the fifth column shows the average change in the share of women if one additional women won a seat, that is $1/\text{council size}$, the sixth column report the average population of the municipality and finally, the last column shows how many elections occurred with the particular number of gender-mixed contests. Note that some large municipalities have multiple election districts, hence, it is possible to have more mixed contests than the number of parties.

Table A.2 OLS: Female spending pooled across election period

	(1)	(2)
Share of female councillors	-0.0082 (0.007)	0.0002 (0.008)
Election year=2014	0.007*** (0.002)	
Lagged share of female spending		0.416*** (0.018)
Constant	0.128*** (0.002)	0.081*** (0.003)
Observations	6503	3221

Note: The dependent variable in both columns is the average share of female spending, averaged over the term period of the legislature. Standard errors, clustered at the municipality level, are reported in parentheses. Model (1) reports the estimation using the full RD sample, controlling for the election year. Model (2) restricts the sample to the 2015-2018 legislatures, controlling for a lagged share of female spending in the previous legislature.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3 Validity: Vote margin by the gender of the winner

	Mean	Median	Min	Max	Obs.
Female winner	2.437	1.301	0.005	58.889	5123
Male winner	-2.684	-1.380	-67.797	-0.003	5562

Note: This table presents the summary statistics of vote margin, grouped by the gender of the winner. It includes 10,685 observations from two election rounds. Since the vote margin is defined as the difference in percentage of votes obtained between the female and the male candidate, a positive margin indicates a female winner, while a male winner gets a negative margin. See more details about margin calculation in Appendix B.

Table A.4 Validity: Covariate balance

	Estimate	Std. err.	p-value	h_{MSE}	CI	Obs.
<i>Panel A: Candidate-level covariates</i>						
Female age	-0.099	0.482	0.840	3.08	[-1.16; 0.94]	10686
Male age	-0.320	0.480	0.540	3.77	[-1.36; 0.71]	10686
Female educated	-0.003	0.022	0.968	3.04	[-0.05; 0.05]	10686
Male educated	-0.015	0.021	0.512	3.10	[-0.06; 0.03]	10686
<i>Panel B: Municipality-level covariates</i>						
Log population	0.055	0.080	0.518	2.98	[-0.12; 0.24]	10686
Council size	0.115	0.394	0.791	3.63	[-0.78; 1.02]	10686
# slates	0.112	0.167	0.543	3.38	[-0.26; 0.50]	10686
# mixed contests	0.012	0.046	0.879	3.37	[-0.09; 0.11]	10686
<i>Panel C: 2011 Census covariates</i>						
% aged 0-14	-0.003	0.002	0.080	2.02	[-0.01; 0.00]	5299
% aged 65+	0.002	0.002	0.321	2.12	[-0.00; 0.01]	5299
% foreign citizens	-0.000	0.004	0.829	3.23	[-0.01; 0.01]	5299
% religious	0.003	0.008	0.580	2.80	[-0.01; 0.02]	5299
% unemployed	-0.001	0.002	0.703	2.97	[-0.01; 0.00]	5299
% female employed	-0.002	0.004	0.556	2.72	[-0.01; 0.01]	5299

Note: This table displays the results of the predetermined covariate balance test. The independent variable for each row is the vote margin. The dependent variables are listed in the left-most column. The subsequent columns display the point estimate, the standard error, the p-value of the point estimate, the optimal bandwidth for the estimation, the robust-adjusted confidence interval (95%) and the number of observations. Panels A and B display the candidate and municipality covariates across election rounds, respectively. Panel C uses Census data from 2011, hence it is only estimated using the 2014 election round.

Table A.5 Validity: Alternative cutoffs

Cutoff	Estimator	Std. err.	p-value	h_{MSE}	CI	Obs.
-0.5	-0.014	0.009	0.077	0.392	[-0.044; 0.002]	5562
-0.4	0.017	0.009	0.069	0.376	[-0.002; 0.044]	5562
-0.3	-0.002	0.009	0.405	0.406	[-0.034; 0.014]	5562
-0.2	0.004	0.009	0.636	0.485	[-0.018; 0.029]	5562
-0.1	-0.008	0.010	0.361	0.492	[-0.045; 0.016]	5562
0.0	-0.001	0.004	0.601	2.337	[-0.010; 0.006]	10685
0.1	-0.000	0.009	0.969	0.451	[-0.027; 0.028]	5123
0.2	0.003	0.008	0.811	0.445	[-0.021; 0.026]	5123
0.3	-0.020	0.009	0.076	0.414	[-0.048; 0.002]	5123
0.4	0.005	0.009	0.158	0.345	[-0.006; 0.035]	5123
0.5	-0.004	0.009	0.725	0.356	[-0.028; 0.019]	5123

Note: This table displays the results of the main estimation using alternative cut-offs ranging from -0.5 to 0.5, listed in the left-most column. The subsequent columns report the estimate, the standard error, the p-value of the estimate, the optimal bandwidth, the robust-adjusted confidence interval (95%) and the number of observations. The estimations with a negative (positive) alternative cut-off use only the observations on the negative (positive) side of the original zero cut-off.

Table A.6 Validity: Donut hole approach

Radius	Estimator	Std. err.	p-value	h_{MSE}	CI	Obs.	Excl. obs.
0.00	-0.0013	0.004	0.601	2.337	[-0.010; 0.006]	10685	[0; 0]
0.01	-0.0013	0.004	0.604	2.346	[-0.010; 0.006]	10669	[11; 5]
0.02	-0.0008	0.004	0.695	2.366	[-0.009; 0.006]	10630	[28; 27]
0.03	-0.0007	0.004	0.717	2.384	[-0.009; 0.006]	10587	[50; 48]
0.04	-0.0002	0.004	0.819	2.444	[-0.009; 0.007]	10536	[77; 72]
0.05	-0.0001	0.004	0.834	2.418	[-0.009; 0.007]	10488	[98; 99]
0.06	0.0003	0.004	0.925	2.493	[-0.008; 0.008]	10427	[126; 132]
0.07	0.0008	0.004	0.965	2.556	[-0.008; 0.008]	10362	[155; 168]
0.08	0.0010	0.004	0.926	2.568	[-0.008; 0.009]	10312	[180; 193]
0.09	0.0003	0.004	0.931	2.587	[-0.009; 0.008]	10247	[219; 219]

Note: This table displays the results of the ‘Donut hole’ test on the main estimation, excluding a selection of observation close to the cut-off. The radius of excluded observations is reported in the left-most column. The subsequent columns report the estimate, the standard error, the p-value of the estimate, the optimal bandwidth, the robust-adjusted confidence interval (95%), the number of observations used and the number of observations excluded on the left and right hand side of the zero cut-off.

Table A.7 Validity: Alternative bandwidths

Alternative h	Estimate	Std. err.	p-value	CI	Obs
1.0	-0.0072	0.005	0.122	[-0.023; 0.003]	10685
1.2	-0.0059	0.004	0.083	[-0.022; 0.001]	10685
1.4	-0.0044	0.004	0.064	[-0.022; 0.001]	10685
1.6	-0.0033	0.004	0.070	[-0.020; 0.001]	10685
1.8	-0.0026	0.004	0.105	[-0.019; 0.002]	10685
2.0	-0.0020	0.004	0.138	[-0.017; 0.002]	10685
2.2	-0.0016	0.004	0.179	[-0.016; 0.003]	10685
2.4	-0.0012	0.004	0.215	[-0.015; 0.003]	10685
2.6	-0.0008	0.003	0.254	[-0.014; 0.004]	10685
2.8	-0.0004	0.003	0.288	[-0.014; 0.004]	10685
3.0	-0.0001	0.003	0.329	[-0.013; 0.004]	10685
3.2	0.0001	0.003	0.385	[-0.012; 0.005]	10685
3.4	0.0001	0.003	0.458	[-0.012; 0.005]	10685
3.6	0.0002	0.003	0.526	[-0.011; 0.006]	10685
3.8	0.0003	0.003	0.569	[-0.010; 0.006]	10685
4.0	0.0005	0.003	0.572	[-0.010; 0.006]	10685
4.2	0.0006	0.003	0.581	[-0.010; 0.006]	10685
4.4	0.0008	0.003	0.569	[-0.010; 0.006]	10685
4.6	0.0010	0.003	0.562	[-0.010; 0.005]	10685
4.8	0.0012	0.003	0.576	[-0.010; 0.005]	10685
5.0	0.0014	0.003	0.602	[-0.009; 0.005]	10685

Note: This table displays the result of the alternative bandwidth sensitivity test, with the width of the alternative bandwidths reported in the left-most column. The subsequent columns display the estimate with its standard error and p-value, the robust-adjusted confidence interval (95%) and the number of observations used.

Table A.8 Sharp RD: Female victory and the yearly share of female spending

	(1)	(2)	(3)	(4)	(5)
	Year 1	Year 2	Year 3	Year 4	All years
Female victory	0.0002 (0.0049)	-0.0004 (0.0049)	-0.0016 (0.0051)	-0.0005 (0.0049)	-0.0013 (0.0036)
Polynomial	Linear	Linear	Linear	Linear	Linear
Bandwidth type	Optimal	Optimal	Optimal	Optimal	Optimal
Bandwidth size	2.748	2.365	2.537	2.744	2.338
Mean control group	0.115	0.123	0.126	0.132	0.124
Robust CI	[-.011 ; .01]	[-.012 ; .009]	[-.013 ; .008]	[-.012 ; .009]	[-.01 ; .006]
Observations	10684	10685	10685	10685	42739

Note: The dependent variable in Models (1)-(4) is the share of female spending in the individual years (1-4) of a legislative term. Model (5) uses all observations from years 1-4. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of gender-mixed contests})$ to correct for municipalities with multiple mixed contests. All models use MSE optimal bandwidth (based on Calonico et al. [2014b]) and a linear fit. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and the number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.9 Robustness: Estimation with annual change in share of female spending

	(1)	(2)	(3)	(4)	(5)
	Year 1	Year 2	Year 3	Year 4	All years
Female victory	-0.0016 (0.0058)	0.0004 (0.0056)	-0.0015 (0.0053)	0.0011 (0.0053)	0.0003 (0.0015)
Polynomial	Linear	Linear	Linear	Linear	Linear
Bandwidth type	Optimal	Optimal	Optimal	Optimal	Optimal
Bandwidth size	2.469	2.588	2.709	3.052	3.033
Mean control group	-0.0160	0.0089	0.0028	0.0053	0.0005
Robust CI	[-.015; .01]	[-.012; .012]	[-.013; .01]	[-.01; .013]	[-.003; .004]
Observations	10684	10685	10685	10685	10685

Note: The dependent variable in Models (1)-(4) is the annual change in the share of female spending between individual years (1-4) of a legislative term (note that Year 1 translates to the change from the year 4 of the previous council to year 1 of the current council). Model (5) uses all observations. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of mixed-gender contests})$ to correct for municipalities with multiple mixed contests. All models use MSE optimal bandwidth (based on [Calonico et al. \[2014b\]](#)) and a linear fit. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and the number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.10 Robustness: Estimation excluding election years

	(1)	(2)	(3)	(4)
Female victory	-0.001 (0.003)	-0.004 (0.005)	0.001 (0.003)	-0.002 (0.004)
Polynomial	Linear	Linear	Linear	Quadratic
Bandwidth type	Optimal	Half	Double	Optimal
Bandwidth size	2.320	1.160	4.640	5.057
Mean control group	0.121	0.118	0.123	0.124
Robust CI	[-.009; .006]	[-.018; .007]	[-.01; .006]	[-.01; .006]
Observations	10685	10685	10685	10685

Note: The dependent variable is the average share of female spending across the first 3 years of the legislative term. The last year of a legislative term, which is the election year, is excluded from the average. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of mixed-gender contests})$ to correct for municipalities with multiple mixed contests. Model (1) and (4) use optimal bandwidth (based on [Calonico et al. \[2014b\]](#)), while Models (2) and (3) use half and double of the optimal bandwidth, respectively. Models (1)-(3) use linear polynomial while the polynomial in Model (4) is quadratic. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.11 Robustness: Estimation in municipalities with only one mixed contest

	(1)	(2)	(3)	(4)
Female victory	-0.002 (0.006)	-0.005 (0.007)	0.001 (0.005)	-0.003 (0.006)
Polynomial	Linear	Linear	Linear	Quadratic
Bandwidth type	Optimal	Half	Double	Optimal
Bandwidth size	3.001	1.500	6.001	6.082
Mean control group	0.130	0.131	0.132	0.132
Robust CI	[-.015; .009]	[-.035; .007]	[-.015; .009]	[-.016; .009]
Observations	3678	3678	3678	3678

Note: The dependent variable is the average share of female spending across the four years of a council's legislative term. The sample is restricted to the municipalities which had exactly one gender-mixed contest. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of mixed-gender contests})$ to correct for municipalities with multiple mixed contests. Model (1) and (4) use optimal bandwidth (based on [Calonico et al. \[2014b\]](#)), while Models (2) and (3) use half and double of the optimal bandwidth, respectively. Models (1)-(3) use linear polynomial while the polynomial in Model (4) is quadratic. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and number of observations used are reported for each estimation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

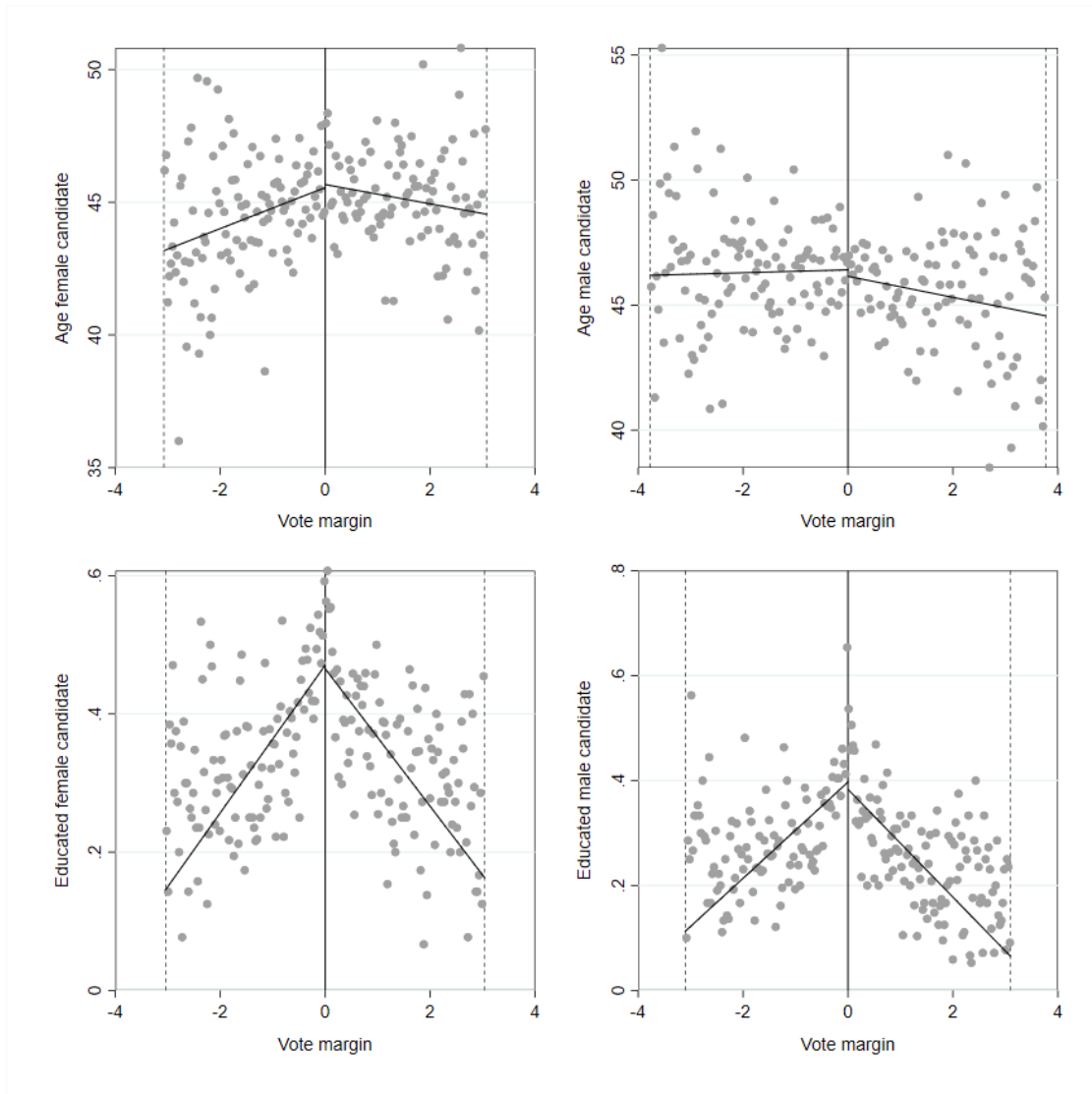
Table A.12 Robustness: Estimation restricted to small municipalities

	(1)	(2)	(3)	(4)	(5)
Female victory	-0.0001 (0.0075)	0.0056 (0.0056)	0.0049 (0.0050)	0.0012 (0.0039)	-0.0002 (0.0038)
Max council size	7	9	11	15	21
Polynomial	Linear	Linear	Linear	Linear	Linear
Bandwidth type	Optimal	Optimal	Optimal	Optimal	Optimal
Bandwidth size	4.375	4.008	4.014	3.226	2.633
Mean control group	0.125	0.125	0.125	0.125	0.124
Robust CI	[-.015; .018]	[-.006; .018]	[-.006; .016]	[-.008; .009]	[-.009; .007]
Observations	3039	5458	6310	9202	10133

Note: The dependent variable is the average share of female spending across the four years of a council's legislative term. In each model, the sample is restricted to a particular maximum council size, reported in the third row. Each cell in the top row reports a conventional RD coefficient, with standard errors in parentheses, clustered by municipality. A triangular kernel is used to weigh the observations based on their distance to the cut-off. Each observation is further weighed by $1/(\text{number of mixed-gender contests})$ to correct for municipalities with multiple mixed contests. Model (1) and (4) use optimal bandwidth (based on [Calonico et al. \[2014b\]](#)), while Models (2) and (3) use half and double of the optimal bandwidth, respectively. Models (1)-(3) use linear polynomial while the polynomial in Model (4) is quadratic. The bandwidth size, mean of the control group, the bias-corrected robust CI (95%) and number of observations used are reported for each estimation.

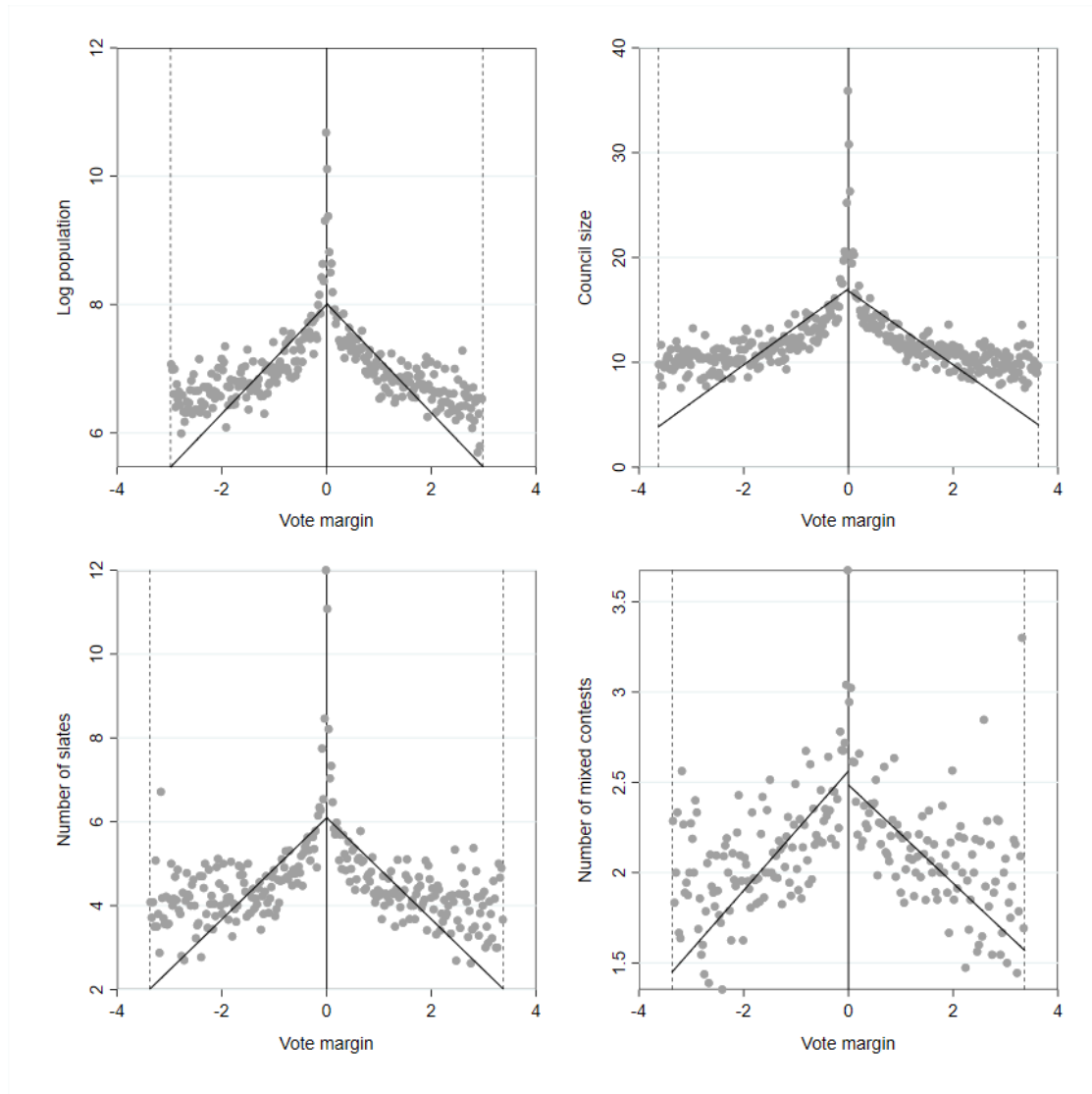
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A.1 Validity: Candidate-level predetermined covariates



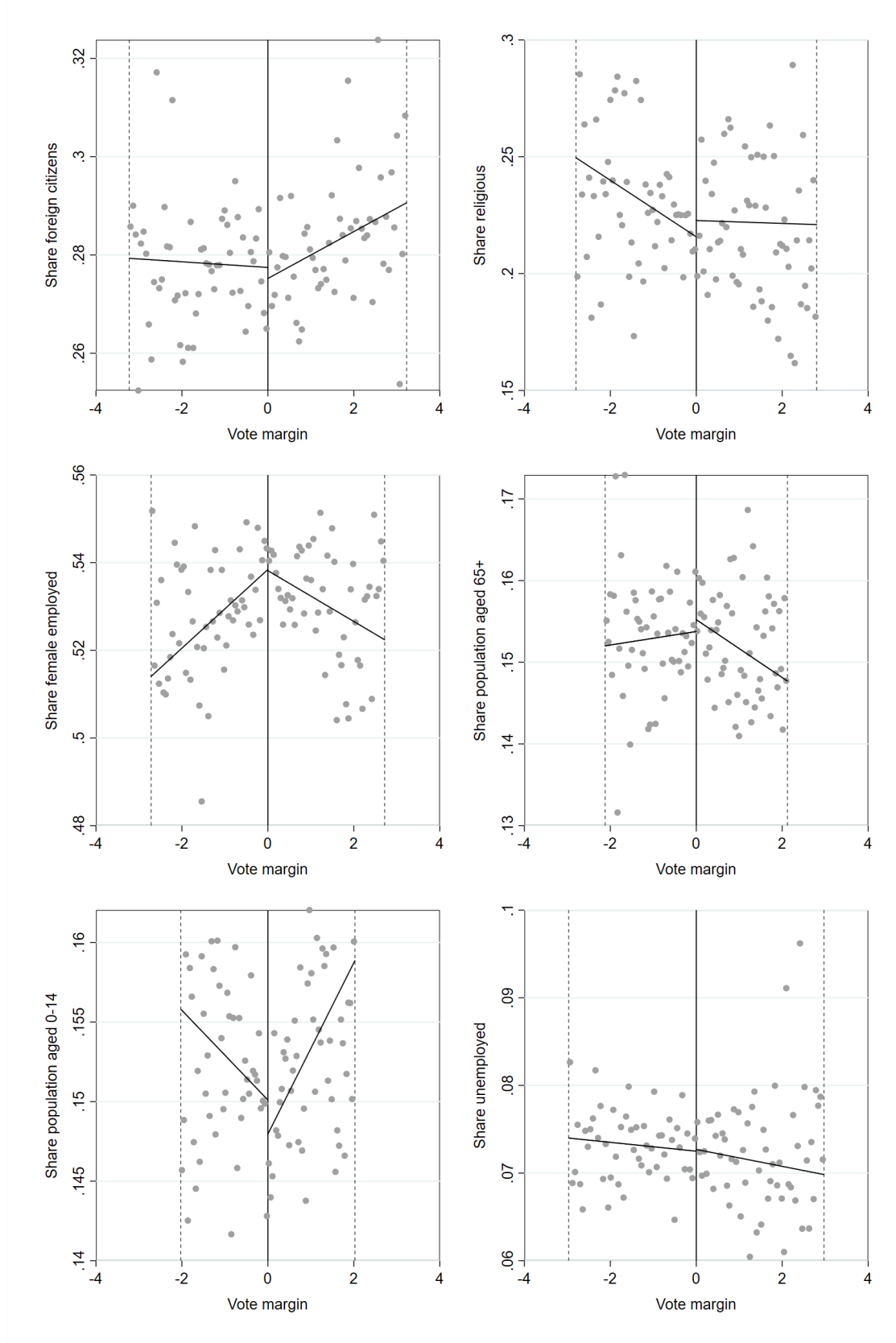
Note: These graphs visually complement the predetermined covariate tests for the candidate-level covariates. The vertical solid line indicates the zero cut-off, while the dashed lines mark the optimal bandwidth.

Figure A.2 Validity: Municipality-level predetermined covariates



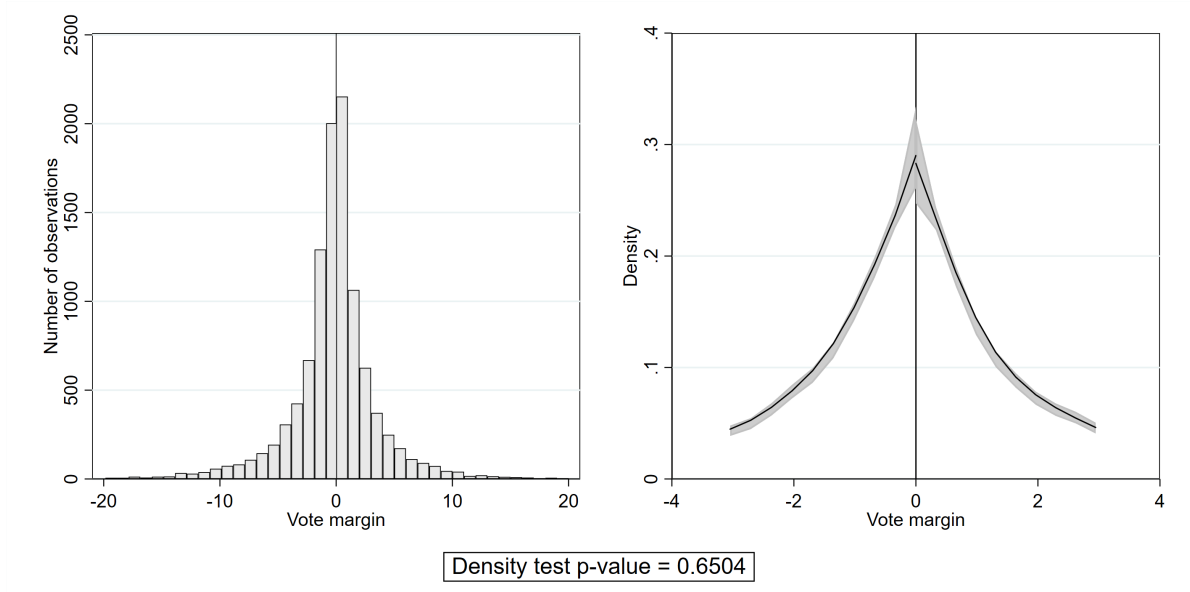
Note: These graphs visually complement the predetermined covariate tests for the municipality-level covariates. The vertical solid line indicates the zero cut-off, while the dashed lines mark the optimal bandwidth.

Figure A.3 Validity: 2011 Census predetermined covariates



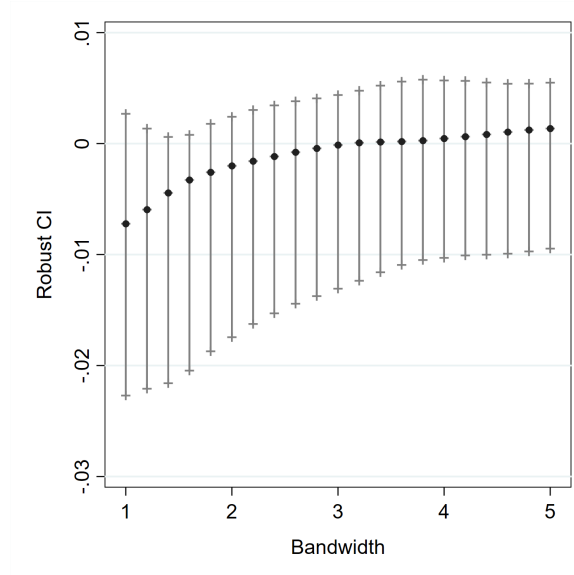
Note: These graphs visually complement the predetermined covariate tests for the municipality-level covariates from the 2011 Census. The vertical solid line indicates the zero cut-off, while the dashed lines mark the optimal bandwidth.

Figure A.4 Validity: Density plots



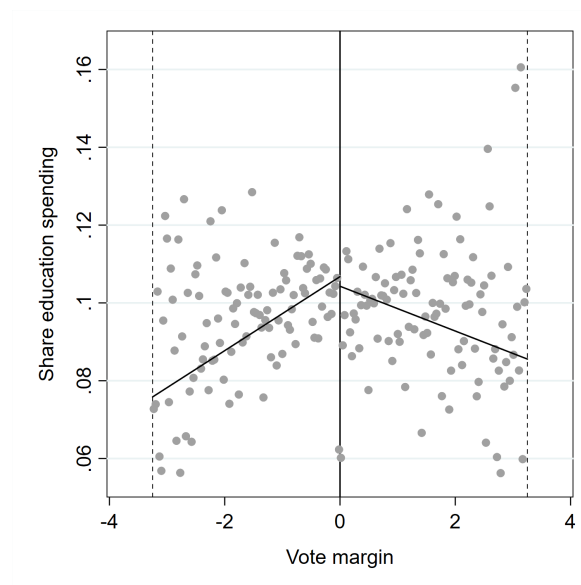
Note: The graphs report the results of the density tests. The left figure shows a histogram of the observations on both sides of the cut-off, while the right figure utilizes the manipulation test using local-polynomial density estimation as described in Cattaneo et al. [2019b]. The p-value of the manipulation test is reported below the table.

Figure A.5 Validity: Alternative bandwidth test



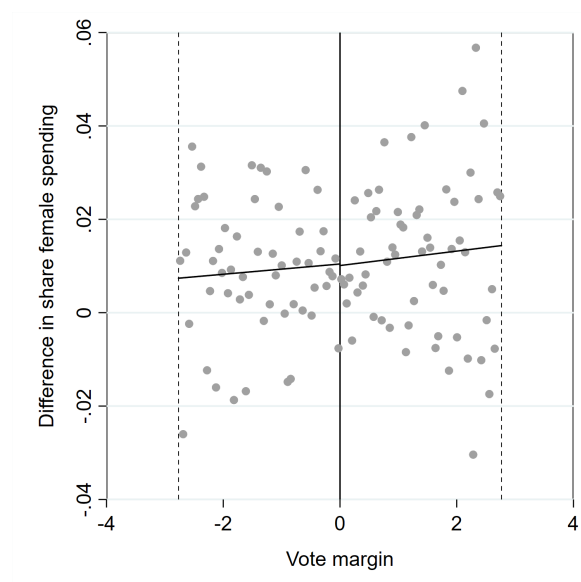
Note: The figure shows the point estimates and CI for a range of bandwidths. The dots represent the conventional point estimates from the main specification with the average share of female spending as a dependent variable. The lines mark the bias-adjusted robust confidence intervals (95%).

Figure A.6 Extensions: Local RD plot of the average share of education spending



Note: Local RD plot of the vote margin against the average share of education spending. The solid vertical line represents the zero vote margin and the dashed vertical lines depict the optimal bandwidth.

Figure A.7 Extensions: Local RD plot of the difference in average share of female spending



Note: Local RD plot of the vote margin against the difference in average share of female spending between the 2010 and 2014 legislatures. The solid vertical line represents the zero vote margin and the dashed vertical lines depict the optimal bandwidth.

B Mandates allocation and margin calculation

B.1 Allocating mandates

The first step in the mandate allocation process is to finalize the rank of the candidates on a slate. Using the total number of votes an election party obtained, an average number of votes per candidate is calculated. Each candidate who received 10% more individual votes than the slate average moves up in the final ranking. Table B.1 shows such a rearranging on a slate from a sample municipality. The total number of votes for the slate is 493, which translates into an average of 98.6 votes. Therefore, the threshold for switching ranks is 108.5 votes,⁵⁵ and is surpassed by candidates B and D. Since Candidate D has more votes than Candidate B, he takes the first position in the final ranking. Notice that although Candidate A received less votes than Candidate C, he still retained his position before Candidate C, as the latter did not manage to reach the threshold of 108.5 votes. In such case, the candidates remained ordered in their initial rank.

Table B.1 Slate order before and after election

Candidate	Initial Rank	Total Votes	Final rank
Candidate A	1	78	3
Candidate B	2	136	2
Candidate C	3	98	4
Candidate D	4	162	1
Candidate E	5	19	5
Total		493	
Average		98.6	
Threshold		108.5	

The Czech municipal election system uses D’Hondt’s method of proportional allocation of mandates. This method relies on calculating vote shares for each candidate. These vote shares are determined by the the total number of votes the candidate’s party received and the candidate’s final rank. In Table B.2, I illustrate the vote share calculation and mandate allocation on a municipality election, where the council has five seats. There are three slates on the ballot, each with five candidates. The candidates are already in their final rank, determined by the method explained above. In the top panel of Table B.2, I show the final distribution of the votes to each of the candidates – for instance, the first-ranked candidate on Slate A received 315 votes. The last column shows the sum of the votes for each of the slates.

After summing the votes, the vote shares for each candidate i on slate s are calculated using the following formula:

$$vote_share_{i,s} = \frac{total_votes_s}{final_rank_{i,s}} \quad (B.1)$$

The middle panel of Table B.2 illustrates this division: for each candidate, it takes the total votes for their slate and divides it by the candidate’s final rank. The bottom panel of Table B.2

⁵⁵To get 10% more than the average votes, I calculate $98.6 \cdot 1.1 = 108.5$.

Table B.2 Post-election vote distribution

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Total votes
Final vote distribution						
Slate A	315	238	204	106	67	930
Slate B	427	389	380	276	182	1654
Slate C	251	236	154	118	93	852
Share calculation						
Slate A	930/1	930/2	930/3	930/4	930/5	930
Slate B	1654/1	1654/2	1654/3	1654/4	1654/5	1654
Slate C	852/1	852/2	852/3	852/4	852/5	852
Final shares and madates allocation						
Slate A	930.0	465.0	310.0	232.5	186.0	–
Slate B	1654.0	827.0	551.3	413.5	330.8	–
Slate C	852	426.0	284.0	213.0	170.4	–

shows the results of these calculations. Finally, the five council mandates are distributed to the candidates with the five highest shares. These are highlighted in the bottom panel in bold: slate B won three seats on the council, while slate A and C won one mandate each. These mandate allocations follow proportional representation: slate B won 48% of the overall votes and 60% of the mandates, whereas slates A and C received 27% and 25% of the votes, respectively, and won 20% of the seats each.

B.2 Calculating margins

As explained in the main text, the vote margin is the margin of the *female* to the *male* candidate within a gender-mixed contests. The calculation of the vote margins is done in two separate ways depending on the distribution of the votes and the final ranking of the candidates. Given the threshold for moving on the slate rank, there can be cases where the last winning candidate has less individual votes than the first losing candidate, essentially making the winning candidate's victory margin negative. In terms of the female vote margin, it could happen that the winning female has fewer votes than the losing male, thus winning with a negative margin. Similarly, a male that won with fewer votes than the female would translate into a positive vote margin for the losing female. This feature is clearly undesirable for the RD design. Therefore, in the cases where the winning candidate got a lower number of votes than the losing candidate, I use the relative distance of the losing candidate to the threshold for changing final rank.

Case 1: The winning candidate received more votes than the losing candidate

In the straightforward scenario where the winner obtained a higher number of votes than the loser, the margin is simply calculated as:

$$vote_margin = \frac{votes_marginal_female - votes_marginal_male}{total_votes_for_slate} \quad (B.2)$$

that is, the distance between the female and the male candidate, in percentage points. Note that if the female candidate won the contests, the vote margin will be positive, while if the male candidate won, the vote margin will be negative.

Case 2: The winning candidate received fewer votes than the losing candidate

In the cases where the losing candidate has more votes than the winning candidate, but does not reach the threshold for moving up the rank, I calculate the margin as distance of the losing candidate to the threshold. This margin stands for how many votes they would need to reach the threshold, move up the rank, and thus be elected to the council.

The calculation in this case differs depending on the gender of the loser. If the female candidate won the race with fewer votes than the (losing) male candidate, the margin is the following:

$$vote_margin = \frac{threshold - votes_marginal_male}{total_votes_for_slate} \quad (B.3)$$

On the other hand, if the male candidate won, the margin of the (losing) female candidate to the (winning) male is:

$$vote_margin = (-1) \cdot \frac{threshold - votes_marginal_female}{total_votes_for_slate} \quad (B.4)$$

Notice that in this calculation, I multiply the fraction on the right hand side by (-1). As the female lost the contest, the margin must be negative, though the size of the margin is still calculated as the distance of the loser to the threshold.

The only limitation of this margin calculation comes in cases where both marginal winner and marginal loser obtained equal number of votes and this number of votes was higher than the threshold for changing the final rank. In such scenario, the final rank is determined by their initial rank on the slate, making the vote margin effectively zero, without any possibility to calculate it using the threshold, since *both* candidates surpassed the threshold. The municipalities in which such contests occur are therefore excluded from further analysis. This concerns 102 municipalities in each election (covering 103 and 110 mixed contests in 2010 and 2014, respectively).

C Power calculations

The results section showed no evidence to support the alternative hypothesis that increased female representation has a statistically significant effect on the share of female expenditures in the municipality budgets. A potential problem that the estimates might suffer from is the possible lack of statistical power which would affect the ability of detecting statistically significant effects. Therefore, this section presents a set of power calculations for the four main estimations presented in this paper.

Cattaneo et al. [2019c] describe a power calculation approach for ex-post RD analysis which is relevant in a setting where the researcher has already finalized data collection. Unlike conventional power calculations in experimental settings, which determine the necessary sample size to achieve a specific level of statistical power, this method is used to analyze the statistical power for a hypothesized treatment effect *given* the sample size of the sample used in the `rdrobust` estimation. Using the statistical package `rdpow`, developed by Cattaneo et al. [2019c], the power of the RD estimation can be calculated for any specific size of the treatment effect, denoted τ . The default size of the effect is equal to half of the standard deviation of the outcome variable for the control group.

Table C.1 Power calculations

	τ	$0.75 \cdot \tau$	$0.5 \cdot \tau$	$0.25 \cdot \tau$	$0.125 \cdot \tau$
<i>Model A: Share of female spending (sharp)</i>					
Effect size	0.036	0.027	0.018	0.009	0.004
Statistical power	>0.999	>0.999	0.998	0.689	0.232
<i>Model B: Share of female spending (fuzzy)</i>					
Effect size	0.036	0.027	0.018	0.009	0.005
Statistical power	0.121	0.089	0.067	0.054	0.051
<i>Model C: Share of education spending</i>					
Effect size	0.038	0.029	0.019	0.010	0.005
Statistical power	>0.999	>0.999	0.998	0.677	0.227
<i>Model D: Change in share of female spending</i>					
Effect size	0.038	0.029	0.019	0.010	0.005
Statistical power	>0.999	>0.999	0.986	0.550	0.181

Note: The table presents the results of power estimations using a 95% significance level. The default size of the effect for which the power is estimated, τ , is equal to half of the standard deviation from the control group. The first column shows the power estimation for the effect of the size τ , while the subsequent columns show the power for a size of three-quarters, one-half, one-quarter and one-eighth of τ , respectively. For each model, the relevant size of the effect is reported in the first row, while the second row show the statistical power using the robust bias-corrected power test. Model A is the baseline sharp RD model from Table 4, Model B comes from the fuzzy RD estimation in Table 6, Model C shows the model from Table 7 which uses education expenditures as the dependent variable, and Model D is taken from the discontinuity-in-difference estimation in Table 8. All Models use optimal bandwidth and linear fit.

Following the method used by Bagues and Campa [2017], I perform the power calculations for a range of hypothetical effects of different sizes, ranging from conservative to rather large. I

use the default size of the effect τ and scale it for the different scenarios. The scenario effect sizes in the five scenarios are τ (the least conservative), 0.75τ , 0.5τ , 0.25τ and 0.125τ (the most conservative). The power test is performed for the four models from the thesis: the baseline sharp RD model of the female victory and the average share of female spending (Table 4), the fuzzy RD estimation with the same outcome variable (Table 6), the sharp RD estimation using the average share of education spending (Table 7), and finally, the sharp RD estimation using the difference in the share of female spending between the legislatures as the outcome variable (Table 8). In Model A, the size of the effect ranges between 0.5 p.p. and 3.6 p.p. for the most and least conservative scenarios, respectively, which is comparable to the effect sizes used in the power calculations by [Bagues and Campa \[2017\]](#).⁵⁶ The hypothetical effect sizes for the most and least conservative scenarios in Models B, C, and D are similar.

Table C.1 reports the results of the power calculations, and Figure C.1 provides visualization of the power test results. It is clear that all the sharp RD models provide much higher power than the fuzzy RD: for the least conservative scenario of an effect of 3.7 p.p., the fuzzy RD model offers a statistical power of only 12%, much lower than that of the sharp RD models. Generally, the sharp RD models offer high power in the scenarios with large effect, which is for all three models just under 4 p.p.; the power remains around 99% also for the effect of half the size, that is 1.8 p.p for Model A or 1.9 p.p for Models B and C. However, for the most conservative scenario of an effect around 0.5 p.p., the results are generally underpowered, ranging between 18.1% and 23.2%.

Table C.2 Minimum detectable effect size (MDES)

	Model A	Model B	Model C	Model D
MDES	0.0109	0.1267	0.0119	0.0150

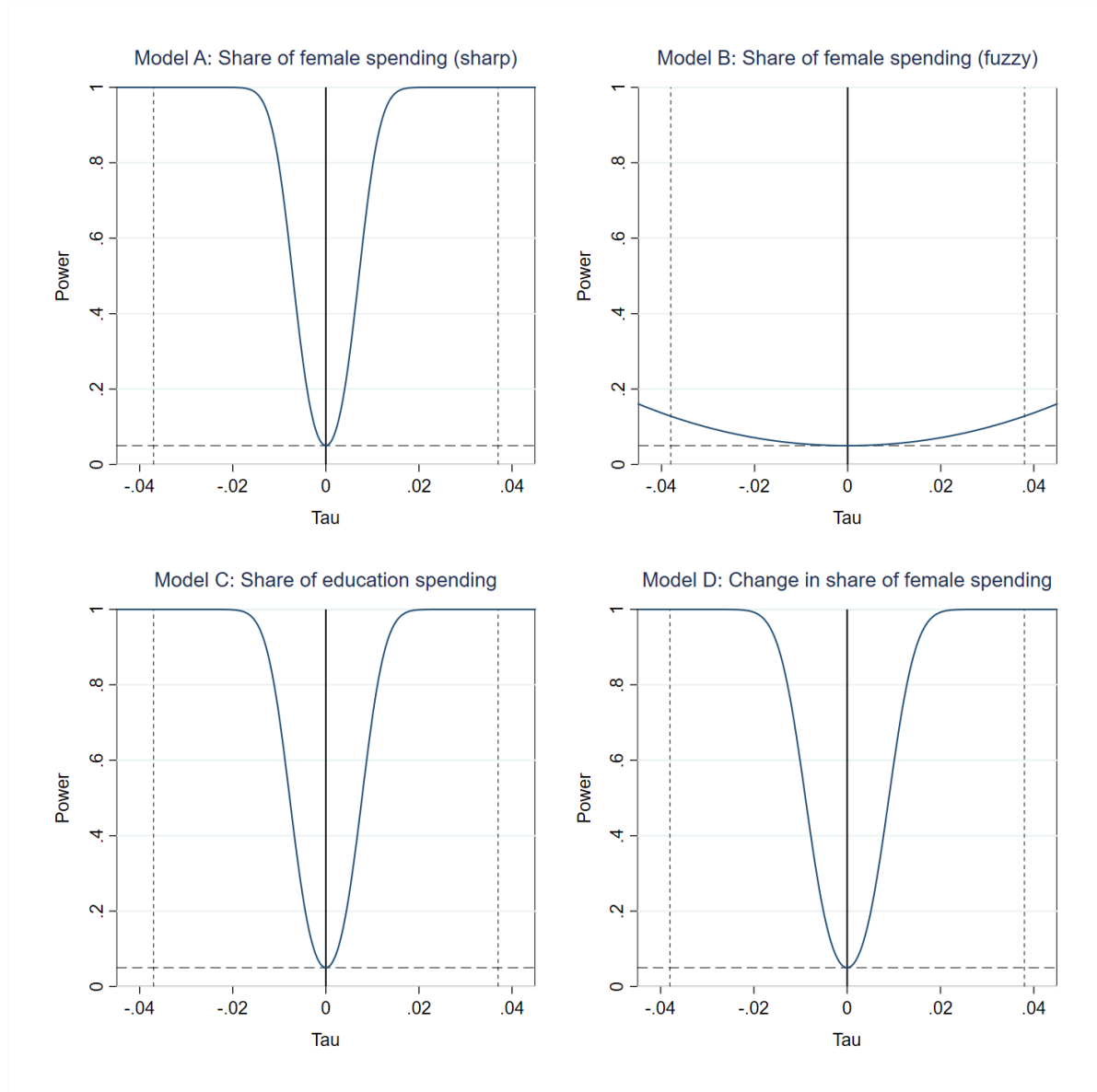
Note: This table presents a calculation of the minimum detectable effect size (MDES) using 5% significance level and 80% power level. The MDES is calculated by multiplying the bias-corrected robust standard errors from the relevant main estimation by the multiplier of 2.8 (given by the significance and power level). Models A, B, C and D come from column 1 in Tables 4, 6, 7 and 8, respectively. Note that the standard errors reported in the main analysis are the *conventional* standard errors and thus they differ from the bias-corrected robust standard errors used for this calculation.

Finally, in addition to the power calculations, I also perform the calculation of the minimum detectable effect size (MDES), which shows the minimum effect that my model has the power to detect, at a given power and significance level. For the calculations in this section, I use the conventional 80% power level and 5% significance level. The MDES is calculated by multiplying the standard error of the RD coefficient from the estimation by 2.8, a multiplier that is determined by the power and significance level. Table C.2 reports the MDES for the four main models. Similar to the power calculations above, it is clear that the fuzzy RD estimation in Model B has much lower power than the remaining models: it is able to detect an effect of 12.7 p.p. at 80% power level. This MDES is an order of magnitude higher than the MDES of the

⁵⁶[Bagues and Campa \[2017\]](#) use three effect sizes: 1 p.p. for the most conservative, 2 p.p. for the intermediate and 3 p.p. for the least conservative scenario.

sharp RD Models A, C and D, which equals 1.1 p.p., 1.2 p.p. and 1.5 p.p., respectively.

Figure C.1 Power calculations



Note: The plots represent the value of the statistical power for the particular models. The dashed horizontal line shows the 0.05 significance level. The vertical dotted lines represent the value of the default τ , equal to half the standard deviation in the control group.

D Data sources

‘Our Society’ survey data

Institute of Sociology at Academy of Sciences of the Czech Republic, Public Opinion Research Center. (2019). Our Society 2019 - February. Online data file. DOI: 10.14473/V1902

Institute of Sociology at Academy of Sciences of the Czech Republic, Public Opinion Research Center. (2017). Our Society 2017 - January. Online data file. DOI: 10.14473/V1701en

Institute of Sociology at Academy of Sciences of the Czech Republic, Public Opinion Research Center. (2016). Our Society 2016 - February. Online data file. DOI: 10.14473/V1602

Institute of Sociology at Academy of Sciences of the Czech Republic, Public Opinion Research Center. (2015). Our Society 2015 - September. Online data file. DOI: 10.14473/V1509

Institute of Sociology at Academy of Sciences of the Czech Republic, Public Opinion Research Center. (2013). Our Society 2013 - June. Online data file. DOI: 10.14473/V1306

Institute of Sociology at Academy of Sciences of the Czech Republic, Public Opinion Research Center. (2013). Our Society 2013 - January. Online data file. DOI: 10.14473/V1301

Institute of Sociology at Academy of Sciences of the Czech Republic, Public Opinion Research Center. (2012). Our Society 2012 - June. Online data file. DOI: 10.14473/V1206

Institute of Sociology at Academy of Sciences of the Czech Republic, Public Opinion Research Center. (2012). Our Society 2012 - January. Online data file. DOI: 10.14473/V1201

Spending Data

Monitor Ministry of Finance of the Czech Republic (2019). Fin 2-12 M - Budget Fulfillment of LGU. Online data file. Retrieved from: <https://monitor.statnipokladna.cz/datovy-katalog/transakcni-data>

Election Data

Czech Statistical Office, Volby.cz Portal. (2014). Open data for local council elections 2014. Online data file. Retrieved from: https://volby.cz/opendata/kv2014/kv2014_opendata.htm

Czech Statistical Office, Volby.cz Portal. (2010). Open data for local council elections 2010. Online data file. Retrieved from: https://volby.cz/opendata/kv2010/kv2010_opendata.htm

Demographic Data

Czech Statistical Office. (2018). Small Lexicon of Municipalities in the Czech Republic 2018. Online data file. Retrieved from: <https://www.czso.cz/csu/czso/maly-lexikon-obci-ceske-republiky-2018>

Czech Statistical Office. (2017). Small Lexicon of Municipalities in the Czech Republic 2017. Online data file. Retrieved from: <https://www.czso.cz/csu/czso/maly-lexikon-obci-ceske-republiky-2017>

Czech Statistical Office. (2016). Small Lexicon of Municipalities in the Czech Republic 2016. Online data file. Retrieved from: <https://www.czso.cz/csu/czso/maly-lexikon-obci-ceske-republiky-2016>

Czech Statistical Office. (2015). Small Lexicon of Municipalities in the Czech Republic 2015. Online data file. Retrieved from: <https://www.czso.cz/csu/czso/maly-lexikon-obci-ceske-republiky-2015>

Czech Statistical Office. (2014). Small Lexicon of Municipalities in the Czech Republic 2014. Online data file. Retrieved from: <https://www.czso.cz/csu/czso/maly-lexikon-obci-ceske-republiky-2014-n-gdc2kaznu1>

Czech Statistical Office. (2013). Small Lexicon of Municipalities in the Czech Republic 2013. Online data file. Retrieved from: <https://www.czso.cz/csu/czso/maly-lexikon-obci-ceske-republiky-2013-nl6r6nhlpw>

Czech Statistical Office. (2012). Small Lexicon of Municipalities in the Czech Republic 2012. Online data file. Retrieved from: <https://www.czso.cz/csu/czso/maly-lexikon-obci-cr-2012-1kd2vridet>

Czech Statistical Office. (2011). Small Lexicon of Municipalities in the Czech Republic 2011. Online data file. Retrieved from: <https://www.czso.cz/csu/czso/maly-lexikon-obci-cr-2011-kbfc6ux0qg>

Czech Statistical Office. (2011). Population Census 2011. Online data file. Retrieved from: https://www.czso.cz/csu/czso/otevrena_data_pro_vysledky_scitani_lidu_domu_a_bytu_2011