UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Essays in Political Economy

Ensaios em Economia Política

Marcel Ferreira de Oliveira

Orientador: Prof. Dr. Ricardo de Abreu Madeira
Co-orientador: Prof. Dr. Marcos Yamada Nakaguma

São Paulo

2021

Prof. Dr. Vahan Agopyan

Reitor da Universidade de São Paulo

Prof. Dr. Fábio Frezatti

Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. José Carlos de Souza Santos

Chefe do Departamento de Economia

Prof. Dr. Wilfredo Fernando Leiva Maldonado

Coordenador do Programa de Pós-Graduação em Economia

Marcel Ferreira de Oliveira

Essays in Political Economy

Ensaios em Economia Política

Tese apresentada ao Programa de Pós-Graduação em Economia do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, como requisito parcial para a obtenção do título de Doutor em Ciências.

Área de Concentração: Teoria Econômica.

Orientador: Prof. Dr. Ricardo de Abreu Madeira

Co-orientador: Prof. Dr. Marcos Yamada Nakaguma

Versão Corrigida

(versão original disponível na Biblioteca da Faculdade de Economia, Administração e Contabilidade)

São Paulo

2021

Catalogação na Publicação (CIP)
Ficha Catalográfica com dados inseridos pelo autor

Oliveira, Marcel Ferreira de.
Ensaios em Economia Política / Marcel Ferreira de Oliveira. - São Paulo, 2021.
148 p.

Tese (Doutorado) - Universidade de São Paulo, 2021.
Orientador: Ricardo de Abreu Madeira.
Co-orientador: Marcos Yamada Nakaguma.

1. Economia Política. 2. Regressão Descontínua. 3. Eleições acirradas. 4. Gênero. 5. Câmaras Municipais. I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

Acknowledgements

Os agradecimentos principais são direcionados aos meus orientadores, os professores Ricardo Madeira e Marcos Nakaguma. Não apenas pela orientação, mas também pela solicitude, paciência e empenho em me auxiliar na confecção desta tese.

Além deles, outro agradecimento especial se deve à Fundação Instituto de Pesquisas Econômicas (FIPE), pelo apoio financeiro durante a maior parte do curso.

Agradecimentos especiais também são direcionados aos demais professores do doutorado em Economia, tanto em Teoria Econômica quanto em Economia Desenvolvimento, da Faculdade de Economia, Administração e Contabilidade, pela diposição em auxiliar e ensinar os alunos do curso. De modo similar, também aradeço aos funcionários do serviço de pós-graduação e de outros serviços essenciais para a manutenção do programa. Bem como aos colegas que fiz durante esses anos de doutorado.

Por fim, agradeço aos meus pais por todo o suporte que recebi até o momento.

Resumo

Esta tese consiste em dois ensaios divididos em capítulos. No primeiro capítulo, nós investigamos se o tamanho das casas legislativas municipais tem alguma influência na composição de gênero das casas legislativas. Para fazer isso, empregamos um design de Regressão Descontínua para explorar uma variação exógena no tamanho das casas legislativas municipais no Brasil. Essa variação resultou de uma regra que determinou de modo preciso o tamanho das casas legislativas de acordo com faixas populacionais para as eleições de 2004 e 2008. Nós encontramos que o número de assentos no legislativo tem um efeito positivo e significante na proporção de mulheres na legislatura. Conforme o número de assentos aumenta, mais candidatos participam nas eleições. No entanto, a maioria desses candidatos são homens e eles competem entre si. Essa competição não afeta tanto as mulheres, o que resulta numa melhora das chances de uma mulher ser eleita. Nós desenvolvemos um modelo teórico para formalizar este argumento. Nós desenvolvemos um modelo de coordenação imperfeita em que os eleitores têm preferências distintas por candidatos homens e mulheres. Sob a hipótese de que é mais custoso para mulheres participarem da eleição e que os eleitores são suficientemente enviesados contra mulheres, nós conseguimos modelar e replicar esse efeito de canibalização. Além disso, esse aumento na representaçõ feminina foi seguido por melhoras na educação infantil, saúde pré-natal e infantil e na assistência comunitária e social.

No segundo capítulo, nós investigamos os impactos de eleger uma mulher para uma casa legislativa municipal em um contexto de eleições acirradas no Brasil. O sistema de representação proporcional por lista aberta para eleições do legislativo municipal no Brasil nos permite delinear um quase-experimento no qual comparamos pares de candidatos que venceram ou perderam por pouco as eleições. Para cada município, existem pares de candidatos desse tipo para cada coalizão. Destes, escolhemos aqueles que possuem um pair de gênero misto e selecionamos dentre eles o que tem a menor margem de vitória do município. Devido à quantidade de incerteza associada às perspctivas de ganhar ou perder para esses pares, é possível comparar municípios onde a mulher venceu com municípios onde o homem venceu. Usamos essa estratégia para estimar o efeito causal de eleger uma mulher na margem em um contexto legislativo na provisão de bens públicos. Em contraste com estudos anteriores, nós não encontramos nenhuma evidência robusta de que haja algum efeito forte de gênero do legislador em resultados políticos relacionados à educação infantil, saúde pré-natal e infantil e no gasto do governo. Essa falta de evidências robustas pode estar relacionado com o nosso desenho quase-experimental, que foca em candidatos que quase ganharam as eleições. Nós argumentamos que simplesmente eleger mulheres pode não ser o suficiente para observar impactos significantes sobre o processo de elaboração de políticas, nos levando a acreditar que outros fatores políticos podem estar servindo de pré-requisitos para sua participação efetiva. **JEL codes:** D72, D78, H41.

Palavras-chave: Economia Política, Regressão Descontínua, Eleições Acirradas , Gênero e Eleições para o legislativo local.

Abstract

This thesis is divided into two essays divided by chapters. In the first chapter, we investigate whether the size of legislative houses has any influence on the gender composition of the legislative house. To do so, we employ a Regression Discontinuity design to exploit an exogenous variation in the size of local legislatures in Brazil. This variation resulted from a rule that determined precisely the size of local legislatures according to population thresholds for the elections of 2004 and 2008. We found that the number of seats in the legislature has a significant and positive impact on the share of women in the legislature. One possible explanation for this result is that male candidates cannibalize each other. As the number of seats increases, more candidates participate in the elections. However, most of these candidates are men and they compete with each other. This competition does not affect women as much, which results in an improvement in the chance that women will be elected. We develop a theoretical framework to formalize this argument. We develop a model of imperfect coordination in which voters have distinct preferences for male and female candidates. Under the assumption that it is costlier for women to participate in the election and that the voters are sufficiently biased against women, we are able to model and replicate this cannibalization effect. Moreover, this increase in female representation was followed by improvements in early education, antenatal and infant health care and social and community assistance.

In the second chapter, we investigate the impacts of electing a woman to a local legislative house in the context of close elections in Brazil. Brazil's open-list proportional representation system for its local legislative elections allows us to design a quasi-experiment in which we compare pairs of candidates that barely won or barely lost the elections. For each municipality, there are such pairs for each coalition. From those, we pick the pairs that have a mixed-gender pair and select the one with the lowest vote margin in the municipality. Due to the amount of uncertainty associated with the prospects of winning or losing associated with these pairs, we can compare municipalities where a woman won with municipalities where a man won. We use this strategy to estimate the causal effect of electing a woman at the margin in a legislative context on the provision of public goods. In contrast to previous studies, we find no strong evidence that there is any effect of the gender of the legislator on policy outcomes related to early education, prenatal and child health care, and government spending. This lack of robust results might be related to our quasi-experimental design, which focuses on candidates that barely won the elections. We argue that simply electing women might not be enough to observe significant impacts in policy-making, leading us to believe other political factors might serve as a prerequisite for their effective participation.

JEL codes: D72, D78, H41.

Keywords: Political Economy, Regression Discontinuity, Close Elections, Gender and Local Legislative Elections.

List of Figures

List of Tables

Contents

1 THE SIZE OF LOCAL LEGISLATURES AND WOMEN'S POLITICAL REPRESENTATION: REVISITING EVIDENCE FROM BRAZIL

1.1 Introduction

Between 1997 and 2020, the proportion of seats held by women in national parliaments worldwide has risen from 11.7% to 25.6%, according to the Inter-Parliamentary Union (IPU)[1]. Despite this steady increase in female representation, the numbers still fall short of full parity. This leads to two questions. First, how relevant is the identity of the politician, gender in particular, to policy-making? Second, which factors hinder or enhance women's participation in politics? The literature has found that in many different settings the gender of politicians indeed leads to different policy outcomes. For instance, Pande (2003) finds that political reservations increase transfers to groups that benefit from the mandate. In the same line, Chattopadhyay and Duflo (2004) find evidence that in the case of men and women, village heads in India tend to invest more in infrastructure that is relevant to the needs of their own gender. The reason for these discrepancies has to do with the fact men and women have, on average, different policy preferences. Women tend to have stronger preferences for policies associated with early education, prenatal and infant health care and redistribution than men (THOMAS; WELCH, 1991; LOTT; KENNY, 1999; ALESINA; FERRARA, 2005), and these preferences lead to differences in policies implemented by female and male politicians. Women elected to office influence the provision of child education, antenatal and childhood health services, and even investment in community and social assistance (MASON; KING, 2001; DUFLO, 2003; SVALERYD, 2009; BHALOTRA; CLOTS-FIGUERAS, 2014; BROLLO; TROIANO, 2016).

These results are in line with citizen-candidate models following Osborne and Slivinski (1996) and Besley and Coate (1997), in which candidates or parties have their own preferences and may not be able to commit to moderate policies, leading to divergence in policies implemented by whoever won the election. As consequence, different groups may be likely to influence policy decisions in the same direction of their own preferences.

Given that the identity of politicians matters for the public provision of services and goods, and that women are under-represented in all political positions, including the legislative houses, we are still left with the question of which factors may hinder or enhance women's ability to participate in the decision-making process of public policy. The possible factors are many,

[1] More information at: ipu.org

ranging from cultural to social, economic and political factors[2]. Bhalotra, Clots-Figueras and Iyer (2018) find that there is a feedback effect in which electing women to state legislatures leads to subsequent political participation of women, with more female incumbents attempting re-election. They also find that in states with more entrenched gender biases there is a backlash effect against new female candidates. There is also evidence that the presence of female political role models propels the engagement of young women (CAMPBELL; WOLBRECHT, 2006). Arvate, Firpo and Pieri (2021) also found that simply having more women participating in politics may not be enough to induce further engagement of women in politics. Instead, this role model effect is tied to the electoral performance of female politicians. A woman winning an election encourages young women to engage in politics, but a defeat has the opposite effect. Their finding applies to mayoral electoral races in Brazil. They also find that this effect is stronger when the legislative branch also displays a higher concentration of women.

The political structure may also affect women's participation in politics, such as the parliament structure (number of chambers, number of seats, maximum length of a legislature, etc) (OAKES; ALMQUIST, 1993), party system (number of parties, how candidates are selected, etc), electoral system (plurality vs. proportional representation) (NORRIS, 2006), length of democratic experiment (TREMBLAY, 2007) and possibly other factors.

In this chapter, we contribute to this strand of literature by investigating the relationship between the size of legislative houses and representation of women[3]. To do so, we exploit an exogenous variation in the size of local legislatures in Brazil. In 2004, there was a change in the way the number of seats in Brazilian municipal councils was determined. This change was also valid for the 2008 municipal elections. The new rule assigned the number of seats according to population intervals, which allows us to employ a sharp regression discontinuity design to compare municipalities around the population thresholds. Despite there being multiple population intervals, 95% of Brazilian municipalities have less than 95.238 inhabitants, which falls around the first threshold, which is our focus here.

We find that an increase by 1 seat in the legislative leads to around a 10% increase in the number of candidates, and most of these entrants are men. However, this entry of male candidates only fosters competition among men. Women candidates' vote share did not change significantly with the entry of new male candidates. As result, more women were elected. To make sense of this chain of events, we develop a model of imperfect coordination in which voters have distinct preferences for male and female candidates. In this setting, voters are biased against women, but there is still a fraction of the electorate that prefers women. However, voters cannot coordinate perfectly to choose a candidate, leading some votes to be 'wasted'. For a certain

[2] Tremblay (2007) does an exploratory research to find which cultural and political factors are correlated to the proportion of women in lower or single houses of parliaments. Among those, she finds these correlations are sensitive to the length of the democratic experiment. Among the factors examined, she finds a positive correlation between the level proportionality of voting systems and participation of women in parliaments for countries that had not a long democratic experiment, meanwhile, in long-lasting democracies, the gender roles seem to have a bigger correlation.

[3] We build upon the work of Correa and Madeira (2014), which had previously studied this topic for Brazilian elections. We extend their findings and provide a theoretical framework to explain these results.

degree of lack of coordination and assuming it is costlier for women to join the elections, we show there it is possible for a situation in which an increase in the mass of seats generates a "cannibalization effect" among men. As the number of seats increases, more men join the election relative to women. However, there is also a relative increase in the proportion of elected female candidates, leading to a more equitable legislature.

Alongside with more elected women, we also find that changes in female participation translate into policy outcomes. We show that larger local councils lead to a higher share of women elected and also to sizable improvements in the provision of public goods and services usually associated with female preferences, such as early education, antenatal health care and community and social assistance. These results are in line with the literature that deals with the relationship between gender and public policy, as mentioned before.

The remainder of this paper is as follows. Section 2 describes our theoretical framework. Section 3 describes the institutional background for the elections examined. Section 4 presents our data and estimation strategy. Section 5 present our main results. Section 6 performs a few robustness checks. Section 7 presents our conclusions.

1.2 Theoretical Framework

In this section, we propose a theoretical framework to examine the impact of an increase in the size of the legislature on the entry of male and female candidates into politics, focusing in particular on candidacy and election rates among men and women.

1.2.1 Model

Suppose that there are two types of candidates, men (M) and women (W) $-$ candidates differ only in terms of their gender, but are otherwise identical. There is a continuum of voters with measure one and voters may belong to one of two groups: (i) type m voters always vote for men and (ii) type w voters always vote for women. The fraction of voters of type m is $\alpha \in (0, 1)$. We focus our analysis on the case where $\alpha > \frac{1}{2}$, so that the parameter α may be interpreted as a measure of the electorate's bias against female candidates.

Suppose that there is a continuum of seats $s < 1$ in the legislature and let the mass of male and female candidates be denoted by n_M and n_W, respectively. We assume that voters are able to coordinate to elect candidates of their preferred type, but only imperfectly. A perfect coordination would correspond to the case where voters of each group are able to precisely allocate their votes in such a way as to maximize the number of elected candidates of their preferred type. For instance, if κ votes are required to elect a candidate in equilibrium, then voters of each group would coordinate to cast exactly κ votes for the largest possible number of candidates belonging to their preferred type.

Our model assumes that coordination among voters becomes more difficult as the mass of candidates of a given type increases. Specifically, we suppose that the total mass of "effective votes" for male and female candidates are given, respectively, by:

$$v_M(n_M) = \alpha(1 - \xi n_M) \tag{1.1}$$

and

$$v_W(n_W) = (1 - \alpha)(1 - \xi n_W), \tag{1.2}$$

where $\xi \in (0,1)$, with $\alpha \xi n_M$ and $(1-\alpha)\xi n_W$ representing the amount of votes "wasted" by each group. Intuitively, the parameter ξ captures, in a reduced form fashion, the existence of coordination frictions among voters. Wasted votes are those which are cast on candidates who have no chances of being elected, while effective votes are those which can be perfectly allocated. We believe that our model is able to capture some of the essential features of the open-list proportional system, which is used to elect members of municipal councils in Brazil. The open-list proportional system is well-known for making coordination among voters is very difficult, since most of the votes are cast directly on individual candidates (COX, 1999).

Given the mass of candidates of each type, n_M and n_W, the amount of votes required to elect a candidate is given by:

$$\kappa(n_M, n_W) = \frac{v_M(n_M) + v_W(n_W)}{s} \Rightarrow \kappa(n_M, n_W) = \frac{\alpha(1-\xi n_M) + (1-\alpha)(1-\xi n_W)}{s}, \tag{1.3}$$

where the electoral threshold $\kappa(n_M, n_W)$ is the total number of effective votes divided by the number of seats.

Under the assumption that effective votes can be perfectly allocated, the mass of male and female candidates elected is given, respectively, by:

$$e_M(n_M, n_W) = \frac{v_M(n_M)}{\kappa(n_M, n_W)} \Rightarrow e_M(n_M, n_W) = \frac{\alpha(1-\xi n_M)}{\alpha(1-\xi n_M) + (1-\alpha)(1-\xi n_W)} s \tag{1.4}$$

and

$$e_W(n_M, n_W) = \frac{v_W(n_W)}{\kappa(n_M, n_W)} \Rightarrow e_W(n_M, n_W) = \frac{(1-\alpha)(1-\xi n_W)}{\alpha(1-\xi n_M) + (1-\alpha)(1-\xi n_W)} s \tag{1.5}$$

We assume that voters choose to coordinate on candidates randomly, given that they are all ex-ante identical. Thus, the probability of election of a particular candidate is:

$$p_M(n_M, n_W) = \frac{e_M(n_M, n_W)}{n_M} \Rightarrow p_M(n_M, n_W) = \frac{\alpha(1-\xi n_M)}{\alpha(1-\xi n_M) + (1-\alpha)(1-\xi n_W)} \frac{s}{n_M} \tag{1.6}$$

and

$$p_W(n_M, n_W) = \frac{e_W(n_M, n_W)}{n_W} \Rightarrow p_W(n_M, n_W) = \frac{(1-\alpha)(1-\xi n_W)}{\alpha(1-\xi n_M) + (1-\alpha)(1-\xi n_W)} \frac{s}{n_W} \tag{1.7}$$

Finally, each agent forms beliefs about the mass of candidates of each type, n_M and n_W, and decide whether to enter or not the electoral race by comparing the expected benefit of doing so with the opportunity cost of running. Suppose that the benefit associated with being in office is $B > 0$ and that the cost of running is c_i for $i \in \{M, W\}$, with $0 < c_i < B$. Therefore, an agent of type i decides to enter the race if, and only if:

$$p_i(n_M, n_W) B \geq c_i \tag{1.8}$$

Following Chattopadhyay and Duflo (2004), we assume that the cost of running for women is larger than the cost of running for men, $c_W > c_M$. This assumption seems to be realistic in the context of Brazilian politics for cultural reasons, especially in the poorer and less-developed regions of the country.

1.2.2 Equilibrium

We solve the model for a Nash equilibrium where agents make their decisions to enter the race optimally. In equilibrium, the mass of candidates of each type must be such that:

$$p_M(n_M, n_W)B = c_M \tag{1.9}$$

and

$$p_W(n_M, n_W)B = c_W \tag{1.10}$$

Intuitively, agents have an incentive to run for office as long as the expected benefit of doing so is larger than the cost of running. Observe that, as additional candidates enter the race, the election probability of each candidate goes down. Thus, in equilibrium, the expected benefit must be exactly equal to the cost of running for candidates of both types.

Hence, in equilibrium, we must have:

$$\frac{\alpha(1-\xi n_M)}{\alpha(1-\xi n_M)+(1-\alpha)(1-\xi n_W)}\frac{s}{n_M}B = c_M$$

$$\frac{(1-\alpha)(1-\xi n_W)}{\alpha(1-\xi n_M)+(1-\alpha)(1-\xi n_W)}\frac{s}{n_W}B = c_W$$

Solving this system of equations for n_M and n_W, we obtain:

$$n_M^* = \frac{\alpha(c_M+c_W)sB\xi + c_M(c_W - sB\xi) - \Delta}{2c_M((-1+\alpha)c_M + \alpha c_W)\xi} \tag{1.11}$$

and

$$n_W^* = \frac{\alpha(c_M+c_W)sB\xi - c_M(c_W + sB\xi) + \Delta}{2c_W((-1+\alpha)c_M + \alpha c_W)\xi} \tag{1.12}$$

where

$$\Delta \equiv \sqrt{4(1+\alpha)c_M c_W(\alpha c_W - (1-\alpha)c_M)sB\xi + (c_M c_W - (\alpha c_W - (1-\alpha)c_M)sB\xi)^2}$$

Note that the expressions above are well-defined for any $\alpha \in \left(\frac{1}{2}, 1\right)$ provided that $\xi < \frac{c_M+c_W}{sB}$; otherwise, the degree of coordination frictions would be so large as to make the number of effective votes zero in equilibrium, i.e. $\alpha(1-\xi n_M^*)+(1-\alpha)(1-\xi n_W^*) = 0$. Hence, we focus our analysis on the case where $\xi < \frac{c_M+c_W}{sB}$ in order to avoid those uninteresting equilibria.

The next proposition provides a general characterization of the equilibrium of the game.

Proposition 1. *Suppose that $c_W > c_M$ and $\xi < \frac{c_M+c_W}{sB}$. In equilibrium, the mass of candidates of each type, n_M^* and n_W^*, is given by (1.11) and (1.12) respectively, with:*

$$n_W^* < n_M^*$$

Moreover, there exists $\overline{\alpha} \in \left(\frac{1}{2}, 1\right)$ *such that if* $\alpha > \overline{\alpha}$*, then:*

$$e_W^* < e_M^*$$

Our model captures some basic features of women participation in politics, particularly the fact that women are less likely to become candidates, $n_W^* < n_M^*$, and win elections, $e_W^* < e_M^*$. Interestingly, observe that the fact that there are more male candidates does not necessarily translate into more men being elected, due to the effect that an increase in the number of candidates has on wasted votes for a particular type. However, we show that in an environment where the bias against women, as captured by the parameter α, is large enough then fewer women will, indeed, be elected. Therefore, our analysis highlights the distinct roles played by the cost of running and bias against women in shaping electoral outcomes.

Next, we examine the effect of an increase in the mass of seats s on the mass of candidates and the share of elected candidates of each type. Our main comparative static result is stated in the following proposition.

Proposition 2. *Suppose that* $c_W > c_M$*. There exist thresholds* $\underline{\xi} > 0$ and $\overline{\xi} < \frac{c_M + c_W}{sB}$ *such that if* $\underline{\xi} < \xi < \overline{\xi}$*, then:*

$$\frac{\partial n_M^*}{\partial s} > \frac{\partial n_W^*}{\partial s}$$

and

$$\frac{\partial (e_M^*/s)}{\partial s} < \frac{\partial (e_W^*/s)}{\partial s}$$

Thus, we show that there exists a region of parameters where an increase in the mass of seats generates a "cannibalization effect" among men, leading to the relative entry of more male than female candidates, while at the same time causing the relative increase in the share of elected female candidates. Interestingly, it is possible to show that if coordination frictions are small, $\xi < \underline{\xi}$, then men enter the race more and gain more seats, whereas, conversely, if $\xi > \overline{\xi}$, then women enter the race more and gain more seats. As discussed above, coordination problems are particularly relevant in open-list proportional systems, so that we expect an increase in the mass of seats to lead to relatively more women being elected under this system.

1.3 Institutional Background

1.3.1 Local politics in Brazil

Brazil is a federative republic whose political and administrative organization comprises the Union, the states, the Federal District and the municipalities, each being treated as autonomous units according to Brazil's 1988 Federal Constitution. Municipalities are the smallest national administrative unit in the federation. There are around 5,000 municipalities distributed among the 26 states. Each municipality has a local government composed of an executive, led by a mayor, and a legislative body. Its representatives are elected simultaneously every four years. The elections for mayors and councilpersons (*vereadores*) are staggered by two years relative to state and federal elections.

Members of the municipal chamber are elected through an open-list proportional representation system. Each party or coalition must present a list of candidates before the elections. For the 2004 and 2008 elections, the electoral law[4] in its article 10 specified that each party is allowed to register candidates up to 150% of the number of seats to be filled in the elections. If parties decide to form a coalition, this number is increased to 200% and is valid for the coalition as a whole. As for the elections itself, each voter can cast a vote for either a party or a specific candidate. A vote to a candidate counts towards the number of votes the party has received and also influences the placement of the candidate chosen in the party list. Voting in Brazil is mandatory, but voters may also register an invalid vote.

After the elections, both votes for candidates and for parties are summed together to decide how many seats each party or coalition will receive. These seats are allocated according to the proportion of votes each party has received and the candidates to occupy these seats are the most-voted candidates in their respective list.

After the elections are over, the mandate begins in the following year. Elected candidates are able to influence public policy in various ways. The elected body is tasked with elaborating and voting the city's organic law, including the municipality budget. They are also tasked with supervising the mayor and may make public service requests to the mayor.

1.3.2 The number of seats in local councils

Before 2004, the legislative house had more flexibility in deciding the number of seats for the next election. They just had to respect the caps defined in the article 29 of the federal constitution, which stated that municipalities with up to one million inhabitants could have between 9 and 21 seats. For populations between one and five million inhabitants, the size can vary between 33 and 41 seats. For five million or higher, the caps were set between 42 and 55 seats. This rule was overridden in 2004 by the Supreme Electoral Court (*Supremo Tribunal Eleitoral - STE*), which issued a resolution[5] that removed from the legislative houses the power to define their own number of seats and instead defined the precise number of seats based on population thresholds. This change applied to the 2004 elections and was later reaffirmed for the 2008 elections. Municipalities with up to 47,619 inhabitants now may only have 9 legislators, but it goes up to 10 for municipalities between 47,620 and 95,238 inhabitants. The full table with the new population thresholds and their respective number of seats can be seen in Table 1. The comparison before and after can be seen in Figure 1, which shows the number of municipalities with each seat. In 2000 this number was varied even at municipalities with a small number of inhabitants, however this change and in 2004 and 2008 we only have municipalities with 9 seats

[4] Law Number 9,504/1997.
[5] Resolution n°21,702/2004.

before the first cutoff and 10 after. That means that for many municipalities the change implied a decrease in the number of seats over time.

Table 1 – The number of local legislators - 2004 and 2008 elections

Population intervals			No of seats	No of obs.	Population intervals			No of seats	No of obs.
0	to	47619	9	9935	1609757	to	1731707	38	2
47620	to	95238	10	623	1731708	to	1853658	39	1
95239	to	142857	11	184	1853659	to	1975609	40	0
142858	to	190476	12	82	1975610	to	4999999	41	6
190477	to	238095	13	64	5000000	to	5119047	42	0
238096	to	285714	14	37	5119048	to	5238094	43	0
285715	to	333333	15	32	5238095	to	5357141	44	0
333334	to	380952	16	25	5357142	to	5476188	45	0
380953	to	428571	17	17	5476189	to	5595235	46	0
428572	to	476190	18	9	5595236	to	5714282	47	0
476191	to	523809	19	10	5714283	to	5833329	48	0
523810	to	571428	20	8	5833330	to	5952376	49	0
571429	to	1000000	21	30	5952377	to	6071423	50	1
1000001	to	1121952	33	2	6071424	to	6190470	51	1
1121953	to	1243903	34	3	6190471	to	6309517	52	0
1243904	to	1365854	35	2	6309518	to	6428564	53	0
1365855	to	1487805	36	4	6428565	to	6547611	54	0
1487806	to	1609756	37	2	6547612	or	more	55	2

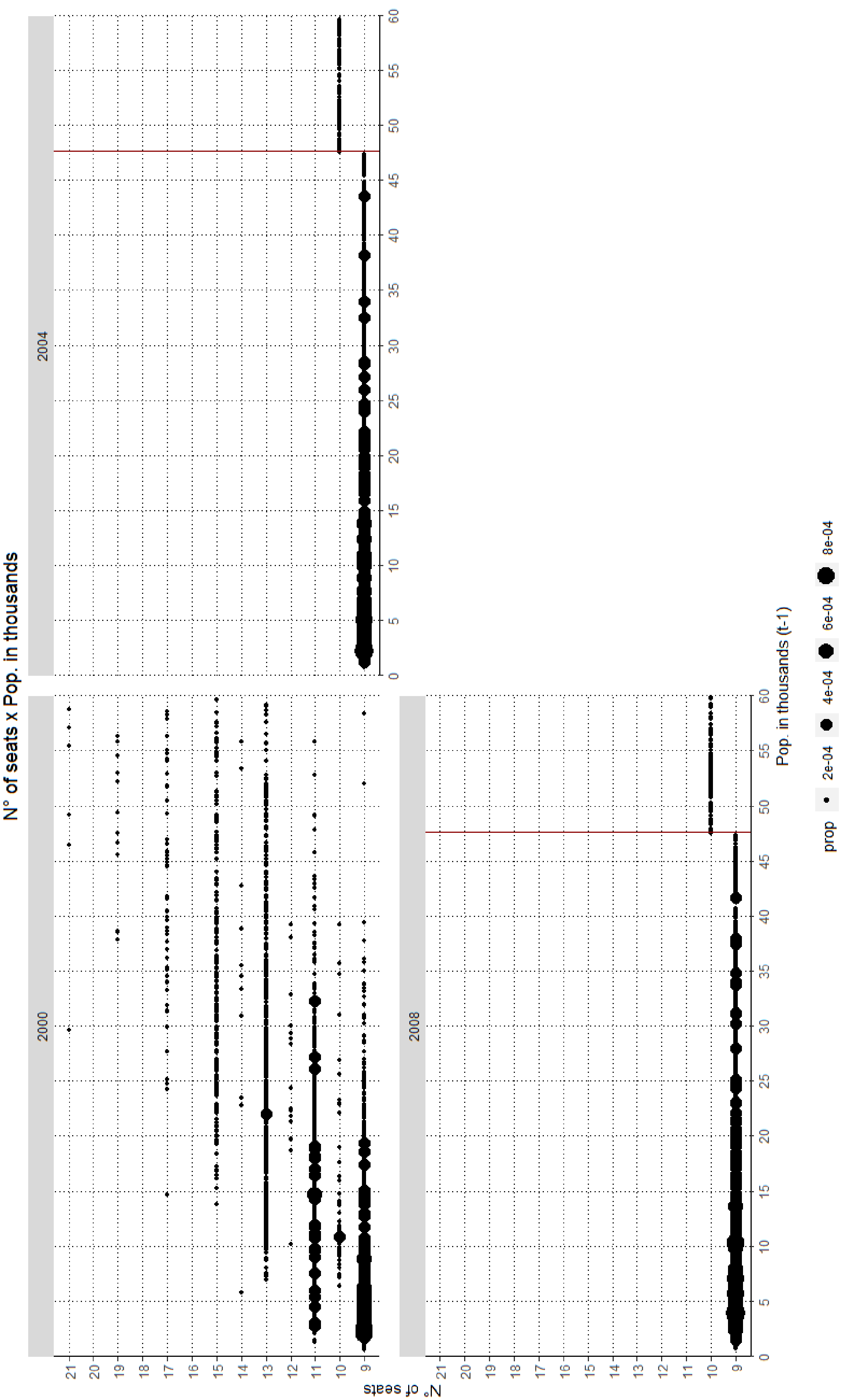
Source: STE Resolution n°21.702/2004.

Notes: This table exposes the rule defining the legislature size in Brazilian municipalities for the 2004 and 2008 elections. As it can be noticed, the number of legislators varies according to 36 intervals in the municipalities' population. Alongside with this information, we present the number of seats and number of municipalities (number of observations) in each strata (pooled for 2004 and 2008).

The number of seats in local councils in the 2004 and 2008 elections was defined based on the populations of, respectively, 2003 and 2007. For 2003, they were estimated using the 2000 Census. For 2007, there was a Population Count. Both the estimates and Population Counts are performed by an independent federal institution: the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatítica - IBGE*). In total, the rule defines 36 population intervals with their respective council size. However, 95% of municipalities have less than 95 thousand inhabitants, which means most municipalities are inside the interval defined for the first cutoff. In fact, even 95 thousand is a number too big. In Figure 2 we plot the distribution of municipalities for different populations, which shows that a large portion of the municipalities has even less than 20 thousand inhabitants.

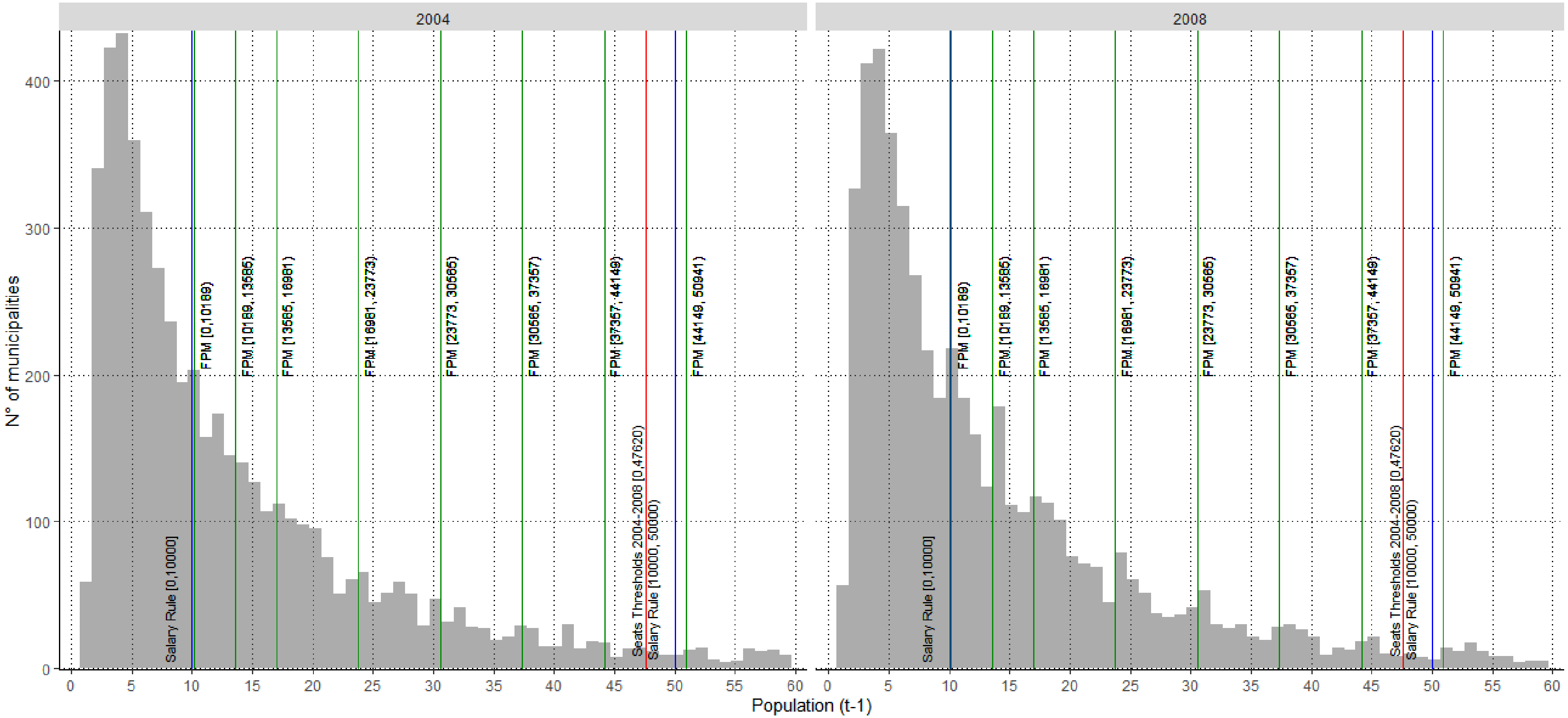
In 2009, the Constitutional Amendment No. 58/2009 further changed the rule determining the number of seats in local councils. Now, instead of an exact number of seats, the new rule went back to setting caps on the maximum number of seats allowed for each threshold, giving back to the legislative houses the power to define their own number of seats, at least to some

Figure 1 – Number of Seats x Population (in thousands)



Notes: This figure displays the number of municipalities in each seat and their population. The number of seats is on the y-axis and the population of the previous year is on the x-axis. The size of each point represents the proportion of municipalities in that point.

Figure 2 – Distribution of municipalities and rules based on population thresholds



Notes: This figure displays the distribution of municipalities ordered by their population in the previous year. We also display the thresholds for the number of seats in municipal legislative houses (red), for the salary of councilpersons (blue) and for the coefficients of the Fundo de Participação dos Municipios (FPM) transfers (green).

extent. It is important to notice, however, that new population thresholds came with this change. Municipalities with up to 10 thousand inhabitants could have only up to 9 seats. Municipalities between 10 and 15 thousand could have up to 11 seats, and so on. Because of this change, we cannot use the elections of 2012, 2016 and 2018 for this study. Although it is possible to examine these new thresholds, we are dealing with a fundamentally different experiment from the ones in 2004 and 2008, since both the population thresholds and the discontinuities in the number of seats changed afterwards. For the 2012 elections and after, the new maximum limits allowed municipalities to have two additional seats instead of one, as they move from one threshold to another.

These changes make it harder to properly identify our parameter of interest. Not only the characteristics of the municipalities around the thresholds are different, but there is also an overlap between the new rule for the number of seats and other rules based on population thresholds. 10 thousand inhabitants is also one of the thresholds for the salary of legislators, as defined by the Federal Constitution - Table 2 displays the population thresholds for these caps. At 50 thousand inhabitants there is another overlap between the new rule and one of the thresholds for the coefficients of the Municipality Participation Fund (*Fundo de Participação dos Municípios - FPM*) - Table 3 displays the thresholds for the FPM coefficients. Because of these issues, we opted to work only with the elections of 2004 and 2008, as this allows for a cleaner experiment.

Table 2 – Caps on local legislators' salary

Population intervals			Cap (% of state legislators salary)	No of seats	No of obs
0	to	10000	20%	9	5280
10001	to	47619	30%	9	4655
47620	to	50000	30%	10	45
50001	to	95238	40%	10	578
95239	to	100000	40%	11+	34
100001	to	300000	50%	11+	342
300001	to	500000	60%	11+	80
Above		500000	75%	11+	68

Notes: this table exposes the rule defining caps on the salary of local legislators (Constitutional Amendment No. 25, 2000). The caps are defined as a percentage of state legislators' salary, which are defined as a percentage of the salary of federal deputies. Alongside with this information, we present the number of seats and number of municipalities (number of observations) in each strata (pooled for 2004 and 2008).

Table 3 – Fundo de Participação dos Municípios - FPM (excluding state capitals) coefficients for each population interval

Population intervals			Coef.	No of seats	No of obs
0	to	10188	0.6	9	5327
10189	to	13584	0.8	9	1175
13585	to	16980	1.0	9	870
16981	to	23772	1.2	9	1143
23773	to	30564	1.4	9	666
30565	to	37356	1.6	9	396
37357	to	44148	1.8	9	267
44149	to	47619	2.0	9	91
47620	to	50940	2.0	10	59
50941	to	61128	2.2	10	187
61129	to	71316	2.4	10	152
71317	to	81504	2.6	10	117
81505	to	91692	2.8	10	77
91693	to	101880	3.0	10/11	80
101881	to	115464	3.2	11	63
115465	to	129048	3.4	11	40
129049	to	142632	3.6	11	32
142633	to	156216	3.8	11	30
Above		156216	4.0	11+	310

Source: Decree Law n°1,881/1981.

Notes: This table exposes the rule defining the coefficients used to determine the amount of federal transfers through the Municipal Participation Fund (FPM) to municipalities excluding the state capitals. For the number of observations in each interval, we pooled 2004 and 2008.

1.3.3 Women's Political Representation in Brazil

Much like in the rest of the world, women are underrepresented in politics in Brazil, and this is particularly true for municipal legislative houses. To amend this situation, the electoral law of 1995 (Law No. 9,100/1995) established that a minimum of 20% of vacancies on party lists should be filled with women. In 1997 this minimum requirement was revisited and changed to a quota that established that party lists should contain a minimum of 30% and a maximum of 70% of candidates of each gender[6].

Nevertheless, the law specified that the proportion of seats reserved for women should be calculated based on the total number of candidacies that the party or coalition may come to launch by electoral constituency instead of the actual number of candidates. What that means is that if a party was meant to have, for example, 100 vacancies, they should ensure that 30 of these vacancies should be reserved for women, but there was no need to actually fill these vacancies. As such, the actual number of candidates could be lower than that if they were not able to fill all the vacancies. This allowed parties to get away with having a percentage of female candidates much lower than the required by the quota. In 2004 and 2008, for example, the number of female candidates varied around 20% across all municipalities.

That said, this number has improved a bit over the years. In 2012 the number of female candidates had finally come closer to the minimum of 30%. The proportion of female candidates in 2012 was 28.32%. This may have been the result of Law No. 12,034/2009, which changed the wording of the law to consider the percentage of actual candidates. As such, parties were obligated to fill 30% of their vacancies with women.

Despite all these efforts, the number of women that actually managed to get elected remained and remains low. In 2004 and 2008 the proportion of women elected to legislative houses across all municipalities was around 12.5%. The changes in 2009 did not change this scenario. In 2012 the proportion of women elected was only 13.56% in our data. Only in 2015 further efforts were put to encourage female candidacies, when Law o. 13,165/15 established that parties and coalitions should increase the resources devoted to female candidacies.

In summary, for this paper, which analyzes the elections of 2004 and 2008, the legislative houses and elections were very male-dominated. As previously said, only around 20% of candidates were women and only around 12.5% of elected candidates were women.

1.4 Data and Estimation Strategy

1.4.1 Electoral and Municipal Data

In this paper, we use data from multiple sources. Data related to candidates was obtained from Superior Electoral Court (*Tribunal Superior Eleitoral - TSE*). This dataset includes information on the gender of candidates, how many votes they were able to obtain,

[6] Law No. 9,504/1997

their party affiliation, whether they managed to get elected or not, and more. We collect data for the 2004 and 2008 elections. However, an issue with datasets from TSE is that they are not perfectly compatible across years, requiring pre-processing corrections. The Center of Policy and Economics of the Public Sector (CEPESP) from the Getulio Vargas Foundation (FGV) provides access to TSE data with these corrections. As such, we use their data[7].

The population thresholds were defined based on the municipal populations of 2003 and 2007 for, respectively, the elections in 2004 and 2008. The information on population is not calculated by the municipalities themselves. Instead, the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística - IBGE*), an independent federal institute, estimates and reports the population for each municipality. These estimates are based on census data, which is performed every 10 years, and on population counts performed between census years. For 2003, the population was estimated based on the 2000 Census. For 2007, there was a Population Count. Both the Census and Population Counts are performed by the IBGE. Alongside with population data, we rely on census data for our balance tests.

We are also interested in examining variables associated with the municipality's prenatal health, infant education and fiscal outcomes during the mandates following the elections. We focus our attention on the third year of mandate after elections. As such, we gather data from 2007 and 2012 for the mandates following the elections of 2004 and 2008, respectively. Fiscal data comes from the National Secretary of Treasury (*Secretaria do Tesouro Nacional - STN*). They make this information available through its dataset called FINBRA (*Finanças do Brasil - Dados Contábeis dos Municípios*). This dataset contains information from declarations received by the National Treasury, which includes annual consolidated data on municipal spending and revenues for all Brazilian municipalities. We use the information on total and current expenditures, as well as information on the composition of spending for the legislature, community assistance, social assistance and child education.

Data on prenatal health outcomes come from the Data Processing Department of the Unified Health System (DATASUS) from the Ministry of Health. We examine data on females hospitalized due to aggression, number of live births, underweight live births (under 2500g), premature live births (under 37 weeks), women with four or more prenatal visits, fetal deaths and infant deaths. The information on hospitalizations due to aggression are provided by the Hospital Information System of the Unified Health System (*Sistema de Informações Hospitalares do Sistema Único de Saúde – SIH/SUS*). Information about live births are from the Information System on Live Births (*Sistema de Informações sobre Nascidos Vivos - SINASC*). Information on infant and fetal mortality come from the Information System on Mortality (*Sistema de Informações sobre Mortalidade - SIM*).

Data on educational outcomes are focused on the early years of education. We gather data from two sources: we use data on educational quality from the Basic Education Assessment

[7]These pre-processing corrections consist mostly of standardizations of certain variables (e.g. adjusting the description of categorical variables to make the datasets compatible across years. The procedures employed by CEPESP can be seen in more detail on their github page: https://github.com/Cepesp-Fgv/tse-dados/wiki.

System (*Sistema de Avaliação da Educação Básica - SAEB*), which is composed of a set of external standardized examinations in mathematics and portuguese. Results from SAEB and school pass rates obtained from the Annual School Census are used to calculate the Municipal Index of Basic Education Development (*Índice de Desenvolvimento da Educação Básica - Ideb*), which is an index of educational quality. As such, for our measures of educational quality in early years (up to 5^{th} grade) we look at the Ideb and the scores in mathematics and portuguese from the SAEB. Moreover, we also collect data on children's enrollment in daycare centers from the Annual School Census. All this information is collected and processed by the National Institute for Educational Studies and Research "Anísio Teixeira" (*Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - INEP*), a special research agency linked to the Ministry of Education.

1.4.2 Identification Strategy and Estimation

The main goal of this paper is to study the relationship between legislature size, feminine electoral representation and their impacts on public policy. The first step is to establish the relationship between the size of local legislatures and variables associated with female participation, such as the number of women elected. However, the electoral success and the decision of candidates to participate in an election are influenced by many other political factors. These other variables may hinder our ability to correctly identify the impact that the number of seats in a given legislature might have on the political representation of women. To avoid this endogeneity problem, we exploit an exogenous change in the number of seats due to Resolution n°21.702/2004, issued by the Supreme Electoral Court, which set the number of seats according to population thresholds. For this paper, we only focus on the first threshold. As such, municipalities below 47,620 inhabitants have 9 seats and those above that and up to 95,238 have 10 seats. This allows us to adopt a sharp regression discontinuity design[8].

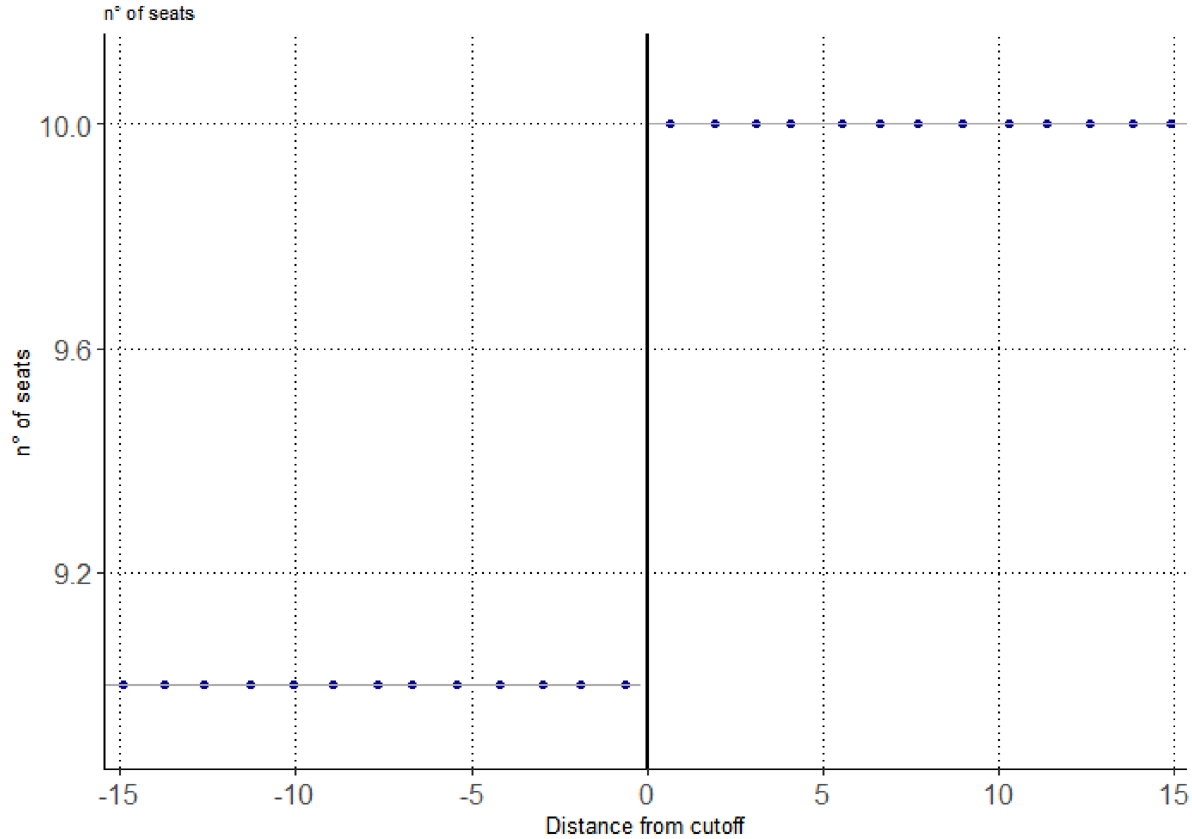
In Figure 3 we can see that this rule completely determines the number of seats for all municipalities below and above the specified threshold in our sample[9]. As such, we have a Sharp design for our Regression Discontinuity approach. In Sharp Regression Discontinuity Designs, the assignment mechanism is assumed to be a deterministic function of the running variable (in this case, the population).

This may be better explained within the context of the counterfactual framework of Rubin (1974). Let our treatment be having one more seat in the legislative house. Then $Y_i(1)$ and $Y_i(0)$ are, respectively, the potential outcome for our variable of interest in municipality i when the municipality has this extra seat and when it does not. We are interested in estimating

[8]See Imbens and Lemieux (2008) and Lee and Lemieux (2010) for an overview of Regression Discontinuity Designs.

[9]Actually, 0.3% of municipalities in our sample did not follow the rule established by Resolution n°21.02/2004. These municipalities violated the law and were prosecuted by electoral authorities, having to adjust the number of seats after the elections. However, due to the extremely low number of deviant cases, we decide to remove these abnormal cases from our sample and proceed with a Sharp RD Design instead of employing a Fuzzy design.

Figure 3 – Number of seats around the threshold (2004 & 2008 pooled)



Notes: Graphical presentation of the RD design. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. We followed (LEE; LEMIEUX, 2010) to construct the bins.

the quantity

$$\tau = E[Y_i(1) - Y_i(0)] \tag{1.13}$$

However, we cannot observe this variable, since we observe only one of these outcomes for each municipality. More specifically, let $T_i \in \{0, 1\}$ be a variable that determines whether the municipality i received the treatment or not. If it received, then $T_i = 1$. Otherwise, the value is zero. Then we only observe $Y_i(1)|T_i = 1$ and $Y_i(0)|T_i = 0$. In the context of a Sharp Regression Discontinuity Design, T_i is believed to be exogenous around the threshold. Let c be the threshold and P our *running variable*. Then

$$T_i = 1(P_i \geq c) \tag{1.14}$$

Where $1(\cdot)$ is an indicator function. This is our treatment variable - a dummy equal to one when a municipality is above the population threshold and 0 otherwise. In our case, $c = 47,619$. If the expectations $E[Y_i(1)|P_i]$ and $E[Y_i(0)|P_i]$ are continuous, then

$$\lim_{c\downarrow 0} E[Y_i|P_i = c] - \lim_{c\uparrow 0} E[Y_i|P_i = c] = E[Y_i(1) - Y_i(0)|P = c] = \tilde{\tau} \quad (1.15)$$

This is our Local Average Treatment Effect (LATE) we are interested in. Notice, however, that this value is not the same as the value we wanted to estimate in Equation 1.13. This is because we are only able to estimate the treatment effect for values around the threshold.

The most straightforward way to estimate the LATE would be to calculate the averages of the variable of interest above and below the cutoff for observations close enough to this cutoff. Then our LATE is simply the difference between both means. This non-parametric approach is akin to running a Nadaraya–Watson (local constant) kernel regression using a rectangular kernel. Kernel regressions are well suited for estimating the regression function at a particular point, such as in this case. However, in a RD setting, the cutoff represents a boundary point and kernel regressions are known to display a systematic bias at boundaries. This bias is related to the curvature of the relationship between our variable of interest, Y, and the running variable, X. One way to reduce the importance of the bias is to include a slope in the kernel regression (local linear regression)[10]. Calonico, Cattaneo and Titiunik (2014) also implements a bias-correction inference procedure for the confidence intervals[11]. As such, for our specifications, we include a simple comparison of means accompanied by more robust estimates using the local linear estimator. For both specifications, we use a rectangular kernel and we also perform the bias-correction mentioned above.

Another concern is bandwidth choice. If we restrict our sample too much, we lose statistical power. However, allowing a too large bandwidth might compromise our ability to correctly identify the parameter of interest. As such, we consider bandwidth choices of 12.5%, 25% and 50% around the cutoff. We also include a data-driven bandwidth that minimizes the Mean Squared Error (CALONICO; CATTANEO; TITIUNIK, 2014). This bandwidth selector is based on a mean squared error (MSE) expansion of the sharp RD estimators, leading to a choice of bandwidth that minimizes the MSE (hence MSE-optimal). This is given by:

$$\begin{aligned} h_{\text{MSE},p} &= C_{\text{MSE},p} n^{-\frac{1}{2p+3}} \\ C_{\text{MSE},p} &= \left(\frac{V_p}{2(p+1)B_p^2}\right)^{\frac{1}{2p+3}} \end{aligned} \quad (1.16)$$

Where p denotes the polynomial order (0 for mean comparison, 1 for linear local), and B_p and V_p are, respectively, the leading asymptotic bias and variance of the RD estimator.

This leaves us with five main specifications for our estimates: mean comparison at around 12.5% of the cutoff; local linear regressions at around 12.5%, 25% and 50%; and local linear regression with MSE-optimal bandwidth. All these specifications are interested in estimating the

[10] See Hansen (2021) Chapters 18-19 for an overview on non-parametric estimation.

[11] Calonico, Cattaneo and Titiunik (2015) presents a package for the Software R which implements the command rdrobust, which implements the bias-corrected robust (to large bandwidths choices) procedures to Confidence Intervals and Bandwidth Selection proposed by Calonico, Cattaneo and Titiunik (2014).

following regression:

$$Y_{i,t} = f(P^c_{i,t-1}) + \tau T_{i,t} + \varepsilon_{i,t} \\ \forall P_{i,t-1} \in [c-h, c+h] \tag{1.17}$$

Where i indexes the municipalities; t are the election years (2004, 2008); $P^c_{i,t-1}$ is the population level in the previous year centered around the cutoff; $f(\cdot)$ represents the RD-polynomial function to capture the relationship between our variable of interest and the running variable; and h is the bandwidth, chosen according to the criteria described previously. For the local constant estimator, we simply ignore this term. For the local linear estimator, we include a linear slope.

As we focus only on the first threshold, our sample is restricted to only municipalities between 10 thousand and 95,238 inhabitants. We pooled together observations from 2004 and 2008. As such, for all specifications, we cluster the standard errors at municipal level.

Nonetheless, there are still two possible concerns related to our identification strategy and estimation: the presence of other rules based on population thresholds[12] and the possibility of auto-selection around the threshold (that is, the concern that municipalities might be able to manipulate their population in order to self-select into the group with 9 or 10 seats).

The latter concern, regarding auto-selection around the threshold, is not a major concern. The population of each municipality is estimated by the Brazilian Institute of Geography and Statistic (*Instituto Brasileiro de Geografia e Estatística - IBGE*), which is an independent federal agency linked to the Ministry of the Economy. Moreover, the electoral authorities only declared the rule valid in the 2004 and 2008 elections in their respective year, but the population estimates were calculated a year before. As such, it is unlikely that municipalities could have known the rule that defined their council sizes before the publication of the population estimates. Nonetheless, we examine the possibility of manipulation by performing the McCrary Test (McCrary, 2008), which examines the discontinuity in the distribution of observations around the threshold of interest. Table 4 and Figure 4 presents the estimates and plots of the density of the population of municipalities in our sample. As it can be seen, we find no evidence of manipulation. Another exercise that we can do to strengthen this result is to show that both treated and controls are similar in observables. In the next section, we check the covariate balance for our sample.

As for the concern regarding other rules that use population thresholds, there are two other rules based on population thresholds that could be a concern for us. Tables 2 and 3 presents, respectively, the caps of the wages of local legislators and the coefficients used for the allocation of federal transfers through the Municipal Participation Fund (FPM) to municipalities excluding state capitals. Both use population thresholds and have cutoffs close to ours. In the case of the legislator's salary, there is a cutoff at 50 thousand inhabitants, which is very close to the cutoff being studied here. This leads the grants received by each municipality to be continuous around the thresholds. Nonetheless, we address these concerns in Section 6, by conducting sensitivity analysis and false thresholds checks as robustness checks. There, we vary the size of our bandwidths, which allows us to check if we can still identify our results for bandwidths

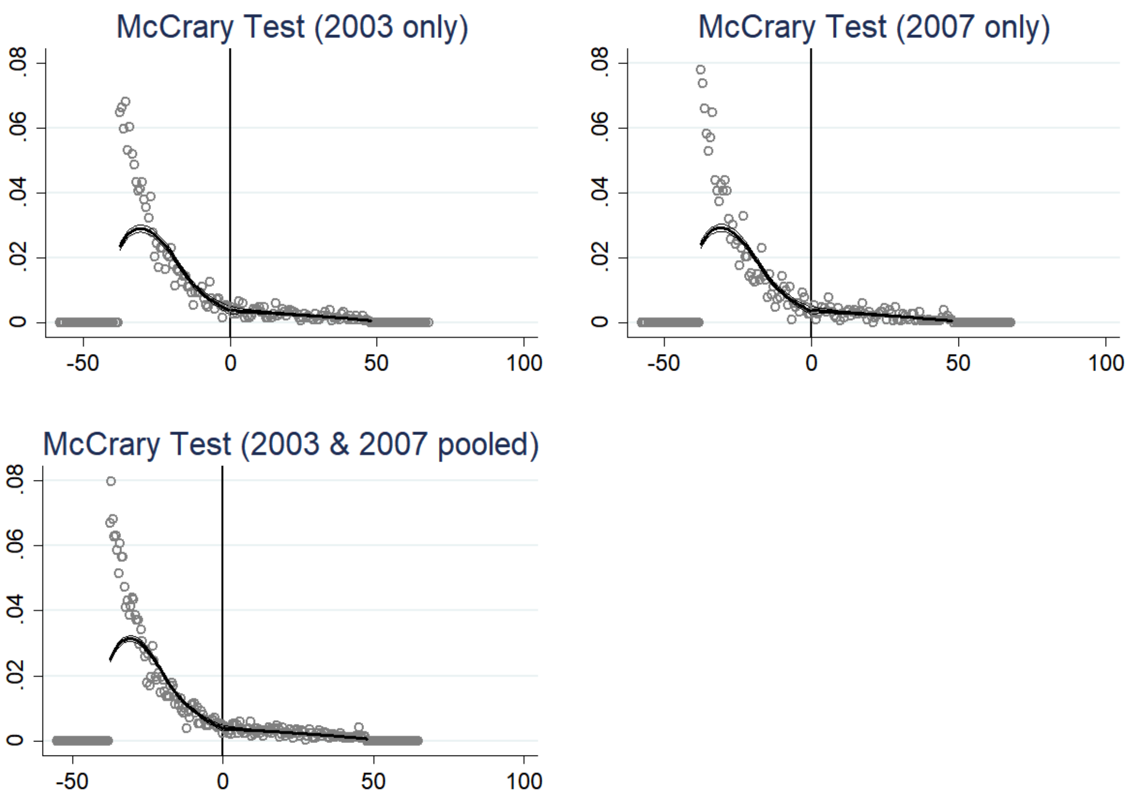
[12]See Eggers et al. (2018) for a discussion on how multiple rules involving the same running variable might invalidate the RD Design and proposed workarounds.

Table 4 – McCrary Test estimates

Discontinuity Estimate			
Pop. Year	2003	2007	2003 & 2007
log diff. in height	-0.035	0.090	0.048
s.d.	(0.238)	(0.258)	(0.163)
t-stat	-0.149	0.349	0.297
p-value	0.882	0.727	0.767
Bandwidth	15.595	14.149	17.500
Bin Size	0.700	0.687	0.490

Notes: This table is the results for the McCrary test (McCrary, 2008). It estimates the difference between the distribution of the running variable (population) at each side of the cutoff of 47,620. The first two columns refer to the elections of 2004 and 2008, respectively, and the last column pools both elections.

Figure 4 – McCrary Test



Notes: These figures are the graphic representation for the McCrary test (McCrary, 2008). It shows the distribution of the population centered around the cutoff of 47,620 for the municipalities in our sample. Each density is estimated separately, which then allows us to identify whether there are any discontinuities in the running variable at the cutoff. The first two figures refer to the elections of 2004 and 2008, respectively, and the last figure pools both elections.

that do not include these other thresholds. And from the false thresholds tests, we can examine whether any effect we find is actually located around the threshold we are examining. We find no evidence that other thresholds could be driving our results.

In the next section, we present some covariate balance tests and then follow up with our main results. Given the concerns discussed here and proposed solutions, we are confident that our empirical strategy indeed captures the parameter of interest and allows us to estimate the causal effects of increasing the number of seats in Brazil's municipal legislative houses from 9 to 10.

1.4.3 Covariate Balance and Validity Checks

The main assumption of our identifying strategy is that municipalities around the cutoff of 47,620 inhabitants did not get to choose which side of the threshold they got to be. To them, being on either side is random. If this is true, we should expect the municipalities around the threshold to be similar in observable characteristics and past variables. To test this, we selected a set of municipal characteristics from the 2000 Census. We also looked at the behavior of the electoral results for the elections in 2000. However, because we are using only data from 2000, we separated our sample and estimated the regressions for 2004 and 2008 separately in order to avoid counting some municipalities twice.

The results can be found in Tables 5 and 6 for, respectively, the years of 2004 and 2008. There are some signs of imbalance, especially for the Avg. Monthly Nominal Salary in 2008, Gini in both years and Herfindahl–Hirschman Index (HHI) for race[13] and schooling for both years. We also noticed many variables are significant for the Linear Local 12.5% BW specification. However, it is also worth noticing the low number of observations and lack of consistency when we vary the bandwidth.

To further examine the balance or lack thereof of these covariates, we performed the Romano-Wolf correction to control for the familywise error rate (FWER)[14]. The idea here is that since we tested many variables, it may be possible that these few significant variables are just false positives. We plotted the Romano-Wolf p-values in Figure 5. The Romano-Wolf p-values are referred by 'rwolf', while the p-values from the regressions are the 'standard' p-values. For both local constant and linear local specifications, we compare municipalities around 25% of the cutoff of 47,620 and we do not perform any bias-correction. As it can be seen, there are more significant variables in the linear constant specification than in the Tables 5 and 6, and even the Romano-Wolf p-values remain significant for many of these variables. However, in our main

[13]Brazil has many ethnic groups. To capture the concentration of one ethnic group over the others, we calculated a Herfindahl–Hirschman Index using the proportion of individuals of each ethnic group in the municipality. The Brazilian Institute of Geography and Statistics (IBGE) defines 5 ethnic groups in the census: *branco* (white), *preto* (black), *amarelo* (yellow, meaning East Asians), and *indígena* (indigenous person, meaning Amerindians).

[14]See Romano and Wolf (2005b), Romano and Wolf (2005a) and Romano and Wolf (2016) for details on the procedure. Clarke, Romano and Wolf (2020) presents an implementation of this procedure for the Stata software

Table 5 – Balance Test for 2004

		Mean Comparison		Local Linear								
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW	
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
N° of Candidates (2000)	98.861	129	11.8056+	286	0.238	7.1216	129	0.1586	306	3.9336	814	13.79
		129	(7.5225)	286	0.238	(13.4084)	129	(22.3708)	306	(16.4701)	814	(11.4481)
N° of Male Cand. (2000)	80.556	129	9.4269+	282	0.235	5.7365	129	-2.0835	306	2.6325	814	10.305
		129	(5.8928)	282	0.235	(10.5131)	129	(17.235)	306	(12.698)	814	(8.9151)
N° of Female Cand. (2000)	18.194	129	2.4898	269	0.224	1.8535	129	2.4391	306	1.4445	814	3.6147
		129	(1.8695)	269	0.224	(3.3402)	129	(5.7544)	306	(4.242)	814	(2.8791)
N° of Females Elected (2000)	1.236	129	0.2551	335	0.280	0.3587	129	1.4275***	306	0.2919	814	0.2834
		129	(0.2001)	335	0.280	(0.3073)	129	(0.4719)	306	(0.4238)	814	(0.2929)
N° of Males Elected (2000)	12.556	129	0.5673	303	0.248	-0.0825	129	-3.0339***	306	-0.6737	814	-0.3125
		129	(0.4433)	303	0.248	(0.7209)	129	(1.1051)	306	(0.8646)	814	(0.6104)
% Legislature Fem. (2000)	9.271	129	1.0559	410	0.326	1.6648	129	10.85***	306	1.6937	814	1.481
		129	(1.4983)	410	0.326	(2.2223)	129	(3.6179)	306	(3.2547)	814	(2.2177)
Elected at Least One Fem. (2000)	0.694	129	0.0599	331	0.276	0.1419	129	0.8237***	306	0.3722**	814	0.1355
		129	(0.0793)	331	0.276	(0.119)	129	(0.1977)	306	(0.1544)	814	(0.1137)
Avg. Monthly Nominal Salary	413.637	142	15.7756	270	0.213	39.7725	142	50.8937	333	31.2873	909	35.0269
		142	(27.2767)	270	0.213	(51.599)	142	(90.6885)	333	(61.0482)	909	(40.534)
Gini	0.581	142	-0.0228**	251	0.202	-0.036**	142	-0.049*	333	-0.0377*	909	-0.0332**
		142	(0.0096)	251	0.202	(0.018)	142	(0.026)	333	(0.0204)	909	(0.0142)
% Households in Urban Areas	71.299	142	2.4617	449	0.325	-1.4016	142	9.913	333	11.5936*	909	0.1034
		142	(3.2861)	449	0.325	(4.7779)	142	(9.7881)	333	(6.8893)	909	(4.6926)
% Women in Municipality	49.926	142	0.1349	334	0.251	-0.1697	142	-0.0693	333	0.2558	909	-0.1359
		142	(0.1776)	334	0.251	(0.286)	142	(0.5256)	333	(0.3706)	909	(0.2555)
Avg. Age in Municipality	26.956	142	0.7602*	475	0.333	0.2644	142	-0.4797	333	0.44	909	0.4324
		142	(0.4443)	475	0.333	(0.6425)	142	(1.2859)	333	(0.9132)	909	(0.6512)
HHI Race in Municipality	5395.029	142	433.1652*	451	0.325	525.2563+	142	1205.1936**	333	483.3204	909	529.3874+
		142	(234.2964)	451	0.325	(332.4075)	142	(575.213)	333	(442.7191)	909	(324.9046)
HHI Schooling in Municipality	1869.214	142	99.0671+	323	0.245	208.2716*	142	326.0041*	333	94.0982	909	161.0157*
		142	(66.8283)	323	0.245	(108.6308)	142	(173.7722)	333	(137.2561)	909	(97.0078)
Avg. Years of Schooling	4.957	142	0.3016+	398	0.302	0.3331	142	0.5233	333	0.5072	909	0.3349
(Population)		142	(0.1843)	398	0.302	(0.2805)	142	(0.5772)	333	(0.3994)	909	(0.2706)
Avg. Years of Schooling	4.742	142	0.3057+	393	0.295	0.3171	142	0.594	333	0.5176	909	0.3782
(Male Population)		142	(0.1983)	393	0.295	(0.3001)	142	(0.6197)	333	(0.4246)	909	(0.289)
Avg. Years of Schooling	5.166	142	0.2961*	406	0.306	0.2946	142	0.4657	333	0.4975	909	0.2957
(Female Population)		142	(0.1731)	406	0.306	(0.2635)	142	(0.5417)	333	(0.3793)	909	(0.2559)
Polynomial Order			Zero			One		One		One		One

Notes: In this table we report our preferred RDD estimates for the balance tests using variables from the 2000 Census and electoral variables from the 2004 elections. Columns (1)-(5) reports our estimates. The *Mean* is the average before the cutoff point for municipalities around 12.5% of the cutoff. Column (1) reports a mean comparison around 12.5% of the cutoff. Tables (2) to (5) reports the Local Linear estimates for different bandwidth. We follow the procedure by *Calonico et al. (2014)* to correct the bias in non-parametric estimates and to calculate the MSE-optimal bandwidth. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6 – Balance Test for 2008

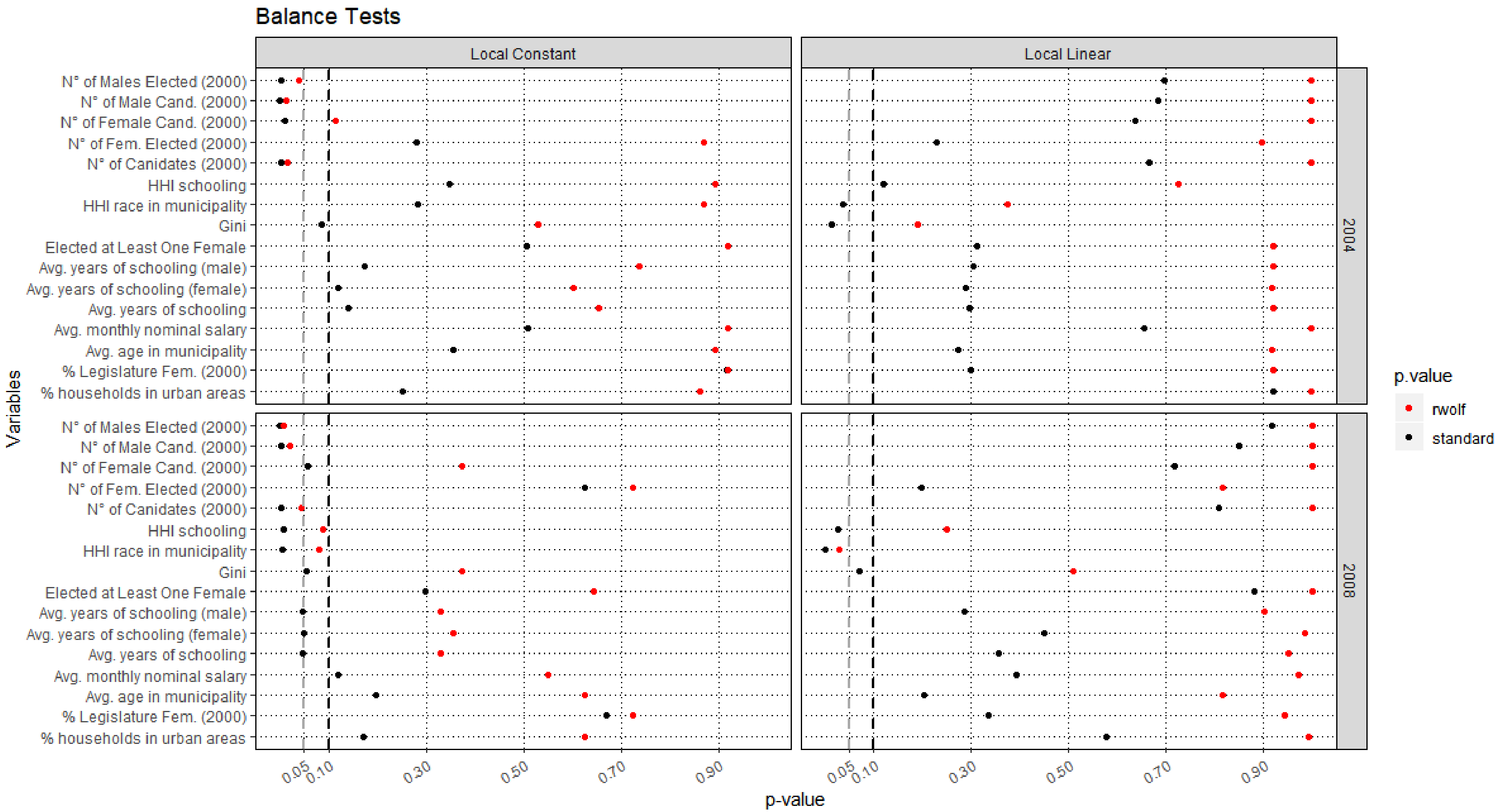
		Mean Comparison		Local Linear									
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW		
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
N° of Candidates (2000)	100.500	138	13.2344*	287	0.247	1.2293	138	9.0412	290	14.7344	808	10.4384
		138	(7.2171)	287	0.247	(14.1701)	138	(22.401)	290	(17.5742)	808	(11.8182)
N° of Male Cand. (2000)	81.608	138	10.9231*	290	0.251	0.6506	138	5.5735	290	10.8016	808	8.4702
		138	(5.6588)	290	0.251	(10.8545)	138	(17.142)	290	(13.432)	808	(9.1409)
N° of Female Cand. (2000)	18.811	138	2.3611	283	0.243	0.9956	138	3.3999	290	4.0303	808	2.032
		138	(1.812)	283	0.243	(3.4007)	138	(6.043)	290	(4.6396)	808	(3.0417)
N° of Females Elected (2000)	1.297	138	0.2965	267	0.228	0.3457	138	1.0625*	290	-0.0024	808	0.3029
		138	(0.2061)	267	0.228	(0.3467)	138	(0.6093)	290	(0.4394)	808	(0.3034)
N° of Males Elected (2000)	12.351	138	0.6486+	288	0.249	0.0872	138	0.2993	290	0.7271	808	0.242
		138	(0.4115)	288	0.249	(0.7657)	138	(1.3978)	290	(0.9783)	808	(0.6478)
% Legislature Fem. (2000)	9.824	138	1.2319	263	0.226	1.664	138	7.4669+	290	-0.6556	808	1.5102
		138	(1.5139)	263	0.226	(2.6415)	138	(4.7047)	290	(3.2562)	808	(2.2208)
Elected at Least One Fem. (2000)	0.770	138	-0.0203	377	0.307	-0.022	138	0.3063	290	-0.0586	808	-0.0437
		138	(0.0735)	377	0.307	(0.1158)	138	(0.288)	290	(0.1806)	808	(0.115)
Avg. Monthly Nominal Salary	397.404	149	46.5059*	278	0.215	64.2409	149	208.0841**	321	151.4973**	901	49.6809
		149	(25.3921)	278	0.215	(48.2113)	149	(83.9802)	321	(59.8117)	901	(40.8326)
Gini	0.582	149	-0.0167*	349	0.272	-0.0317*	149	-0.0804**	321	-0.026	901	-0.0268*
		149	(0.009)	349	0.272	(0.017)	149	(0.0384)	321	(0.0255)	901	(0.0155)
% Households in Urban Areas	72.297	149	0.6621	323	0.253	-3.1304	149	13.2023	321	7.3067	901	-1.1764
		149	(3.3477)	323	0.253	(6.5976)	149	(11.9415)	321	(8.6834)	901	(5.6577)
% Women in Municipality	50.043	149	-0.0501	418	0.307	-0.0462	149	-0.4898	321	-0.0723	901	-0.1078
		149	(0.1597)	418	0.307	(0.2965)	149	(0.7137)	321	(0.4536)	901	(0.2806)
Avg. Age in Municipality	27.082	149	0.6332+	323	0.255	0.8712	149	2.1782	321	2.1663**	901	0.8076
		149	(0.4284)	323	0.255	(0.8075)	149	(1.6646)	321	(1.0975)	901	(0.7161)
HHI Race in Municipality	5,160.689	149	740.4002***	426	0.316	1021.2074***	149	1723.0723***	321	1536.4172***	901	936.9607***
		149	(209.6706)	426	0.316	(318.7212)	149	(558.3098)	321	(447.8344)	901	(318.6208)
HHI Schooling in Municipality	1,812.103	149	172.8426***	354	0.273	252.2217**	149	482.751**	321	294.5911**	901	261.7268***
		149	(65.8737)	354	0.273	(108.1682)	149	(194.5533)	321	(140.2547)	901	(99.7955)
Avg. Years of Schooling	4.987	149	0.3075*	295	0.227	0.3185	149	1.3017*	321	1.0058**	901	0.3936
(Population)		149	(0.1831)	295	0.227	(0.3491)	149	(0.6698)	321	(0.4556)	901	(0.2996)
Avg. Years of Schooling	4.755	149	0.348*	281	0.218	0.4709	149	1.4625**	321	1.0948**	901	0.4673+
(Male Population)		149	(0.1949)	281	0.218	(0.3707)	149	(0.6978)	321	(0.478)	901	(0.3159)
Avg. Years of Schooling	5.213	149	0.2687+	312	0.242	0.2525	149	1.1509*	321	0.9208**	901	0.3226
(Female Population)		149	(0.1739)	312	0.242	(0.3267)	149	(0.648)	321	(0.4378)	901	(0.2867)
Polynomial Order			Zero			One		One		One		One

Notes: In this table we report our preferred RDD estimates for the balance tests using variables from the 2000 Census and electoral variables from the 2008 elections. Columns (1)-(5) reports our estimates. The *Mean* is the average before the cutoff point for municipalities around 12.5% of the cutoff. Column (1) reports a mean comparison around 12.5% of the cutoff. Tables (2) to (5) reports the Local Linear estimates for different bandwidth. We follow the procedure by *Calonico et al. (2014)* to correct the bias in non-parametric estimates and to calculate the MSE-optimal bandwidth. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

specification - local linear, we find that most of these variables are not significant. And the ones remaining are not significant when we use the Romano-Wolf p-values, except for the HHI of race. This lack of balance, however, is only present in 2008.

As such, we believe there is a good balance between the two sides of the threshold. The significant results do not seem robust to different specifications nor when we control for the FWER. These results reinforce our confidence that we can identify our parameters of interest using this RD design - that is, that our main results, which are going to be presented in the next section, indeed stem from an exogenous increase in the number of seats

Figure 5 – Balance Statistics



Notes: In this figure, we display the balance p-values for municipality characteristics and candidate characteristics. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate two p-values. The standard p-value is calculated from the estimated regression model and the 'rwolf' p-value refers to p-values calculated with the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variables are the variables in the y-axis and the explanatory variable is a dummy variable for the cutoff of 47,620 inhabitants. Both local constant and linear regressions assume a 25% bandwidth around the cutoff point and we use a rectangular kernel. These estimates were not corrected with the bias-correction confidence intervals of (CALONICO; CATTANEO; TITIUNIK, 2014).

1.5 Main Results

1.5.1 Political Entry and Electoral Outcomes

In the theoretical framework we developed in Section 2, it was assumed that it is costlier for women to participate in an election, that some voters are biased against women and that there is a lack of coordination between voters. From these assumptions, we drew out two testable empirical predictions. The first one is that an increase in the number of seats in a legislative house should have a bigger effect on the entry of male candidates compared to the number of female candidates. The second prediction, however, states that this increase in seats has a bigger effect on the proportion of elected female candidates compared to the proportion of elected male candidates. We now test these predictions using a Regression Discontinuity approach.

In Figure 6 we selected variables associated with political entry and electoral outcomes. The point-estimates for different specifications can be seen in Table 7. We also include the average value of each variable to the left of the cutoff for a bandwidth of 12.5%. By visual inspection of the graphs, there seems to be a slight increase in the number of candidates, particularly male candidates. However, there is also a decrease in their average votes and the Herfindahl–Hirschman Index (HHI). Women, on the other hand, do not experience any noticeable increase in entry, but the overall number of elected women increases as well as the proportion of women in the legislature.

The regressions partially confirm what was shown in the graphs. In the mean comparison, only the average votes for male candidates do not behave as we had seen in the graphs, displaying a positive increase, but the estimates are negative for the other specifications and none of them is significant. As for the Local Linear specifications, the increase in the number of candidates is only significant at 15% for the 50% bandwidth. The increase in the number of male candidates is similar, displaying significant results for the optimal bandwidth and at a 50% bandwidth. That said, in all cases the point-estimates display a similar behavior, with all point-estimates being positive and being similar in magnitude. As for the size of these effects, we can compare the averages around the cutoff. Going by the mean comparison estimates, there is around 12.78%[15] more candidates on the right side of the cutoff, and out of these 10.73 extra candidates, around 85% are men.

Despite the mean comparison displaying a positive increase in the average votes for male candidates, this result is not robust. In fact, none of the estimates are significant at any level. But the magnitude of the effects is mostly negative. That said, the HHI of vote for males is significant in the mean comparison and in the Local Linear specification with a 12.5% bandwidth. The results are not significant for other bandwidths, but the magnitudes all have the same sign, suggesting an overall increase in competitiveness among male candidates. This is in line with our prediction that an extra seat motivates party affiliates to run for office, but because the cost

[15] This is the marginal increase, 10.73, divided by the average before the cutoff point, indicated in the table (84.012). Both the average value before the cutoff and the marginal increase of the mean comparison use a 12.5% bandwidth.

of entry is higher for women and part of the electorate is biased against women, more men end up entering. This is also in line with the argument that women face higher barriers to entry (CASAS-ARCE; SAIZ, 2011; BHALOTRA; CLOTS-FIGUERAS; IYER, 2018).

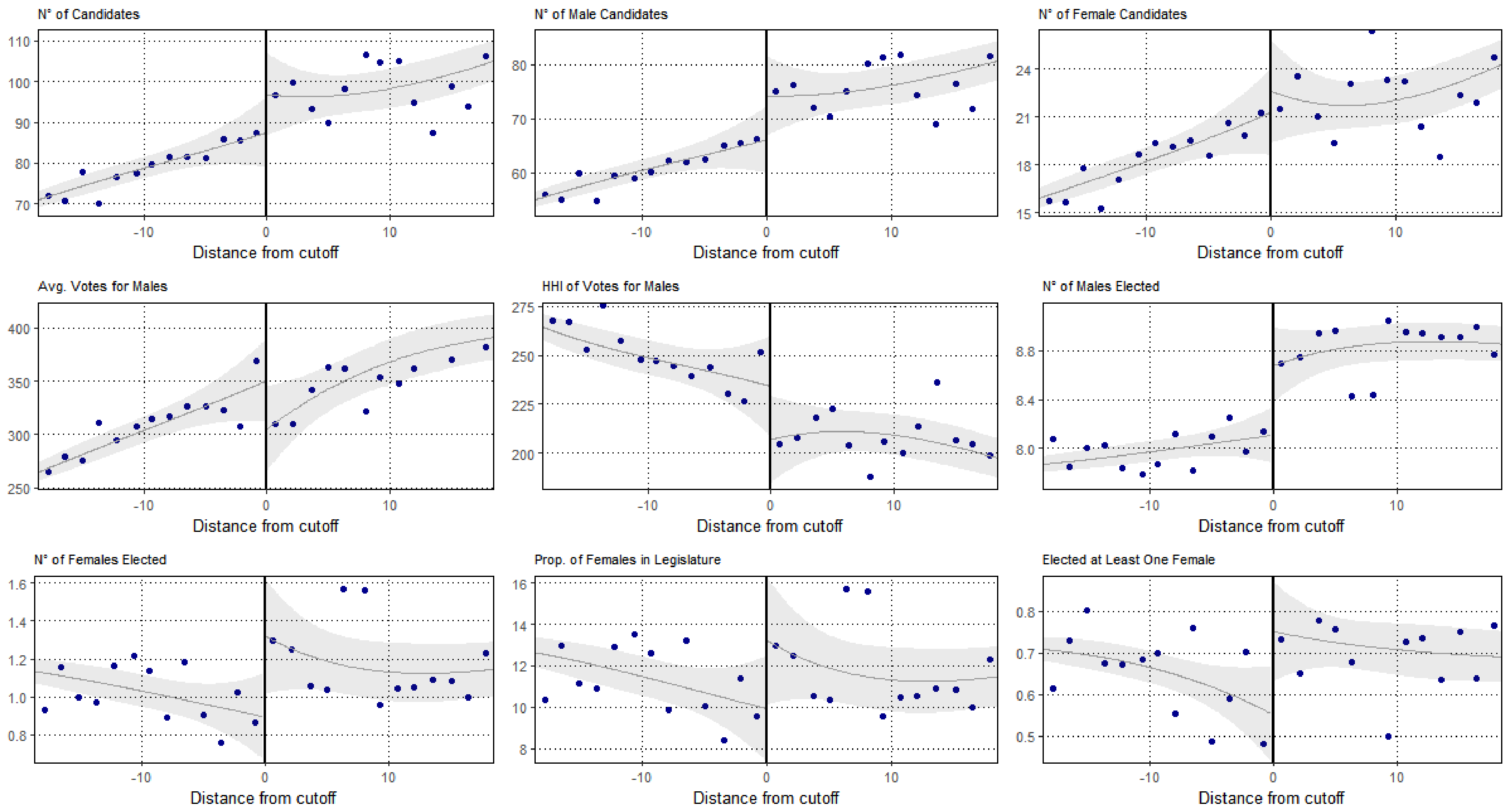
As for elected candidates, there is a noticeably increase in the number of male candidates being elected. The same is observed in female candidates for the mean comparison and for the Local Linear specification with 12.5% and 50% bandwidth. Although the magnitude of this increase is higher for men compared to women, women experience a bigger increase in relation to themselves. In the mean comparison, the number of elected male candidates is increased by only 8.8% in relation to the municipalities with 9 seats, but the increase for women is 33%. This also reflects in the probability of electing at least one female candidate, which is significant across all specifications. However, the increase in the share of women in the legislature is only significant in the mean comparison. The magnitude of the effect is positive for all specifications, though. Going by the mean comparison, an extra seat increases the percentage of women in the legislature from around 9.5% to 11.3%, almost a 20% increase in relation to the counterfactual. These results are still in line with idea that there is a cannibalization effect here. As more men enter the election, the greater the competition between them is. It doesn't affect women as much. Thus, even though the increase in the number of elected men is higher than women, the relative increase in the number of women elected is higher. This also increases the probability of a woman being elected. The RD graphs and comparison of averages suggest that this also results in a more equitable legislative house, but this result is sensitive to the bandwidth choice and regression specification.

1.5.2 Policy Outcomes

We just have seen that an increase in the number of seats increases women's representation in the legislative house. A follow-up question would be: what happens as a result of this increase in representation? To examine this question, we selected a set of variables associated with prenatal health care and child education, which are areas commonly thought to be associated with female preferences by the literature (MASON; KING, 2001; DUFLO, 2003; MILLER, 2008; SVALERYD, 2009; BHALOTRA; CLOTS-FIGUERAS, 2014; BROLLO; TROIANO, 2016).

We first start with the health outcome variables, for which our results are displayed in Figure 7 and Table 8. The RD plots seem to display a discontinuity in the number of live births per 1000 inhabitants, the proportion of fetal deaths and the proportion of infant deaths. The birth rate and proportion of fetal and infant deaths are all negative and significant in the mean comparison and in the Local Linear specification with optimal bandwidth. The proportion of fetal deaths is also significant at 25% and 50% bandwidths, and the % infant deaths are significant in all specifications. The birth rate is decreased by around 1.1 children per 1000 inhabitants in the mean comparison and 2 in the Local Linear specification. This represents a 6.5% to 12% decrease in relation to the average before the cutoff. The effect on fetal and infant deaths have similar magnitude, with both representing almost a 12% decrease in relation to the left side in the man comparison. The point-estimates are reasonably bigger in the Local Linear

Figure 6 – Main Results - Political Variables



Notes: RD graph for our main results for the political variables. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 7 – Effects of Local Council Size on Political Variables

		Mean Comparison		Local Linear									
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW		
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
N° of	84.012	289	10.7338**	692	0.268	7.0882	289	10.1123	650	6.8892	1802	10.2575+
Candidates		289	(4.2656)	692	0.268	(7.525)	289	(14.9264)	650	(10.2574)	1802	(6.8387)
N° of Male	64.184	289	9.181***	819	0.306	7.4479+	289	8.2196	650	6.3096	1802	8.9709*
Candidates		289	(3.0863)	819	0.306	(4.9402)	289	(10.6101)	650	(7.3185)	1802	(4.8777)
N° of Female	19.828	289	1.5527	534	0.213	-0.2239	289	1.8927	650	0.5796	1802	1.2885
Candidates		289	(1.3233)	534	0.213	(2.7136)	289	(4.6482)	650	(3.2058)	1802	(2.1808)
Avg. Votes	333.386	289	2.6723	723	0.279	-27.5597	289	-81.8762	650	-36.9344	1802	-38.2603
for Males		289	(18.3368)	723	0.279	(31.9141)	289	(59.5275)	650	(42.0081)	1802	(28.3529)
HHI of Votes	240.962	289	-26.86**	781	0.295	-22.8098	289	-73.3347*	650	-31.9453	1802	-23.2729
for Males		289	(10.9642)	781	0.295	(18.1328)	289	(38.7993)	650	(25.6779)	1802	(17.3374)
N° of Males	8.147	289	0.7178***	480	0.195	0.645***	289	0.3871	650	0.6768**	1802	0.686***
Elected		289	(0.1172)	480	0.195	(0.2467)	289	(0.3845)	650	(0.2832)	1802	(0.1958)
N° of Females	0.853	289	0.2822**	480	0.195	0.355	289	0.6129+	650	0.3232	1802	0.314+
Elected		289	(0.1172)	480	0.195	(0.2467)	289	(0.3845)	650	(0.2832)	1802	(0.1958)
% Legislature Fem.	9.475	289	1.8741+	480	0.195	2.6975	289	5.351	650	2.3245	1800	2.1788
		289	(1.2275)	480	0.195	(2.5776)	289	(4.0949)	650	(2.9776)	1800	(2.034)
Elected at Least	0.558	289	0.1719***	512	0.205	0.1726+	289	0.3658**	650	0.2075+	1802	0.1629*
One Female		289	(0.0562)	512	0.205	(0.1097)	289	(0.1832)	650	(0.1276)	1802	(0.0871)
Polynomial			Zero		One			One		One		One

Notes: In this table, we report our preferred RDD estimates for the effect of the council size on political variables associated with the entry of citizens to politics and its electoral results. We use municipalities as observations. Columns (1)-(4) reports our estimates. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW.* reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis.
$^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

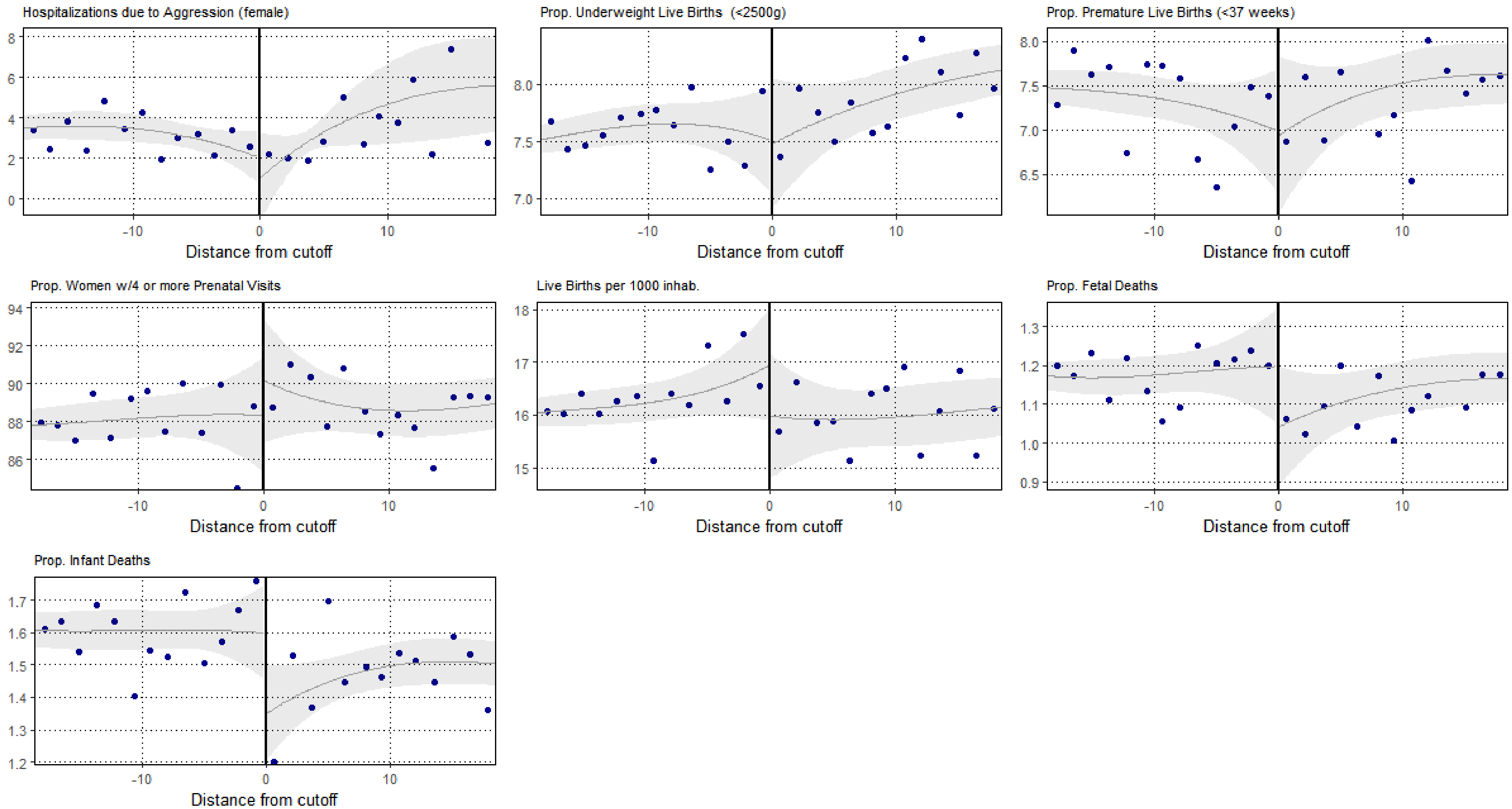
specification for infant deaths.

Figure 8 and Table 9 reports the results for child education. We find very noticeable improvements in all variables. The SAEB score for early years of education (up to 5th grade) is increased by 5.5 for mathematics and 4.3 for portuguese in the man comparison, which represents a 2.9% and 2.5% increase in relation to the average before the cutoff. For the Local Linear specification these effects are even stronger in magnitude, varying between 12 and 26 points for mathematics and 8.9 and 21 points for portuguese. As expected, this also translates to the average SAEB score and the Ideb score. In the mean comparison, there is a 3.8% and 6.31% increase in relation to the average before the cutoff. As in the previous case, the point-estimates are generally higher in the Local Linear specification. As for the enrollment in childcare, there is an increase of 1.5 children by 1000 inhabitants in the mean comparison (15% of the mean) and between 2.73 and 7.67 in the Local Linear estimations. These results are in accordance with much of the literature that relates to women's empowerment and child education, which sees that female politicians are more likely to emphasize early childhood education (SVALERYD, 2009; CLOTS-FIGUERAS, 2012).

Now we turn to fiscal variables to examine how the extra seat and higher female representation could have an effect on government size and expenditure composition. Our results are displayed in Table 10 and Figure 9. We find no change in total and current superavit. An extra seat, and consequently a higher percentage of women, does not seem to affect the local government's revenues and spending. It is also interesting to notice that the legislative expenditures also do not change as a result of an extra seat. The only significant results are at 15% for the mean comparison for the Total Superavit and Legislature Expenditures, but none of the estimates for the Local Linear specification are significant. This lack of effect could be explained by the Fiscal Responsibility Law (Complementary Law No. 101/2000), which limits expenditures with employees to 6% of net current revenue for the legislative house in municipalities. That said, the point-estimates are positive, despite not being significant.

Despite not having any effects on overall spending and revenues, the higher female representation in the legislative houses seems to have an effect on the composition of government expenditures. Likewise, the results for expenditure associated with community and social assistance are not robust. However, we find that municipalities above the cutoff spend more on child education. In the mean comparison, we find that an extra seats lead to more 0.9% expenditures in child education as a proportion of the current expenses, which represents an increase of 27.8% over the average before the cutoff. In the Local Linear specification, this value is even higher, with estimates that vary between 1.19% and 3.46%. This result resonates with much of the literature that states that women are more likely than men to invest in children (THOMAS; WELCH, 1991; LOTT; KENNY, 1999; ALESINA; FERRARA, 2005) and that this reflects into policy decisions when they are better represented (BESLEY; CASE, 2003; CHATTOPADHYAY; DUFLO, 2004; BHALOTRA; CLOTS-FIGUERAS, 2014). It also corroborates with and helps to explain our previous findings regarding educational outcomes.

Figure 7 – Main Results - Policy Outcomes (health)



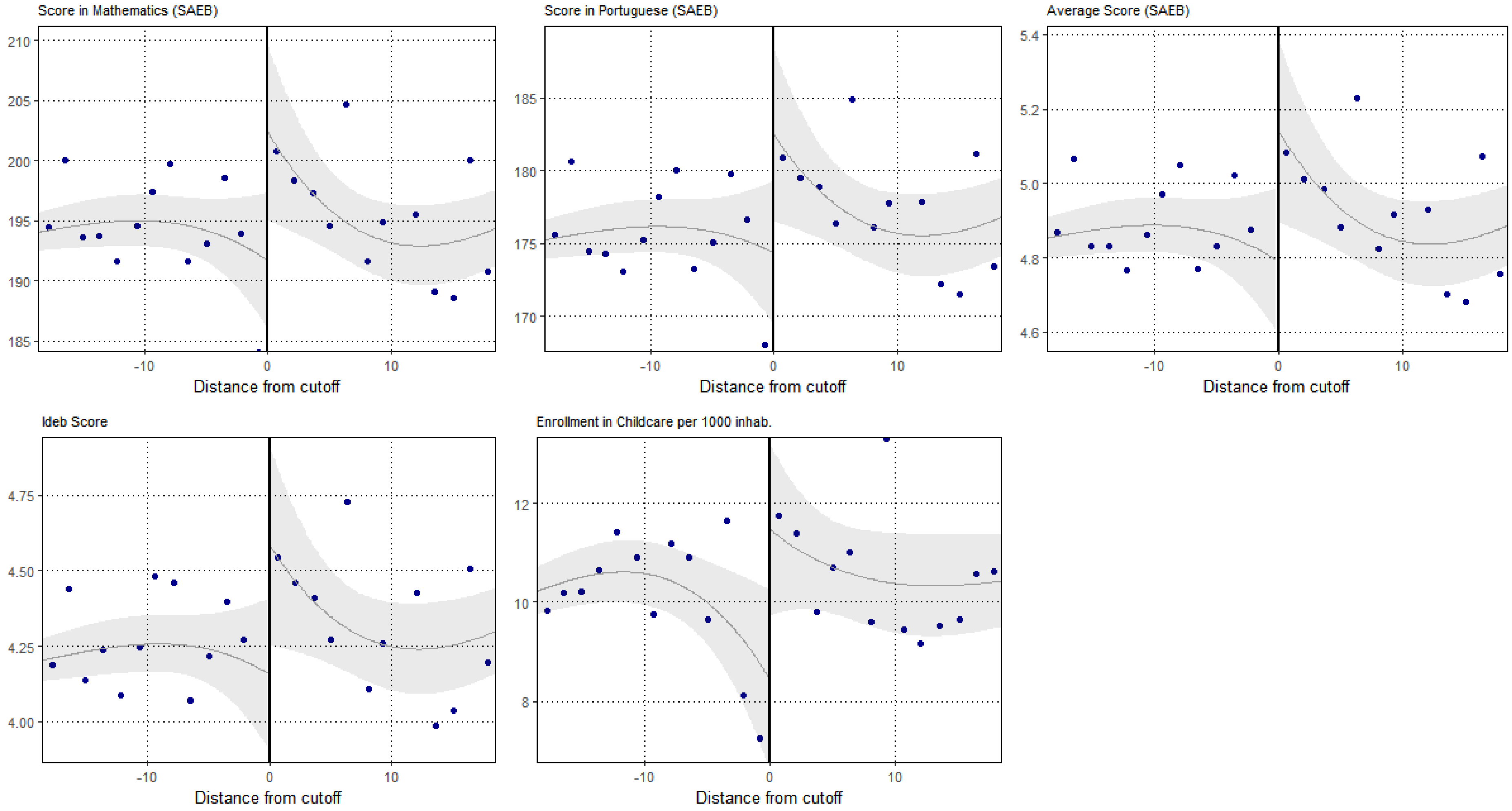
Notes: RD graph for our main results for the health outcomes. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 8 – Effects of Local Council Size on Political Variables - Health Outcomes

		Mean Comparison		Local Linear								
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW	
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
Hospitalizations	2.718	178	-0.4918	214	0.150	-1.3552	178	0.3424	368	-0.7075	879	-1.6681+
due to Aggression (fem.)		178	(0.4548)	214	0.150	(1.2454)	178	(1.3355)	368	(1.0834)	879	(1.067)
% Underweight	7.469	292	0.2085	748	0.286	-0.1239	292	-0.3275	655	0.1569	1812	-0.1036
Live Births (<2500g)		292	(0.2439)	748	0.286	(0.4119)	292	(0.6758)	655	(0.5375)	1812	(0.3865)
% Premature	6.983	292	0.3141	638	0.244	0.3988	292	-0.1388	655	-0.7335	1812	0.2411
Live Births (<37 weeks)		292	(0.3607)	638	0.244	(0.709)	292	(1.2871)	655	(0.9332)	1812	(0.5954)
% Women w/4 or	87.964	292	1.8081	455	0.189	2.5187	292	2.5762	655	3.8319	1812	2.5921
more Prenatal Visits		292	(1.3713)	455	0.189	(3.0375)	292	(4.1876)	655	(3.2717)	1812	(2.2837)
N° of Live Births	16.874	292	-1.1036**	559	0.217	-2.097**	292	-0.8961	655	-1.4331	1812	-1.1245
per Capita		292	(0.5329)	559	0.217	(0.9615)	292	(1.4549)	655	(1.1041)	1812	(0.8188)
% Fetal Deaths	1.226	292	-0.147**	584	0.225	-0.2282*	292	-0.3166	654	-0.255+	1799	-0.1731+
		292	(0.0696)	584	0.225	(0.1321)	292	(0.2586)	654	(0.1635)	1799	(0.1104)
% Infant Deaths	1.616	292	-0.1925***	300	0.129	-0.6088***	292	-0.6675***	654	-0.5313***	1808	-0.3077***
		292	(0.0682)	300	0.129	(0.1441)	292	(0.1914)	654	(0.1428)	1808	(0.1064)
Polynomial			Zero		One			One		One		One

Notes: In this table, we report our preferred RDD estimates for the effect of the council size on policy variables associated with health outcomes. Columns (1)-(4) reports our estimates. We use municipalities as observations. Sample sizes in columns (3) and (4) for Hospitalization due to Aggression (Female), Prop. of Fetal Deaths and Infant Deaths are, respectively, 368 and 879, 655 and 1801 and 655 and 1810. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW.* reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 8 – Main Results - Policy Outcomes (education)



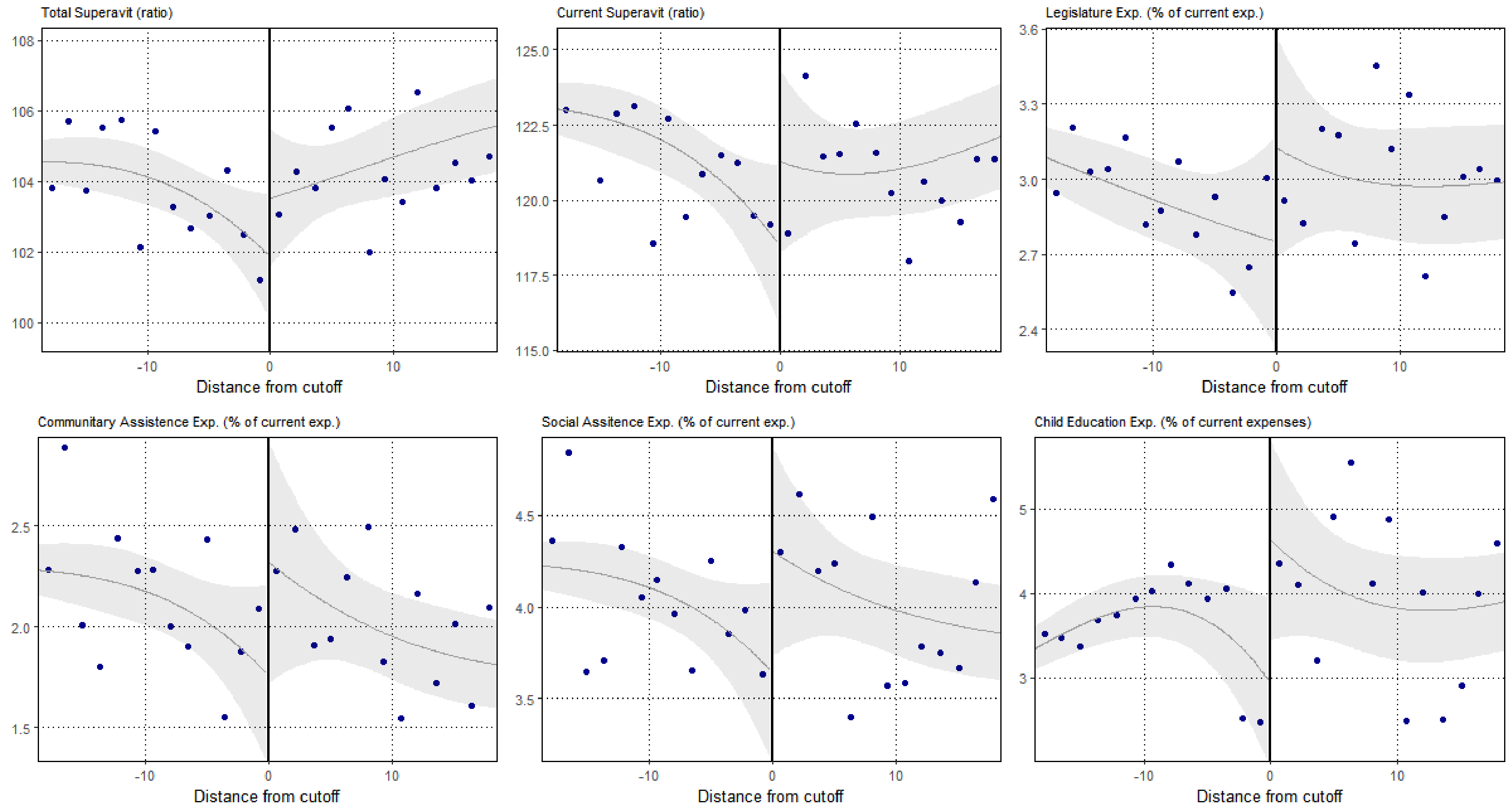
Notes: RD graph for our main results for the educational outcomes. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 9 – Effects of Local Council Size on Political Variables - Educational Outcomes

		Mean Comparison		Local Linear								
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW	
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
Score in	193.281	288	5.5382*	482	0.198	12.1205**	288	26.3765***	644	14.1721**	1772	12.0798**
Mathematics (SAEB)		288	(2.9404)	482	0.198	(5.9456)	288	(9.3576)	644	(6.988)	1772	(4.8278)
Score in	175.540	288	4.3222*	562	0.220	10.4073**	288	21.4862***	644	10.0441*	1772	8.8926**
Portuguese (SAEB)		288	(2.4881)	562	0.220	(4.7245)	288	(7.8181)	644	(5.8252)	1772	(4.0387)
Avg. Score	4.844	288	0.1843*	559	0.219	0.4456**	288	0.894***	644	0.4531*	1772	0.3922**
(SAEB, standarized)		288	(0.1009)	559	0.219	(0.1945)	288	(0.3195)	644	(0.2383)	1772	(0.1649)
Ideb Score	4.207	288	0.2654**	515	0.206	0.6016**	288	1.1402***	644	0.6699**	1772	0.4782**
		288	(0.1334)	515	0.206	(0.2563)	288	(0.4081)	644	(0.3084)	1772	(0.2182)
Enrollment in Childcare	9.608	292	1.4856*	410	0.175	4.1332***	292	7.6701***	655	4.8381***	1812	2.7346**
per 1000 inhab.		292	(0.845)	410	0.175	(1.5423)	292	(2.3344)	655	(1.7358)	1812	(1.2807)
Polynomial			Zero		One			One		One		One

Notes: In this table, we report our preferred RDD estimates for the effect of the council size on policy variables associated with educational outcomes. Columns (1)-(4) reports our estimates. Sample sizes in columns (3) and (4) for Enrollment in Childcare per Capita are, respectively, 656 and 1814. We use municipalities as observations. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW*. reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 9 – Main Results - Policy Outcomes (fiscal)



Notes: RD graph for our main results for the fiscal outcomes. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 10 – Effects of Local Council Size on Political Variables - Fiscal Outcomes

		Mean Comparison		Local Linear									
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW	
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
Total Superavit	103.052	286	1.3831+	603	0.237	1.9087	286	2.9597	644	1.472	1774	1.863
(ratio)		286	(0.8554)	603	0.237	(1.4659)	286	(2.7728)	644	(1.903)	1774	(1.3228)
Current Superavit	120.559	286	1.1268	410	0.177	-0.7078	286	1.2581	644	0.0035	1773	2.7072
(ratio)		286	(1.3312)	410	0.177	(2.8562)	286	(3.574)	644	(2.7109)	1773	(2.0677)
Legislature Expenditures	2.721	286	0.2919+	460	0.192	0.2821	286	-0.3908	644	0.084	1773	0.2948
(% of curr. expenses)		286	(0.1949)	460	0.192	(0.3768)	286	(0.5917)	644	(0.4239)	1773	(0.3091)
Communitary Assist. Exp.	1.930	286	0.2018	624	0.243	0.6791+	286	0.3332	644	0.3581	1773	0.4818
(% of curr. expenses)		286	(0.2441)	624	0.243	(0.4318)	286	(0.7095)	644	(0.5499)	1773	(0.3909)
Social assistance Exp.	3.946	286	0.3294	387	0.168	0.6871	286	0.7549	644	0.6309	1773	0.5659+
(% of curr. expenses)		286	(0.2297)	387	0.168	(0.5182)	286	(0.6948)	644	(0.5443)	1773	(0.389)
Child Education Exp.	3.357	286	0.9323+	416	0.178	1.6322+	286	3.4633***	644	2.0267**	1773	1.1984+
(% of curr. expenses)		286	(0.5661)	416	0.178	(1.0191)	286	(1.2751)	644	(1.0134)	1773	(0.795)
Polynomial			Zero		One			One		One		One

Notes: In this table, we report our preferred RDD estimates for the effect of the council size on policy variables associated with fiscal outcomes. Columns (1)-(4) reports our estimates. We use municipalities as observations. Sample sizes in columns (3) and (4) for Total Superavit and Current Superavit are, respectively, 451 and 1244, and 637 1764. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW*. reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

1.6 Robustness checks

We found that an extra seat leads to an increase in female representation in the municipal legislative house, as described by our theoretical framework, and this, in turn, leads to policies that cater to female preferences, such as prenatal health care and child education. Now we check if these claims are robust and indeed stem from our experiment. To validate our findings from the previous section, we perform two tests: we perform a sensitivity analysis for different bandwidths and a falsification test with false thresholds.

For each variable found significant in our main specification, we perform a sensitivity analysis in which we run various regressions for different bandwidths. The results for the political variables can be found in Figure 10. For each variable, we estimate Local Linear Regressions with CCT's robust bias-corrected confidence intervals and an uniform kernel. Each observation in the figures represents the point-estimate of a regression for a specific bandwidth. The vertical hashed line marks the optimal bandwidth for the corresponding regression.

The estimates for the number of candidates and the number of candidates and male candidates are not significant for most regressions, but there are still some significant results for larger bandwidths. Moreover, the point-estimates are consistently positive except for very low bandwidths, which at this point might suffer from a small sample size. On the other hand, the estimates for the number of female candidates are not significant and around zero. Although we cannot confirm the significance of the estimates for candidates and male candidates, the point-estimates corroborate with the argument that an increase in the number of candidates is mostly driven by male candidates. The average number of votes and HHI of vote for males are also not significant for most regressions, but some are and most of the point-estimates are well behaved and negative, as we would expect. As for the number of female candidates, proportion of women in the legislature and probability of electing at least one woman, all of these variables have mostly positive point-estimates and many of the regressions are significant, particularly so for higher bandwidths.

Moving on to Figures 11 and 12 we find the same exercise for the political outcome variables. The policy outcomes are mostly unchanged when we vary the bandwidth. The birth rate is only significant at larger bandwidths, but most estimates are negative. The same is true for the proportion of fetal deaths. Likewise, the expenditures with community and social assistance are not significant only at higher bandwidths, but the majority of point-estimates are similar in magnitude and positive. The remaining variables are very robust. They are all significant for most specifications and the sign of the estimates are in accordance with our main specification.

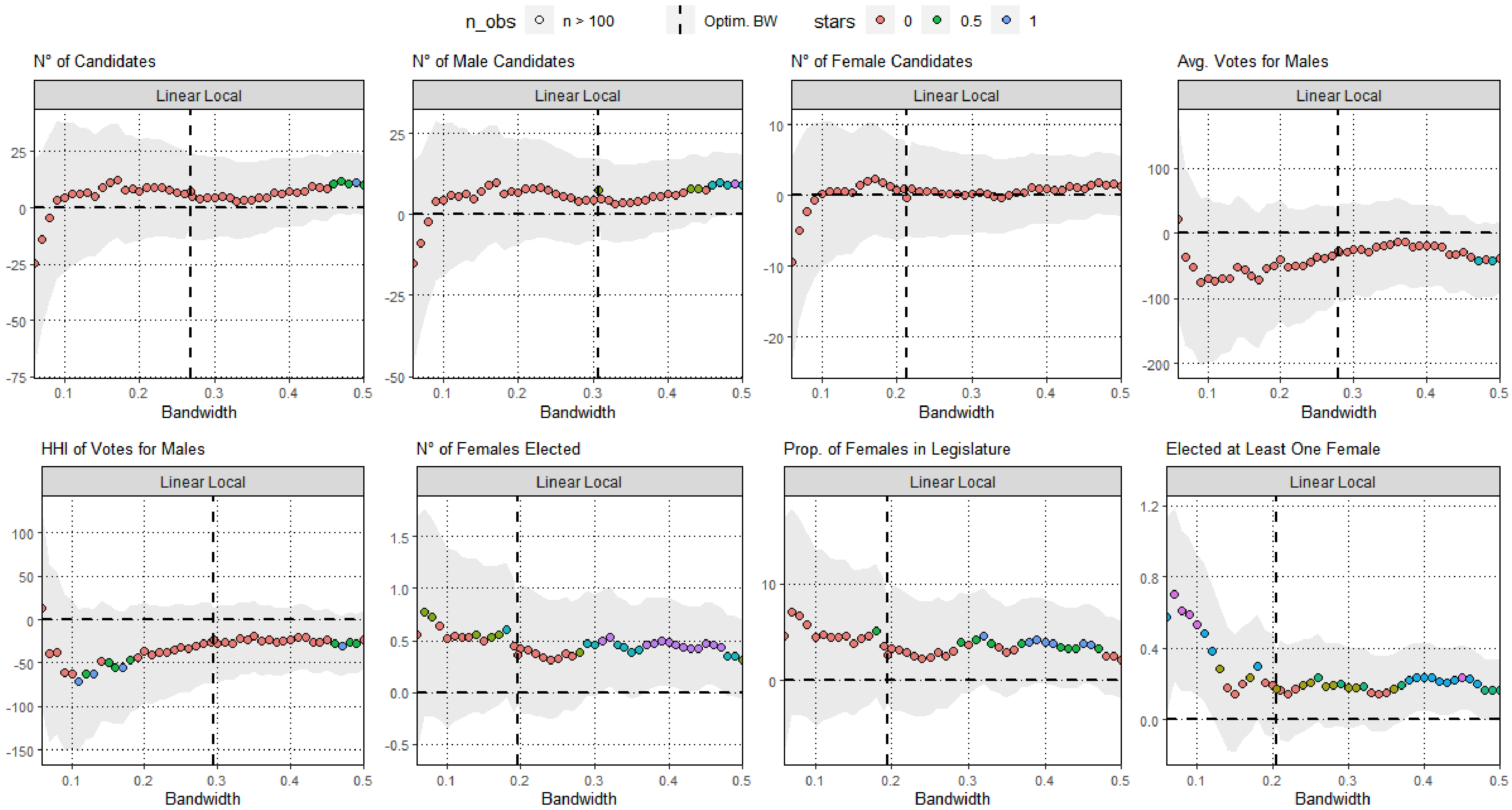
In summary, varying the bandwidths does not seem to change our results. We still find a cannibalization effect on most regressions and a subsequent increase in women's representation. The policy outcomes that follow are also largely unchanged when we vary the bandwidth.

Now we turn to the exercise with false thresholds. In this exercise, we perform several placebo regressions that assume different cutoff points. We consider false cutoff points between

90% and 110% of the true cutoff of 47,620. For each variable, we estimate Local Linear regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014). We estimate these regressions for a fixed bandwidth of 25% and a MSE-optimal bandwidth. This procedure allows us to check if our results are spurious and whether other rules based on population thresholds may be driving our results. Figures 13 and 14 presents the exercise for political entry electoral results variables and Figures 15, 16 and 17 for the policy outcomes. As it can be seen, the effect of the treatment effects are maximized at or very close to the true cutoff for all of our variables except for the number of female candidates, which is a variable we expected to be not have any effect anyway. In particular, the results for educational outcomes are very noticeable. To complement this result we plot the CDF of t-statistics for each variable and specification. Here we ran 500 regressions at discontinuities between 50% and 150% of the true threshold, using increments of 0.2 percentage points[16]. This exercise can be seen in Figures 18 to 22. The vertical hashed line represents the t-statistic of our main specification at the true cutoff. As it can be seen, this t-statistic is always around the extremes of the distribution. Therefore, both false thresholds exercises seem to corroborate that our results are only identified at the true cutoff.

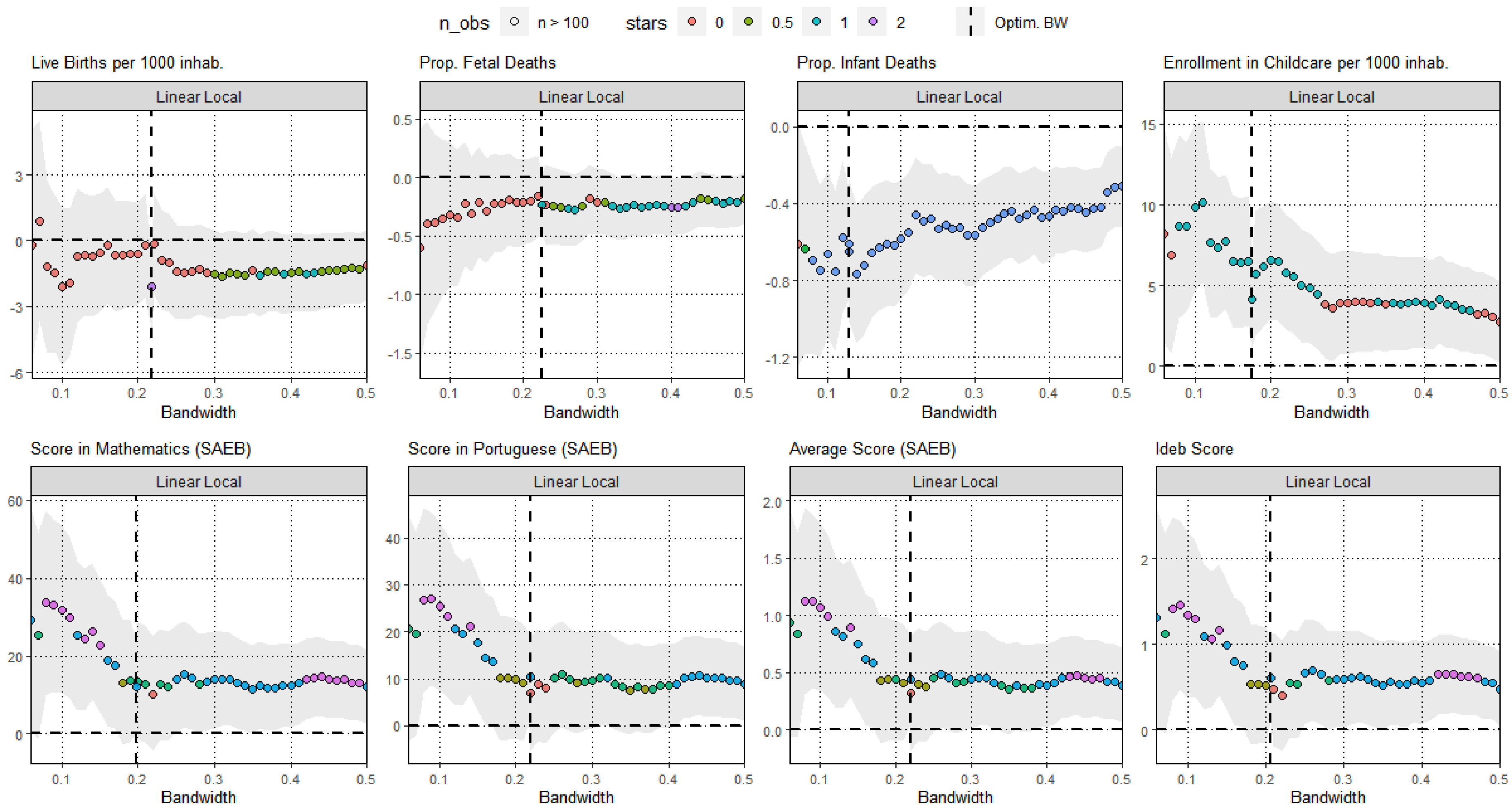
[16]This exercise was initially based on falsification tests by Vigna and Ferrara (2010).

Figure 10 – Sensitivity analysis - Political Variables



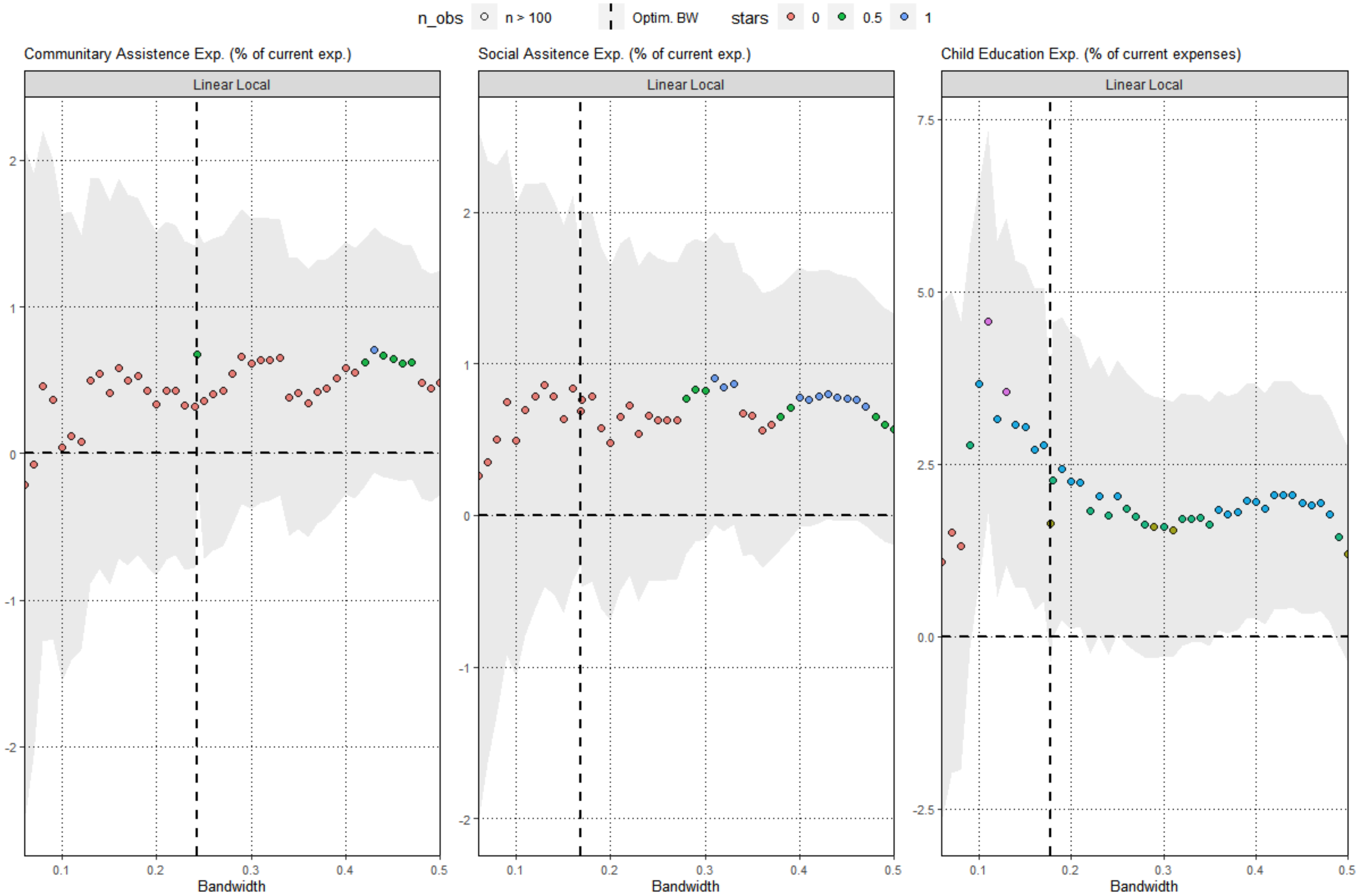
Notes: In this figure, we display a sensitivity analysis for our main RDD specification for political outcomes. We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

Figure 11 – Sensitivity analysis - Policy Outcomes (1/2)



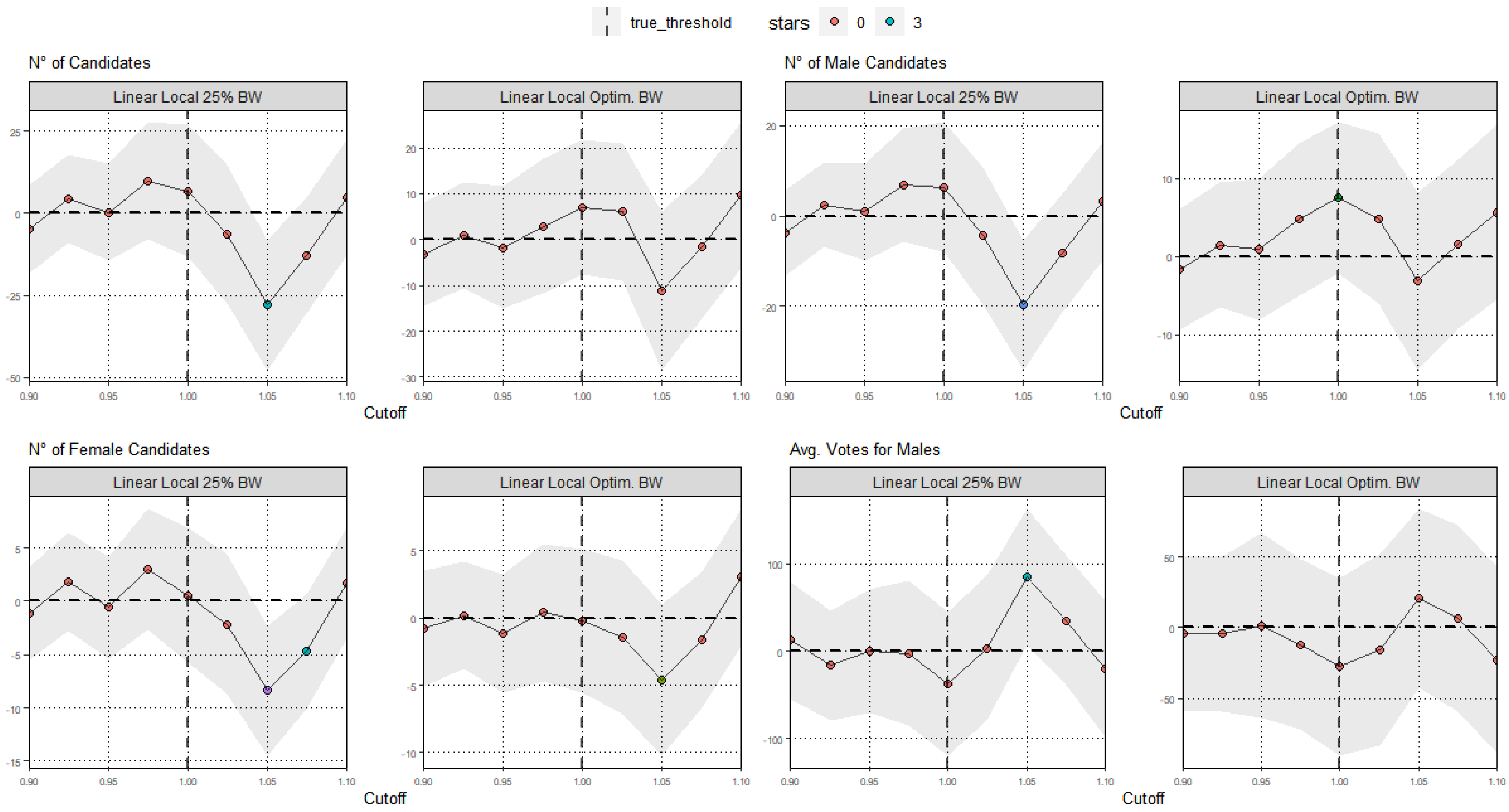
Notes: In this figure, we display a sensitivity analysis for our main RDD specification for policy outcomes. We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

Figure 12 – Sensitivity analysis - Policy Outcomes (2/2)



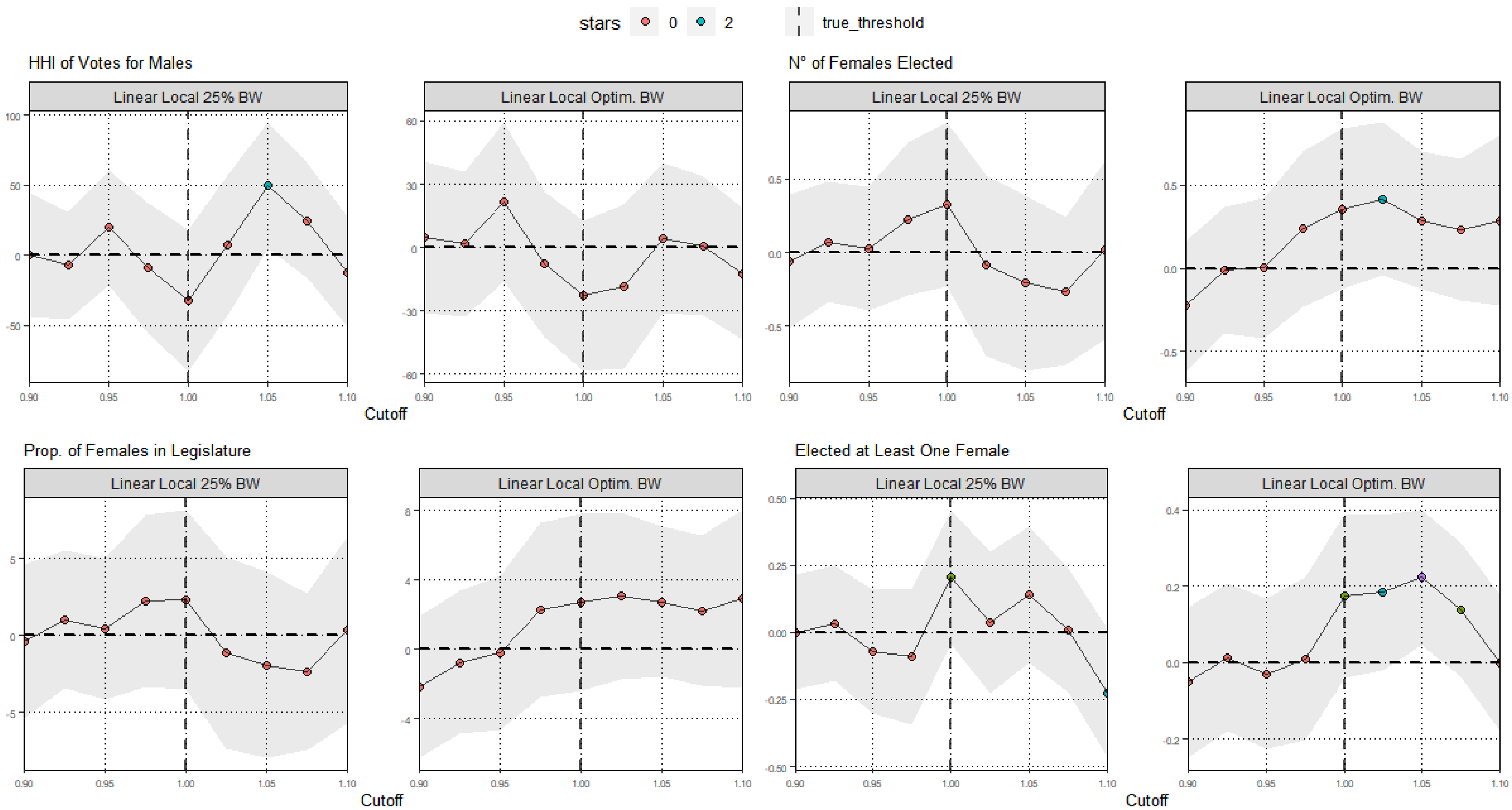
Notes: In this figure, we display a sensitivity analysis for our main RDD specification for policy outcomes. We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

Figure 13 – False Thresholds - Political Variables (1/2)



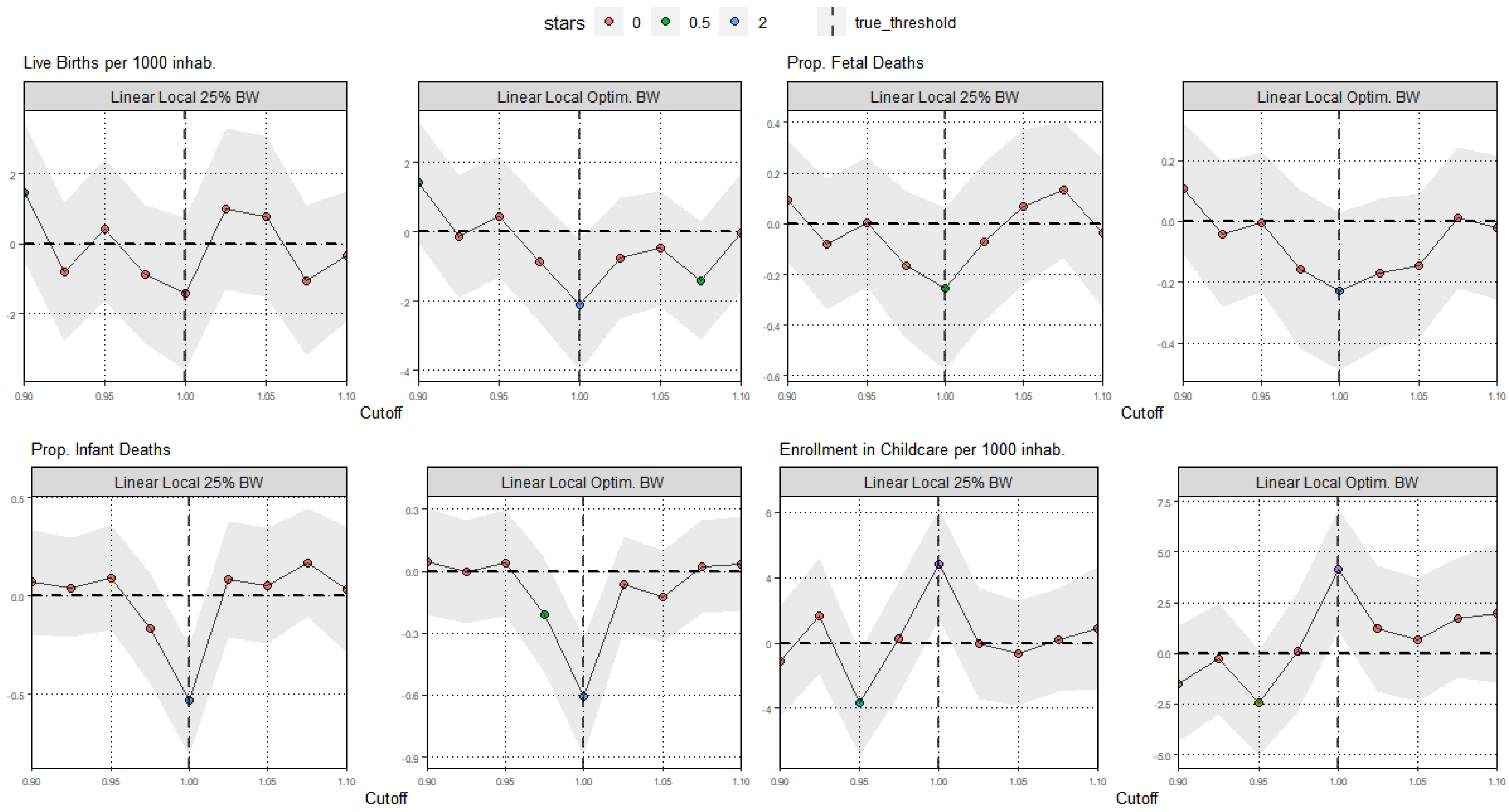
Notes: In this figure, we display the results of our main RDD specifications for political outcomes at false thresholds near the true cutoff point. Each point represents a point-estimate from a regression at a false threshold. We consider false cutoffs between 90% and 110% of the true cutoff of 47,619. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The shape of each point-estimate represents the number of observations (n).

Figure 14 – False Thresholds - Political Variables (2/2)



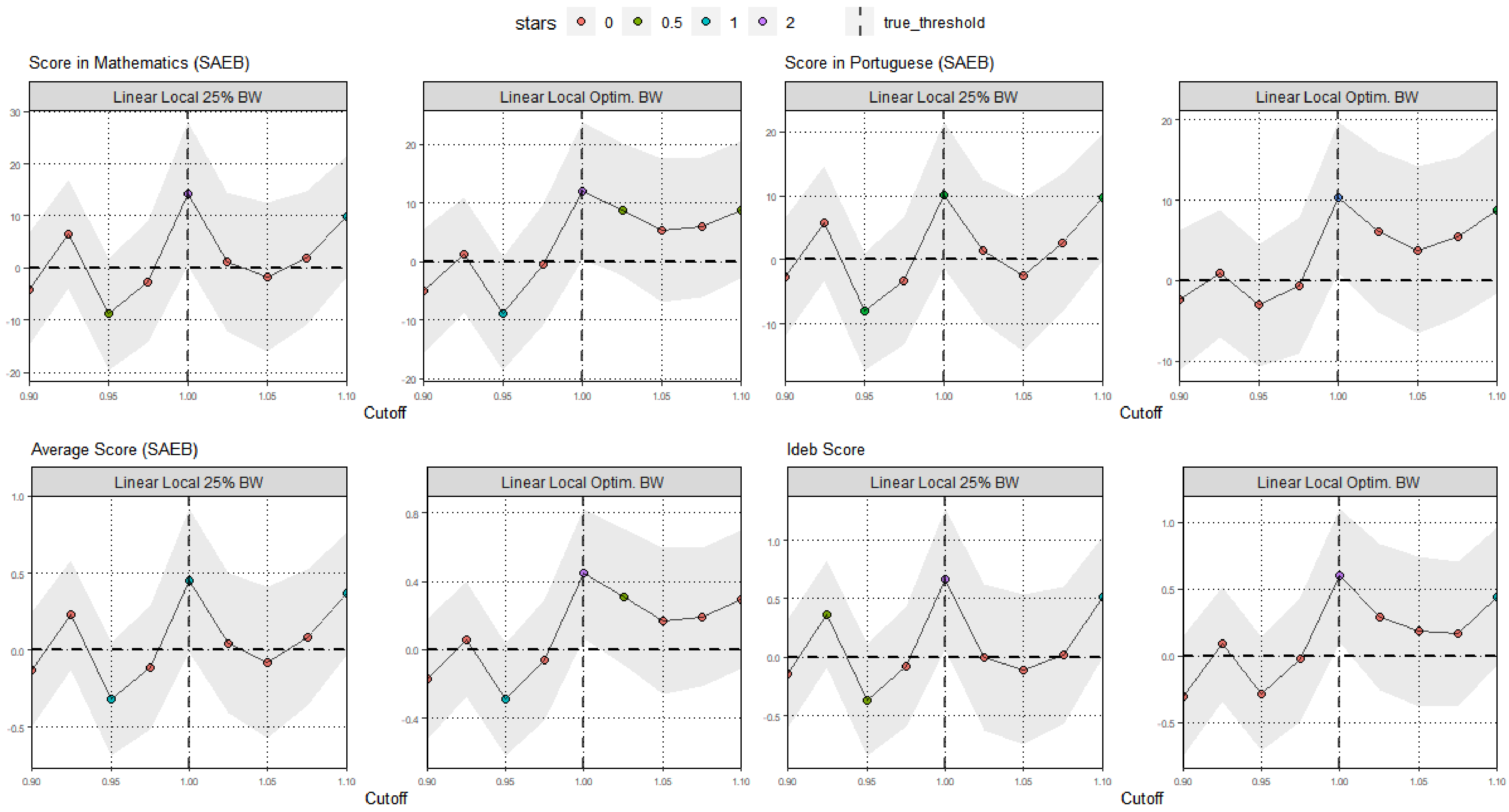
Notes: In this figure, we display the results of our main RDD specifications for political outcomes at false thresholds near the true cutoff point. Each point represents a point-estimate from a regression at a false threshold. We consider false cutoffs between 90% and 110% of the true cutoff of 47,619. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The shape of each point-estimate represents the number of observations (n).

Figure 15 – False Thresholds - Policy Outcomes (1/3)



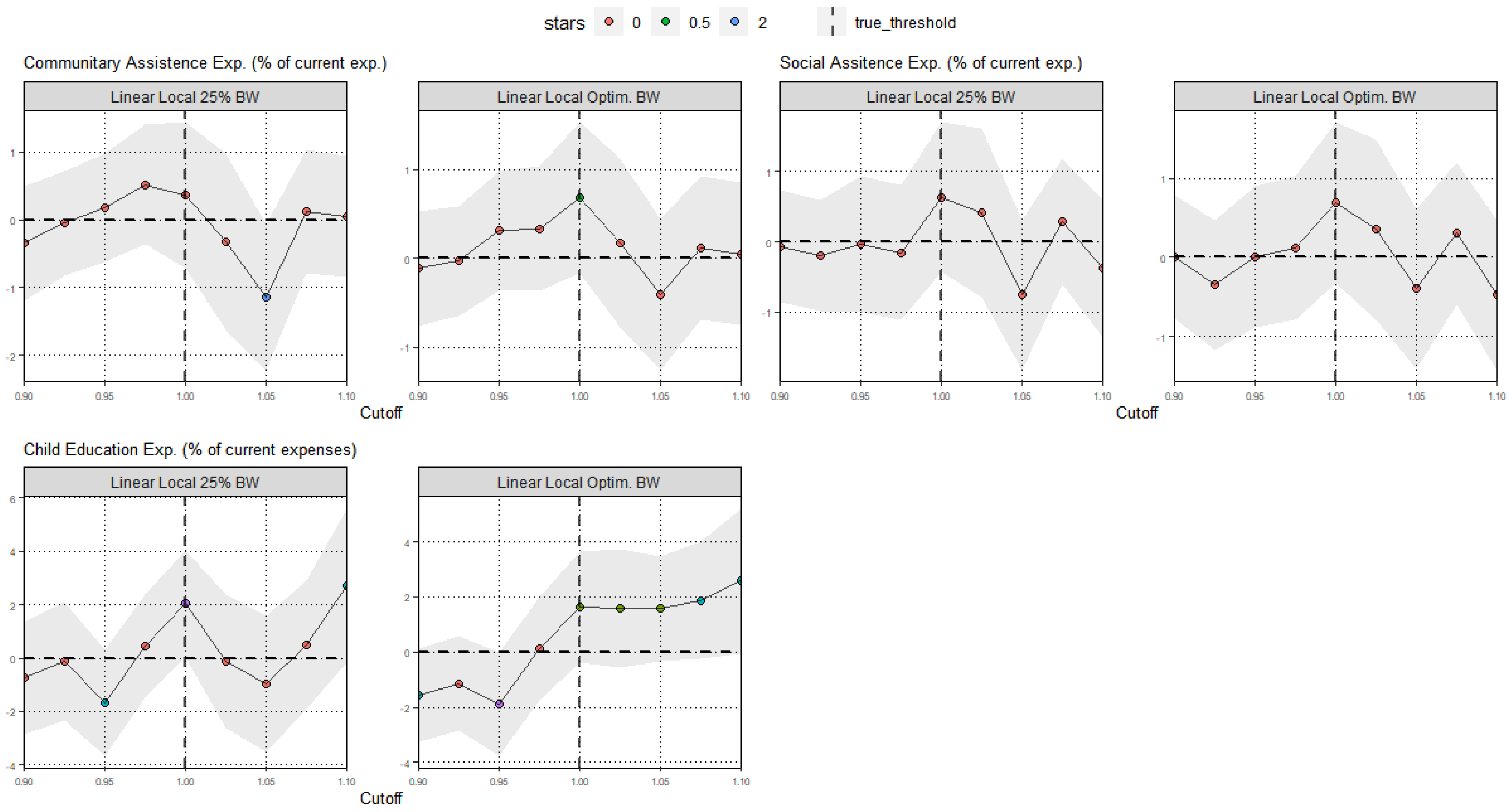
Notes: In this figure, we display the results of our main RDD specifications for policy outcomes at false thresholds near the true cutoff point. Each point represents a point-estimate from a regression at a false threshold. We consider false cutoffs between 90% and 110% of the true cutoff of 47,619. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The shape of each point-estimate represents the number of observations (n).

Figure 16 – False Thresholds - Policy Outcomes (2/3)



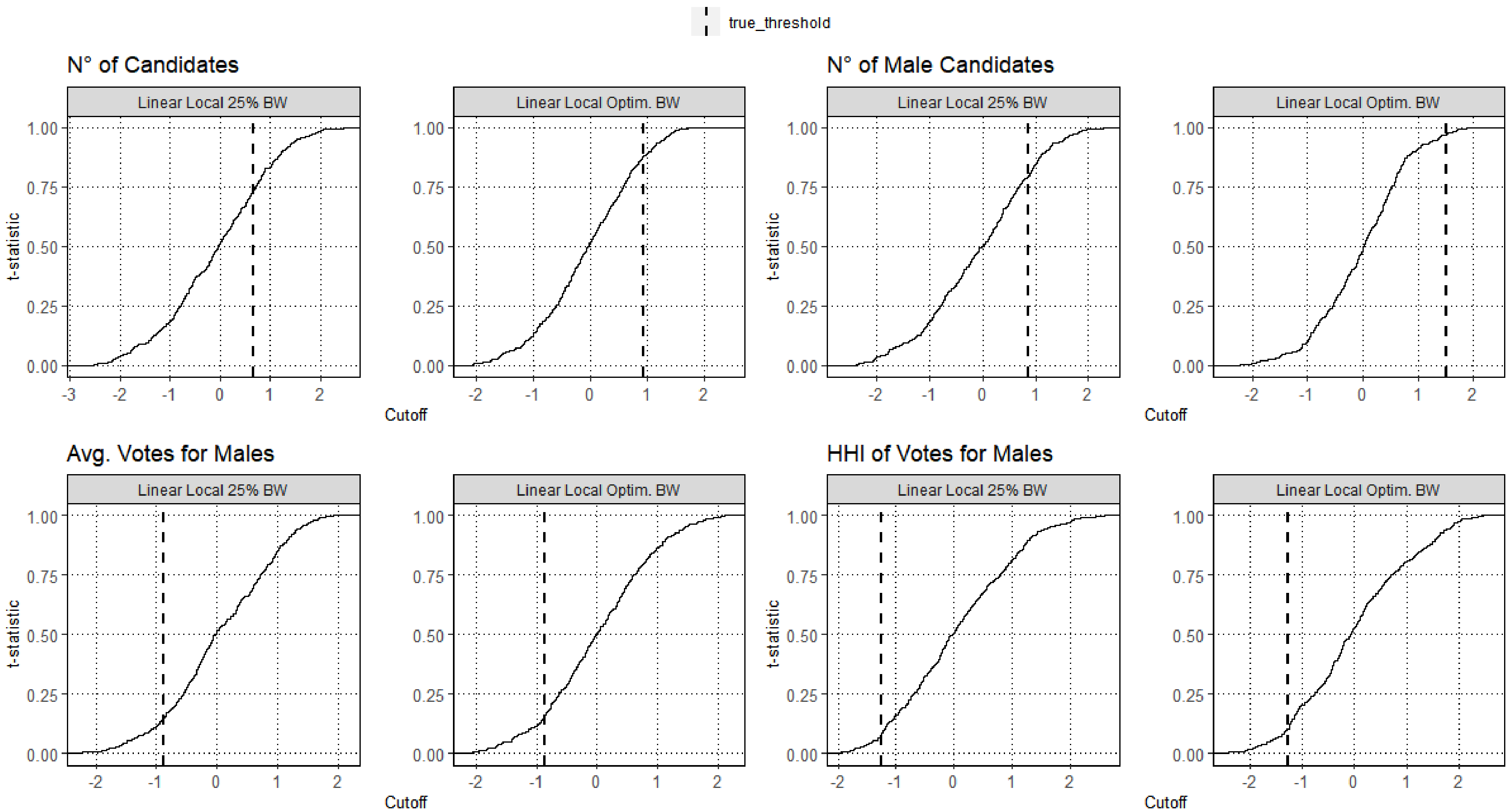
Notes: In this figure, we display the results of our main RDD specifications for political outcomes at false thresholds near the true cutoff point. Each point represents a point-estimate from a regression at a false threshold. We consider false cutoffs between 90% and 110% of the true cutoff of 47,619. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The shape of each point-estimate represents the number of observations (n).

Figure 17 – False Thresholds - Policy Outcomes (3/3)



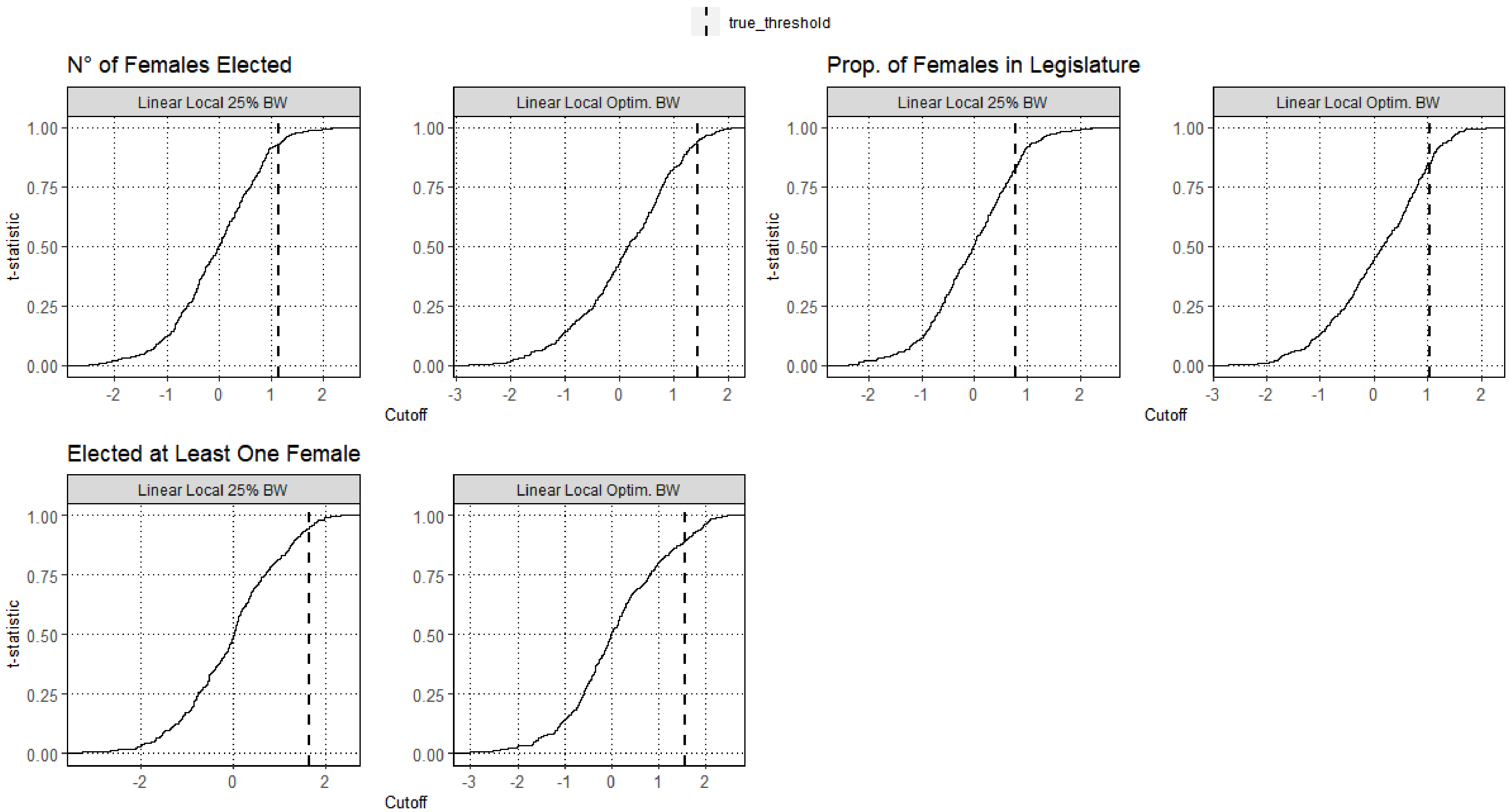
Notes: In this figure, we display the results of our main RDD specifications for political outcomes at false thresholds near the true cutoff point. Each point represents a point-estimate from a regression at a false threshold. We consider false cutoffs between 90% and 110% of the true cutoff of 47,619. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The shape of each point-estimate represents the number of observations (n).

Figure 18 – CDF of t-statistic (500 placebos) - Political Variables (1/2)



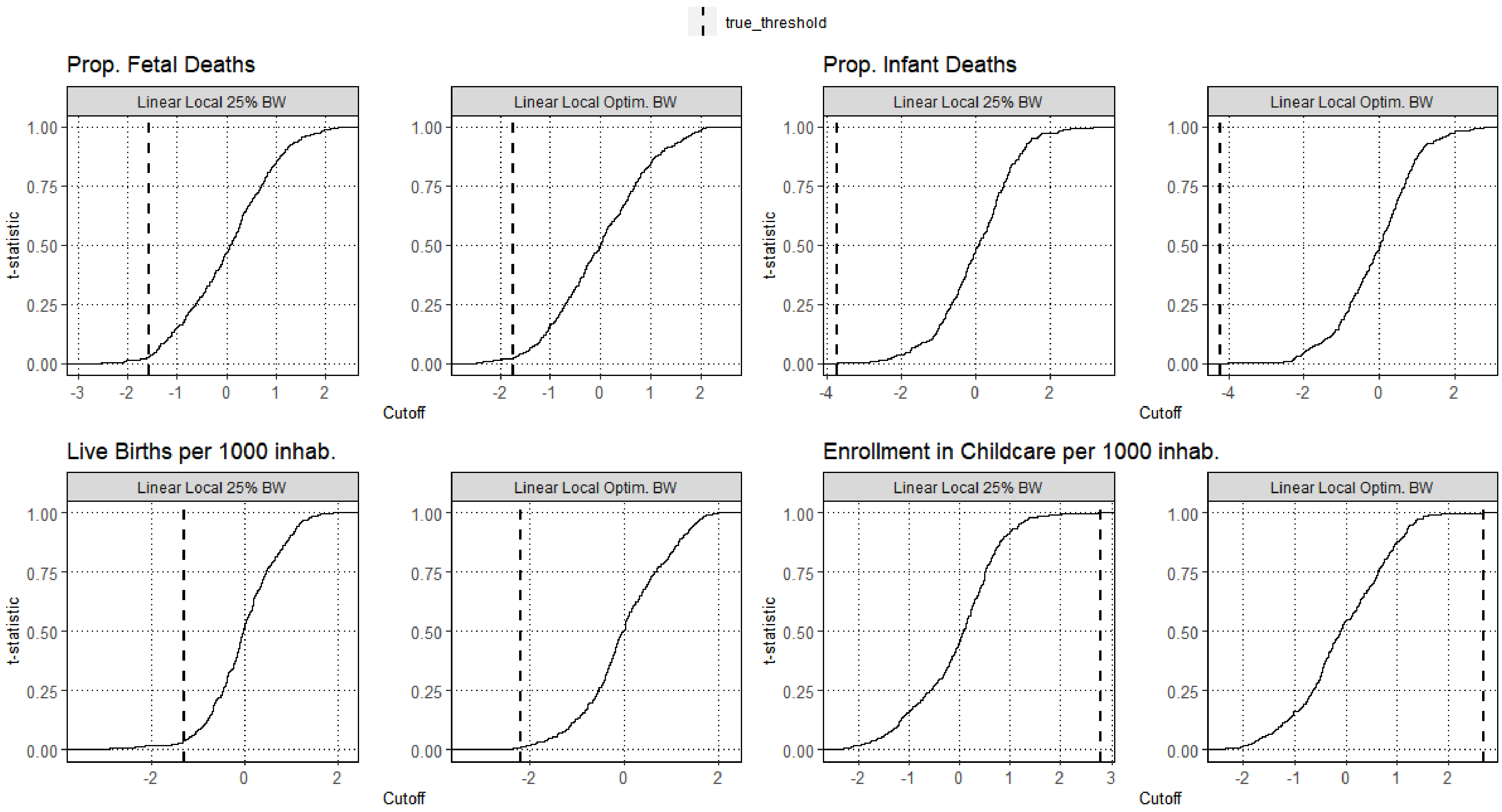
Notes: In this figure, we report the cumulative density function of the t-statistic from the RD estimations at various false thresholds for the political variables. We estimated 501 regressions at discontinuities between 50% and 150% of the true cutoff, using increments of 0.2%. The dashed vertical line represents the t-statistic of the specification at the true threshold. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth.

Figure 19 – CDF of t-statistic (500 placebos) - Political Variables (2/2)



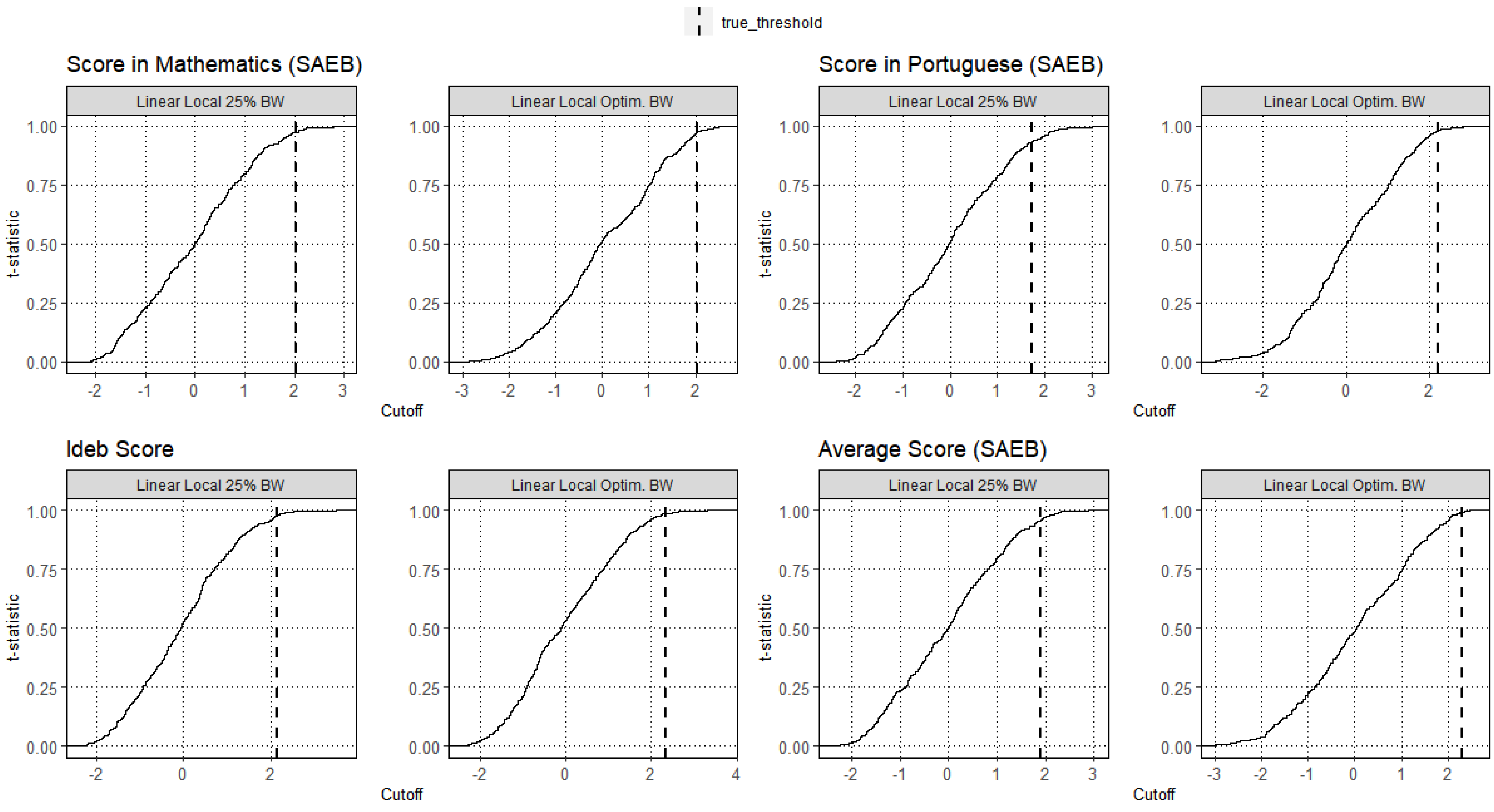
Notes: In this figure, we report the cumulative density function of the t-statistic from the RD estimations at various false thresholds for the political variables. We estimated 501 regressions at discontinuities between 50% and 150% of the true cutoff, using increments of 0.2%. The dashed vertical line represents the t-statistic of the specification at the true threshold. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth.

Figure 20 – CDF of t-statistic (500 placebos) - Policy Outcomes (1/3)



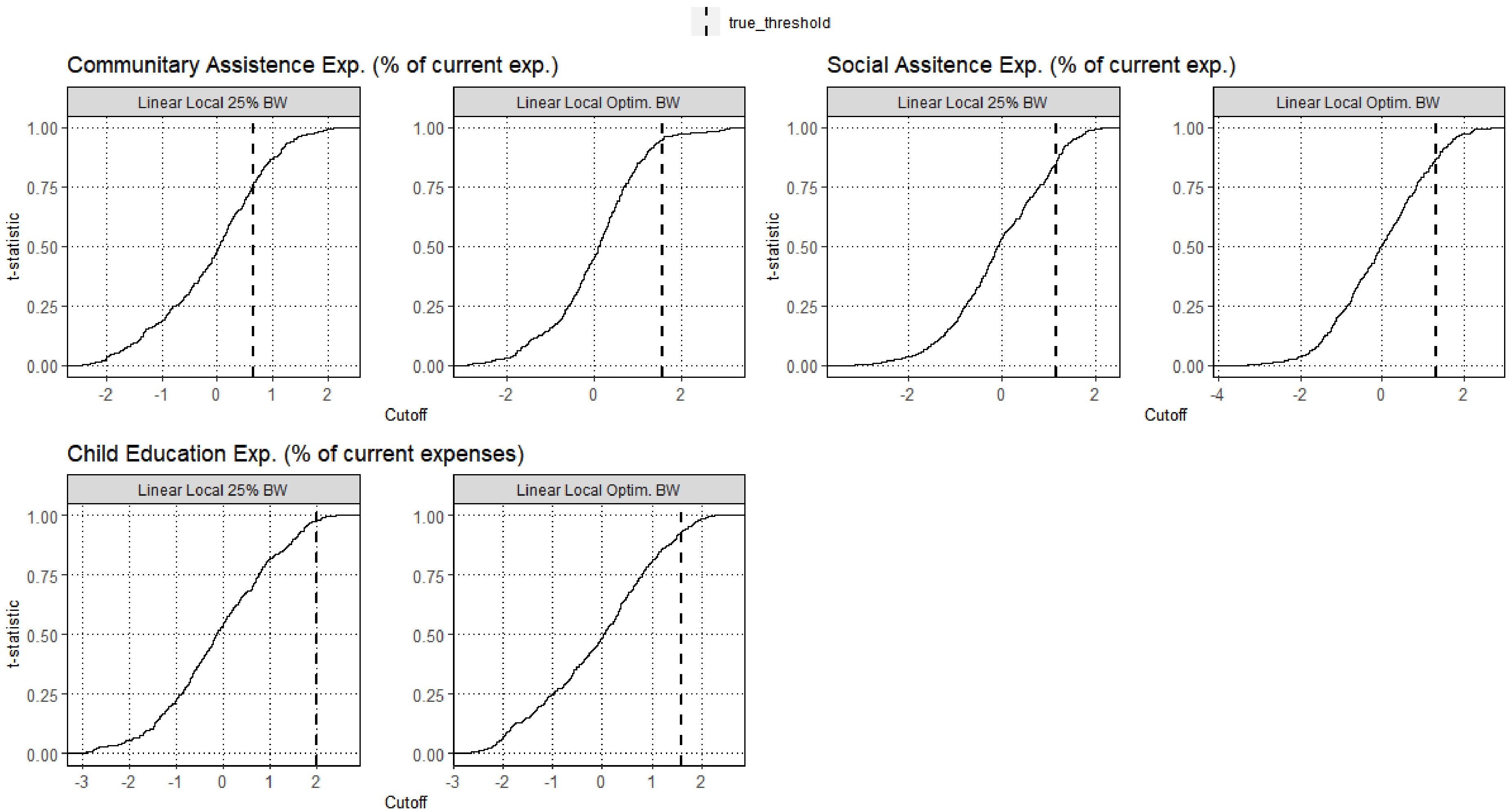
Notes: In this figure, we report the cumulative density function of the t-statistic from the RD estimations at various false thresholds for the political variables. We estimated 501 regressions at discontinuities between 50% and 150% of the true cutoff, using increments of 0.2%. The dashed vertical line represents the t-statistic of the specification at the true threshold. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth.

Figure 21 – CDF of t-statistic (500 placebos) - Policy Outcomes (2/3)



Notes: In this figure, we report the cumulative density function of the t-statistic from the RD estimations at various false thresholds for the political variables. We estimated 501 regressions at discontinuities between 50% and 150% of the true cutoff, using increments of 0.2%. The dashed vertical line represents the t-statistic of the specification at the true threshold. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth.

Figure 22 – CDF of t-statistic (500 placebos) - Policy Outcomes (3/3)



Notes: In this figure, we report the cumulative density function of the t-statistic from the RD estimations at various false thresholds for the political variables. We estimated 501 regressions at discontinuities between 50% and 150% of the true cutoff, using increments of 0.2%. The dashed vertical line represents the t-statistic of the specification at the true threshold. For each variable, we estimate Local Linear Regressions with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and uniform kernel. We consider two specifications. One with a fixed bandwidth of 25% and another with a MSE-optimal bandwidth.

1.7 Heterogeneity in attitudes towards female candidates

In this Section we investigate how political entry and electoral outcomes are affected by the presence of gender prejudice. In the theoretical framework presented in Section 1.2, the proportion of the electorate that is biased against women candidates was an important parameter to draw conclusions about the entry of candidates and subsequent electoral outcomes. To investigate this matter, we perform two exercises. In the first exercise, we divide our sample between municipalities that had elected at least one woman in the 2000 elections and those who did not. In a second exercise, we divide our sample between municipalities that were below or above the median proportion of votes cast for women in 2000. In both exercises, we restricted the sample to contain only municipalities that had observations for both 2004 and 2008, and did not have more than 95,238 inhabitants in both years.

An histogram of the percentage of votes cast for women candidates in 2000 can be seen in Figure 23. As it can be seen, there is a reasonable amount of heterogeneity in the proportion of votes across municipalities, with most of them having between 0% and 30% votes, with even some outliers where women managed to get around 40% of the votes.

Given this heterogeneity in the percentage of votes for women in 2000 and presence of women in the legislature in 2000, we can proceed to split the sample and see how the results for political entry and electoral results fare in comparison to the full sample. Results for the first exercise, with the proportion of votes cast for women in 2000, are displayed in Figures 24 and 25, and Tables 11 and 12. In the municipalities where the proportion of votes for women in 2000 were lower, we find a stronger effect of the legislature size on the electoral success of women: more women elected, higher proportion of women in the legislature and a higher chance of having at least one woman elected. Despite that, these variables are not consistent across all specifications in this sample. The magnitude for these estimates are consistent and also suggest an increase, however.

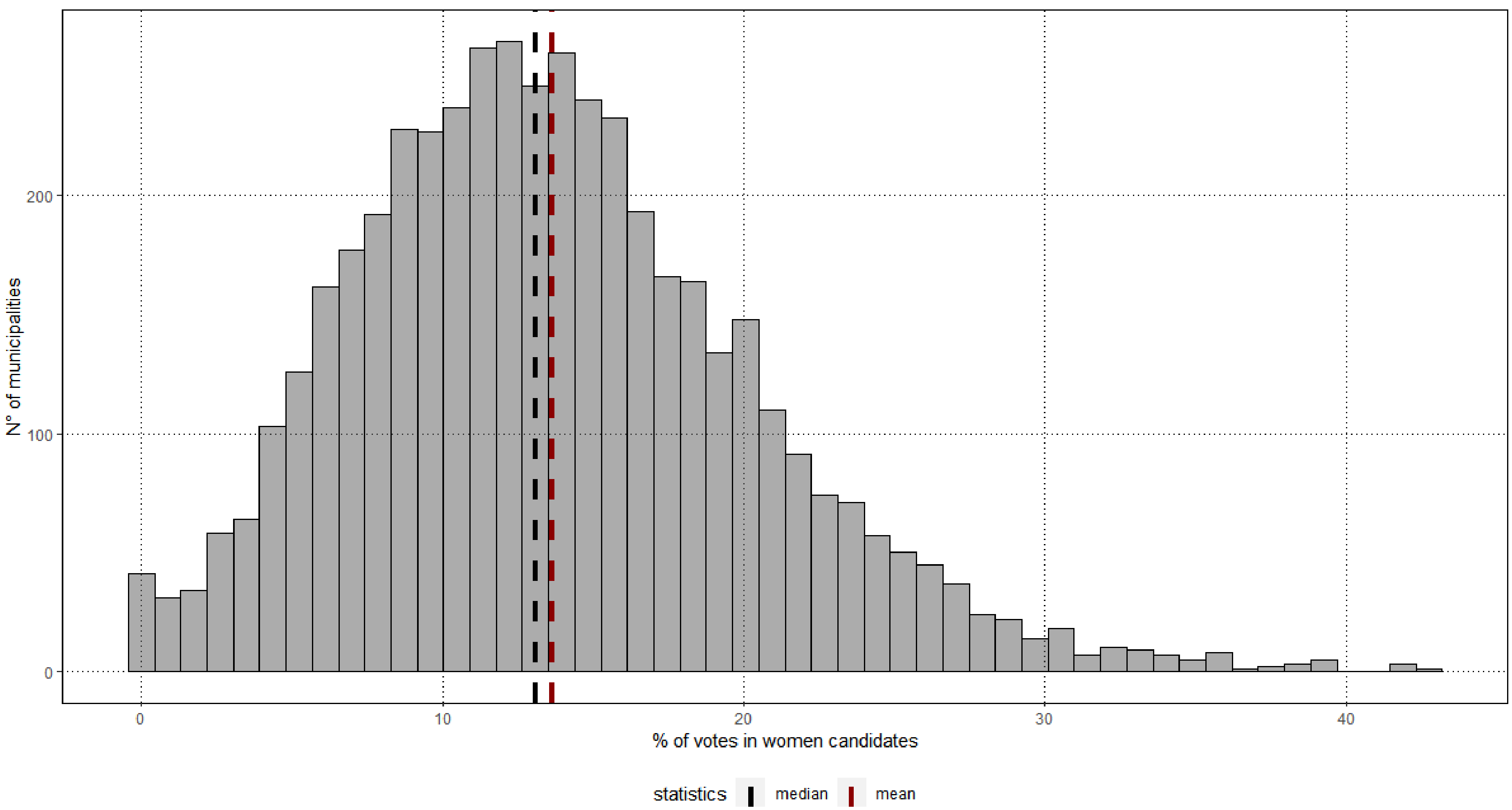
As for the sample of municipalities where female candidacies received a higher proportion of votes than the median in 2000, we find opposite results in relation to the previous sample. In this case, there are discontinuities for the number of candidates, male candidates, average votes for male candidates and HHI of votes for male candidates. The point-estimates also agree with the main specification, suggesting an increase in the number of male candidates, but at the cost of lower average votes and lower HHI. There seems to be some discontinuity in the variables associated to women's electoral success, but the values are more scattered. As in the previous case, the point-estimates, despite being consistent across specifications and agreeing with the RD graphs, most estimates are not significant.

In Figures 26 and 27 and Tables 13 and 14 we did the same exercise, but using the presence or not of women elected in 2000 to split the sample. The results are very similar to the previous exercise, with the exception that we do not find a clear discontinuity for the HHI of votes for males in the sample where at least one woman had been elected in 2000. Still in this sample, the observations for the variables of electoral success are also more scattered. As such, we only find a clear discontinuity for these variables in the sample where no woman had been elected in 2000.

We also tested the sensitivity of these findings in Figures 28, 29, 30 and 31. Despite not having significant results for most variables, the magnitude and direction of the point estimates are consistent for the variables for which we had found discontinuities in the graphs.

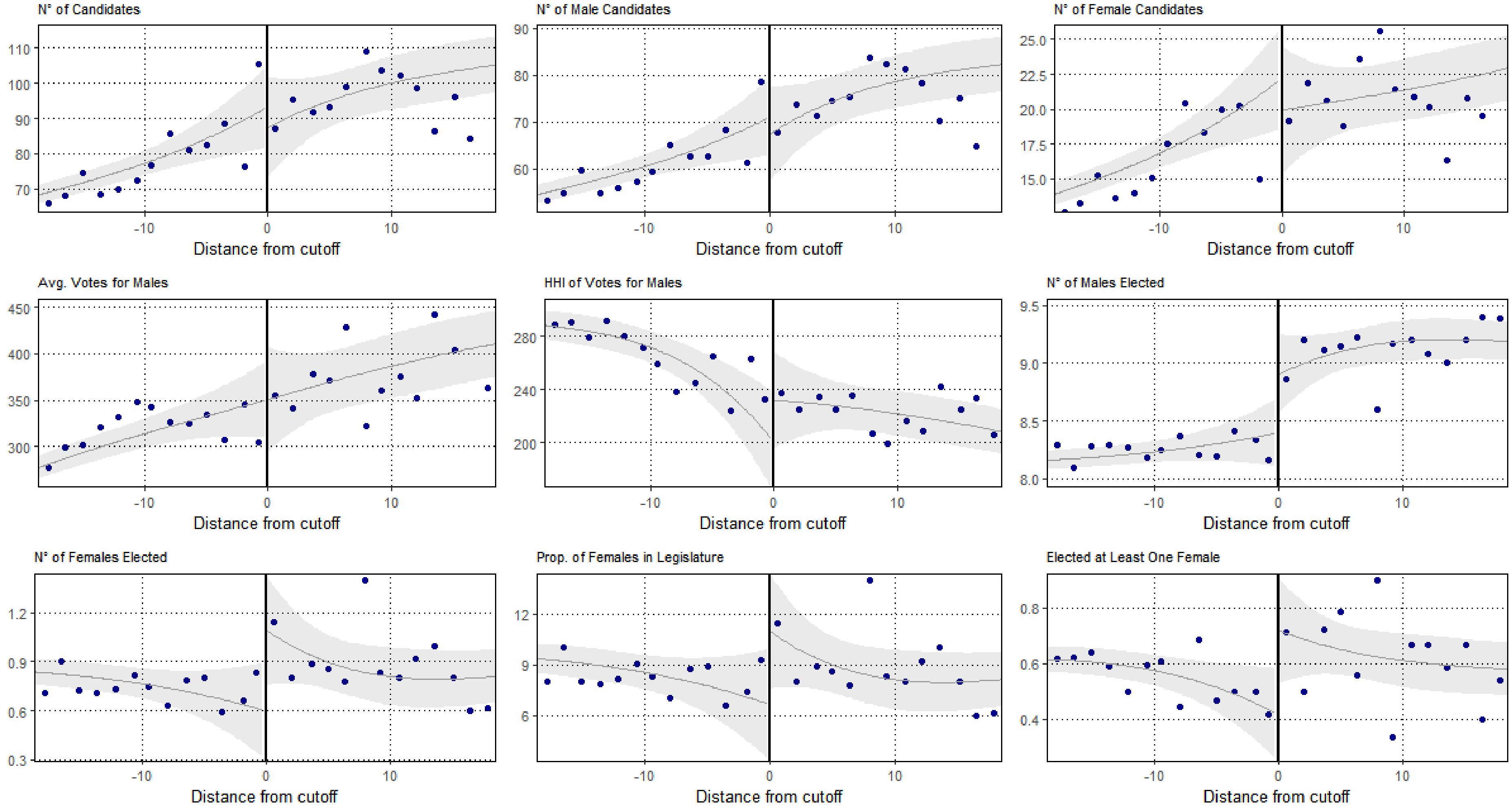
These results suggests that the impact of an increase in the size of the legislature on the electoral success of women is higher on municipalities that have a higher bias against female candidacies. However, in the same group of municipalities, we do not observe the cannibalization effect. Rather, this effect is only present in the sample of municipalities that have relatively less bias against female candidacies. These results are not fully explained by our theoretical framework. As such, it may be interesting to explore further the relationship between the size of legislative houses and the level of bias against female candidates.

Figure 23 – Histogram of the proportion of votes in women candidates in 2000.



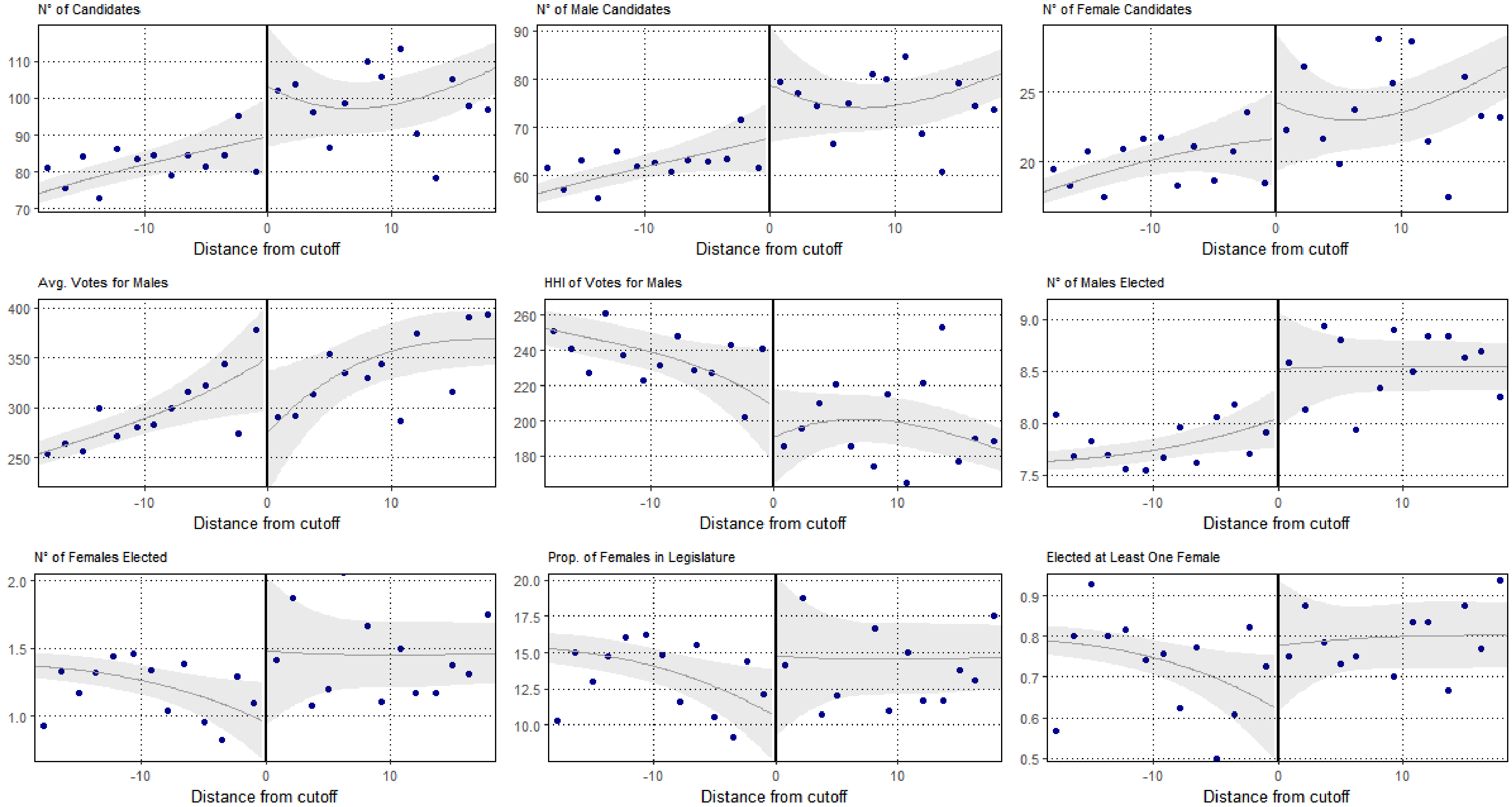
Notes: In this figure, we report a histogram for the proportion of votes on women candidates in 2000. We use a sample of municipalities that had observations for both 2004 and 2008, and did not have more than 95,238 inhabitants.

Figure 24 – Main Results - Political Variables - Votes cast for women in 2000: below the median



Notes: RD graph for the political variables for municipalities that had a low percentage of votes in women candidates in 2000 (below the median). In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Figure 25 – Main Results - Political Variables - Votes cast for women in 2000: above the median



Notes: RD graph for the political variables for municipalities that had a high percentage of votes in women candidates in 2000 (above the median). In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 11 – Effects of Local Council Size on Political Variables - Votes cast for women in 2000: below the median

Votes cast for women in 2000: below the median													
		Mean Comparison		Local Linear									
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW		
Variables	Mean	Obs	(1)	Obs	BW	(2)	Obs	(3)	Obs	(4)	Obs	(5)	
N° of Candidates	85.2113	135	8.5075	275	0.2387	-9.0546	135	-30.6531	294	-13.2945	805	-0.9969	
		135	(6.5006)	275	0.2387	(11.3571)	135	(22.5899)	294	(14.5227)	805	(9.7865)	
N° of Male Candidates	65.6620	135	7.338+	219	0.2017	-9.0431	135	-20.0841	294	-9.6625	805	0.8680	
		135	(4.6732)	219	0.2017	(8.5334)	135	(15.9185)	294	(10.2511)	805	(6.9319)	
N° of Female Candidates	19.5493	135	1.1695	401	0.3220	-1.5731	135	-10.569+	294	-3.6319	805	-1.8659	
		135	(2.0272)	401	0.3220	(3.1846)	135	(7.0531)	294	(4.5957)	805	(3.1221)	
Avg. Votes for Males	338.4617	135	20.9392	175	0.1689	47.3605	135	41.8018	294	35.2100	805	12.9726	
		135	(26.0312)	175	0.1689	(45.1047)	135	(65.0873)	294	(47.7656)	805	(36.2922)	
HHI of Votes for Males	254.8095	135	-26.6258+	443	0.3449	4.0228	135	-13.9570	294	-4.2568	805	-0.0966	
		135	(18.1559)	443	0.3449	(26.313)	135	(62.4247)	294	(40.2589)	805	(26.8005)	
N° of Males Elected	8.3239	135	0.8167***	334	0.2824	0.8416***	135	0.8418+	294	0.7662**	805	0.8541***	
		135	(0.1396)	334	0.2824	(0.2646)	135	(0.5417)	294	(0.3682)	805	(0.2375)	
N° of Females Elected	0.6761	135	0.1833	334	0.2824	0.1584	135	0.1582	294	0.2338	805	0.1459	
		135	(0.1396)	334	0.2824	(0.2646)	135	(0.5417)	294	(0.3682)	805	(0.2375)	
Prop. of Females in Legislature	7.5117	135	1.0820	358	0.2947	0.2344	135	0.6930	294	1.5272	804	0.6663	
		135	(1.4834)	358	0.2947	(2.7446)	135	(5.8771)	294	(3.9312)	804	(2.5054)	
Elected at Least One Female	0.4789	135	0.1774**	353	0.2910	0.1819	135	0.2436	294	0.2360	805	0.2139+	
		135	(0.0842)	353	0.2910	(0.1451)	135	(0.2873)	294	(0.196)	805	(0.1333)	

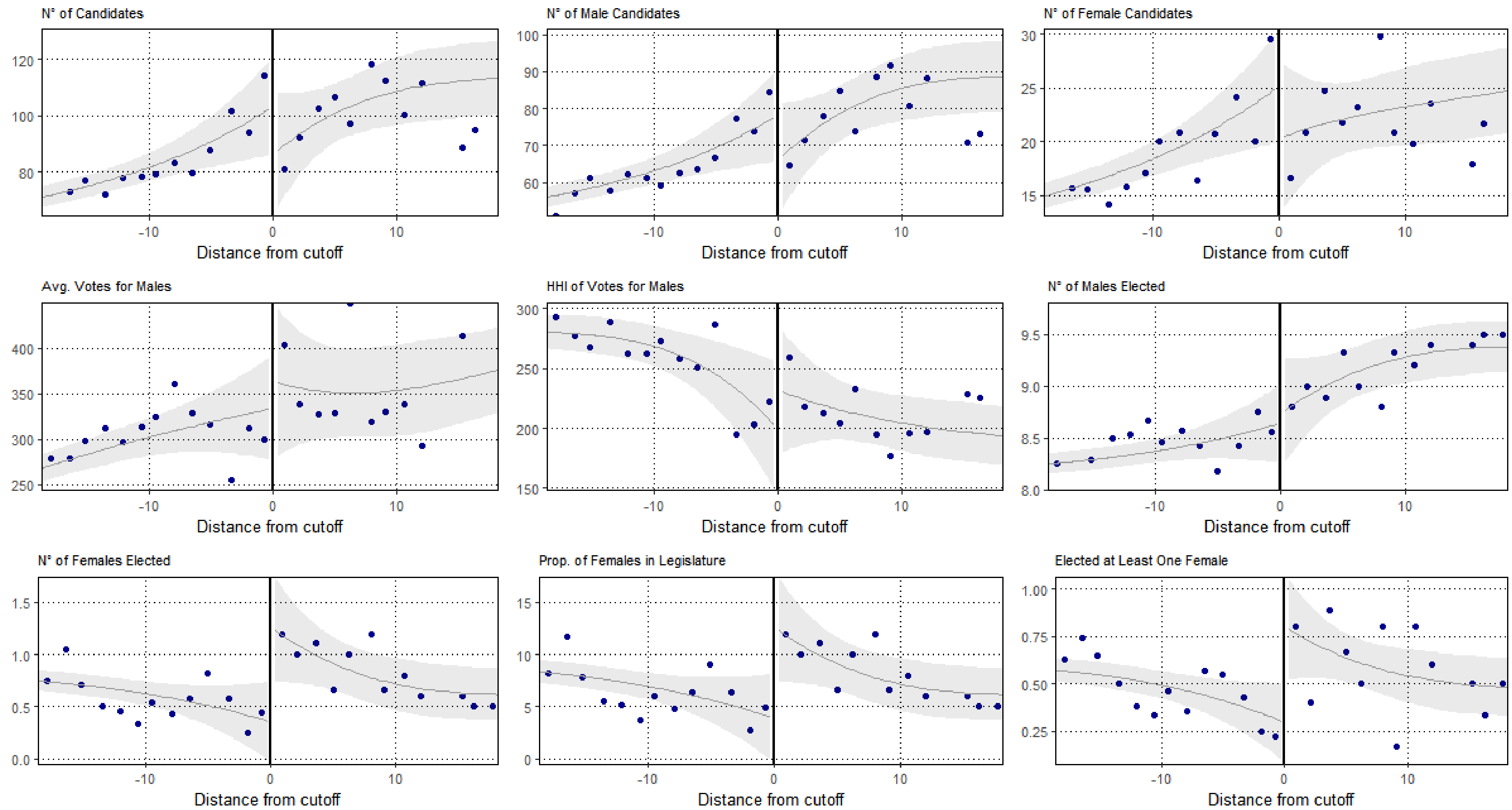
Notes: In this table, we report the RDD estimates for the effect of the council size on political variables associated with the entry of citizens to politics and its electoral results for municipalities that had a low percentage of votes in women candidates in 2000 (below the median). Columns (1)-(4) reports our estimates. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW.* reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 12 – Effects of Local Council Size on Political Variables - Votes cast for women in 2000: above the median

		Votes cast for women in 2000: above the median											
		Mean Comparison		Local Linear									
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW		
Variables	Mean	Obs	(1)	Obs	BW	(2)	Obs	(3)	Obs	(4)	Obs	(5)
N° of Candidates	86.1176	120	8.4977	271	0.2347	11.2134	120	45.2444*	284	16.1555	793	20.0038*
		120	(6.7358)	271	0.2347	(13.6203)	120	(25.168)	284	(19.0133)	793	(11.9486)
N° of Male Candidates	65.1618	120	7.569+	282	0.2482	12.5485	120	32.9019*	284	14.3447	793	15.6703*
		120	(4.9419)	282	0.2482	(9.5725)	120	(18.8491)	284	(13.7811)	793	(8.657)
N° of Female Candidates	20.9559	120	0.9287	238	0.2140	0.1532	120	12.3426*	284	1.8109	793	4.3391
		120	(2.0167)	238	0.2140	(4.8534)	120	(7.4546)	284	(5.8684)	793	(3.7187)
Avg. Votes for Males	320.3589	120	5.9913	313	0.2787	-53.4448	120	-118.7044	284	-60.6275	793	-57.4739
		120	(29.18)	313	0.2787	(54.9316)	120	(128.6588)	284	(81.7932)	793	(50.0487)
HHI of Votes for Males	224.7034	120	-18.2134	231	0.2074	-31.9192	120	-95.1876*	284	-31.6933	793	-36.1112+
		120	(14.4782)	231	0.2074	(30.114)	120	(54.0993)	284	(35.7815)	793	(24.4024)
N° of Males Elected	7.9853	120	0.6109***	307	0.2721	0.2064	120	0.3062	284	0.7398	793	0.4904
		120	(0.2001)	307	0.2721	(0.3732)	120	(0.8074)	284	(0.5249)	793	(0.3612)
N° of Females Elected	1.0147	120	0.3891*	307	0.2721	0.7936**	120	0.6938	284	0.2602	793	0.5096
		120	(0.2001)	307	0.2721	(0.3732)	120	(0.8074)	284	(0.5249)	793	(0.3612)
Prop. of Females in Legislature	11.2745	120	2.7640	310	0.2736	6.8839*	120	5.7058	284	1.3872	792	4.0092
		120	(2.0819)	310	0.2736	(3.8249)	120	(8.331)	284	(5.4138)	792	(3.7011)
Elected at Least One Female	0.6324	120	0.1753**	292	0.2634	0.1653	120	0.0813	284	0.0413	793	0.1161
		120	(0.0827)	292	0.2634	(0.1427)	120	(0.3151)	284	(0.1903)	793	(0.1257)

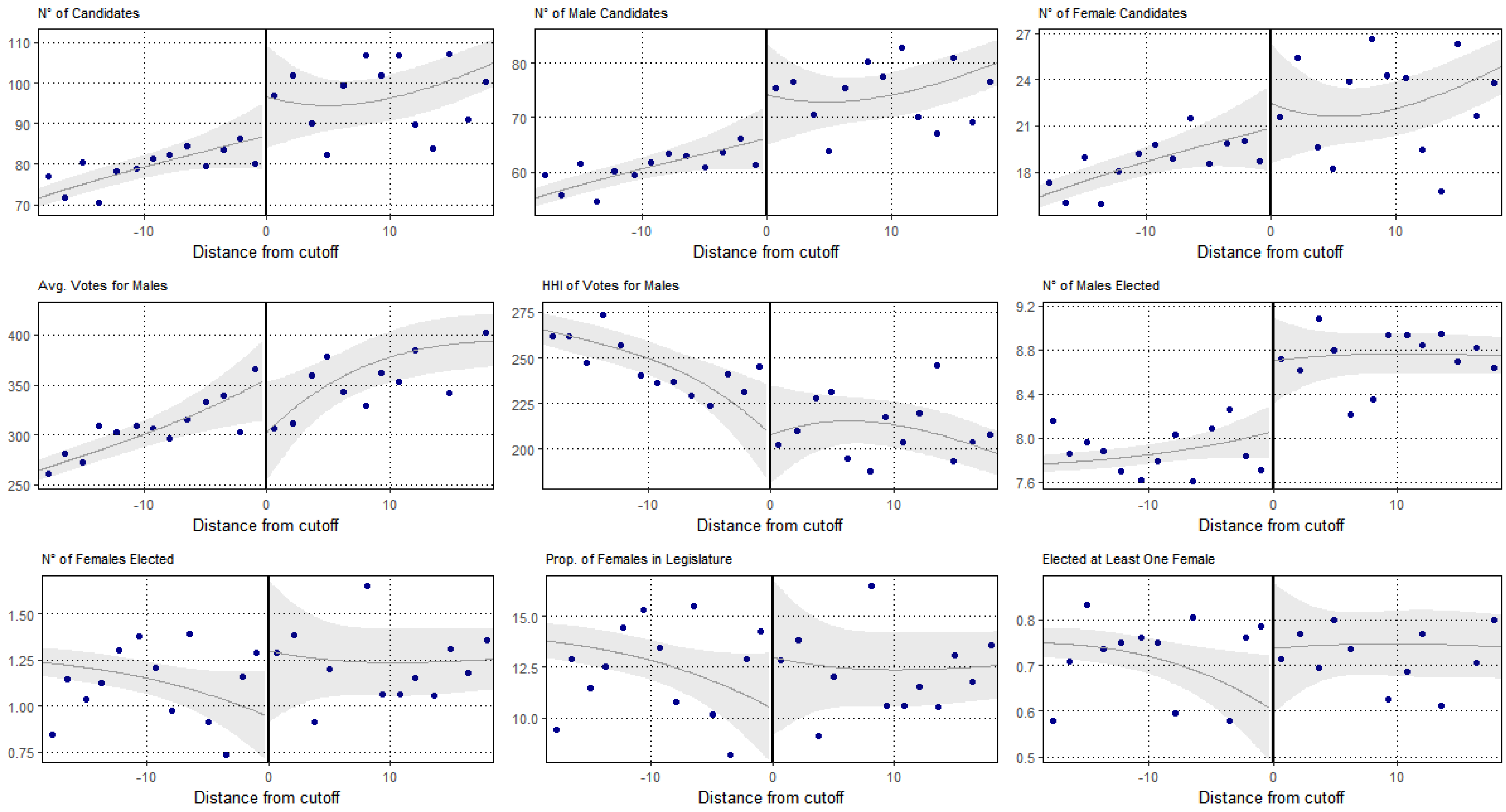
Notes: In this table, we report the RDD estimates for the effect of the council size on political variables associated with the entry of citizens to politics and its electoral results for municipalities that had a high percentage of votes in women candidates in 2000 (above the median). Columns (1)-(4) reports our estimates. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW.* reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 26 – Main Results - Political Variables - No woman elected in 2000



Notes: RD graph for the political variables for municipalities that did not elect any woman in 2000. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Figure 27 – Main Results - Political Variables - Had at least one woman elected in 2000



Notes: RD graph for the political variables for municipalities that had elected at least one woman in 2000. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 13 – Effects of Local Council Size on Political Variables - No woman elected in 2000

Elected at least one woman to the legislative: no													
		Mean Comparison		Local Linear									
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW		
Variables	Mean	Obs	(1)	Obs	BW	(2)	Obs	(3)	Obs	(4)	Obs	(5)	
N° of Candidates	95.0286	65	6.3714	125	0.2168	-22.2468	65	-45.7160	144	-42.6935*	371	-14.5574	
		65	[9.6415]	125	0.2168	[18.2175]	65	[36.5506]	144	[22.4802]	371	[15.5725]	
N° of Male Candidates	72.4000	65	6.2333	122	0.2150	-18.2822	65	-28.5513	144	-29.612*	371	-8.8773	
		65	[6.9563]	122	0.2150	[13.0409]	65	[25.7767]	144	[16.0412]	371	[11.0211]	
N° of Female Candidates	22.6286	65	0.1381	135	0.2416	-4.8532	65	-17.1647+	144	-13.0815*	371	-5.6801	
		65	[2.9842]	135	0.2416	[5.4928]	65	[11.2815]	144	[6.8677]	371	[4.9132]	
Avg. Votes for Males	316.0921	65	19.9993	137	0.2436	92.9142*	65	81.5345	144	100.4708	371	61.4450	
		65	[35.1919]	137	0.2436	[54.3349]	65	[121.2131]	144	[83.9085]	371	[55.0884]	
HHI of Votes for Males	250.6594	65	-34.8020	141	0.2482	27.7864	65	27.0377	144	54.8453	371	26.6166	
		65	[27.5571]	141	0.2482	[47.6967]	65	[102.8162]	144	[68.6746]	371	[42.3344]	
N° of Males Elected	8.4571	65	0.6095***	152	0.2701	0.4702	65	0.4837	144	0.2422	371	0.2849	
		65	[0.2027]	152	0.2701	[0.4042]	65	[0.8126]	144	[0.6042]	371	[0.3619]	
N° of Females Elected	0.5429	65	0.3905*	152	0.2701	0.5298	65	0.5163	144	0.7578	371	0.7151**	
		65	[0.2027]	152	0.2701	[0.4042]	65	[0.8126]	144	[0.6042]	371	[0.3619]	
Prop. of Females in Legislature	6.0317	65	3.3016+	158	0.2863	4.6520	65	4.7924	144	7.2055	371	6.6383*	
		65	[2.1603]	158	0.2863	[4.1221]	65	[8.5727]	144	[6.2878]	371	[3.7836]	
Elected at Least One Female	0.3714	65	0.3286***	153	0.2725	0.4391**	65	0.3814	144	0.6174**	371	0.53***	
		65	[0.113]	153	0.2725	[0.2061]	65	[0.4681]	144	[0.3022]	371	[0.1971]	

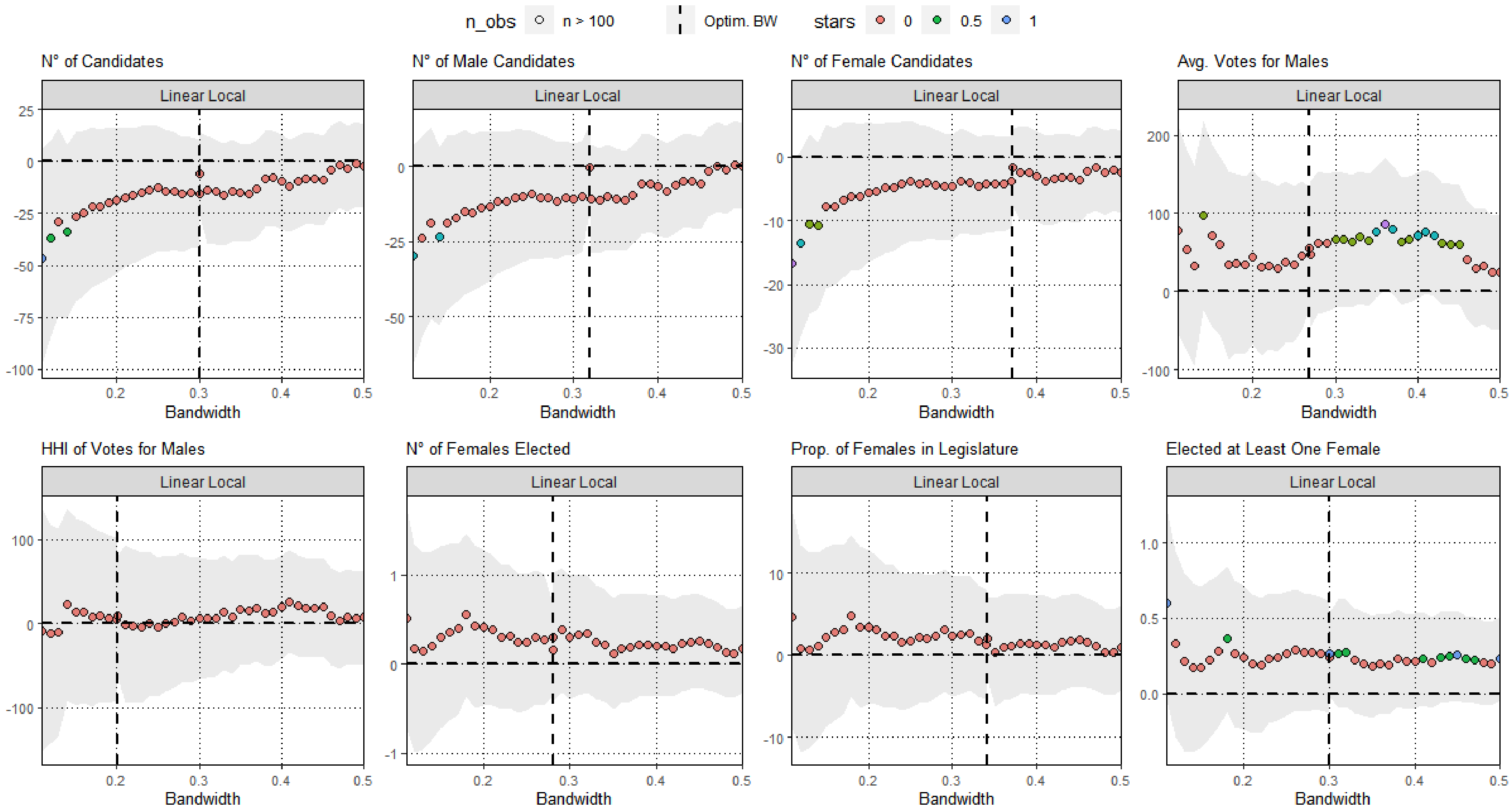
Notes: In this table, we report the RDD estimates for the effect of the council size on political variables associated with the entry of citizens to politics and its electoral results for municipalities that had no women elected in 2000. Columns (1)-(4) reports our estimates. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW.* reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 14 – Effects of Local Council Size on Political Variables - Had at least one woman elected in 2000

Elected at least one woman to the legislative: yes													
		Mean Comparison		Local Linear									
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW		
Variables	Mean	Obs	(1)	Obs	BW	(2)	Obs	(3)	Obs	(4)	Obs	(5)	
N° of Candidates	82.5000	190	9.0814*	443	0.2577	11.0936	190	22.1849	434	13.8915	1227	15.1635*	
		190	[5.2352]	443	0.2577	[9.2562]	190	[16.1338]	434	[12.3093]	1227	[8.5048]	
N° of Male Candidates	63.0673	190	7.8048**	435	0.2519	7.5356	190	15.8554	434	10.9195	1227	12.1604**	
		190	[3.7943]	435	0.2519	[6.9899]	190	[11.8658]	434	[8.8689]	1227	[6.1072]	
N° of Female Candidates	19.4327	190	1.2766	407	0.2335	1.0504	190	6.3296	434	2.9719	1227	3.0061	
		190	[1.6191]	407	0.2335	[3.2022]	190	[4.9903]	434	[3.8477]	1227	[2.6986]	
Avg. Votes for Males	334.1535	190	13.3945	531	0.2953	-31.7945	190	-41.9247	434	-37.8124	1227	-40.7510	
		190	[23.3801]	531	0.2953	[38.8633]	190	[85.0527]	434	[55.2094]	1227	[36.2689]	
HHI of Votes for Males	236.5215	190	-17.1549	435	0.2527	-22.8805	190	-52.0132	434	-37.4866	1227	-26.1317	
		190	[13.0736]	435	0.2527	[23.1231]	190	[43.8794]	434	[29.5623]	1227	[20.9638]	
N° of Males Elected	8.0577	190	0.7795***	337	0.2023	0.7685***	190	1.1695**	434	1.0789***	1227	0.8929***	
		190	[0.1491]	337	0.2023	[0.2948]	190	[0.4832]	434	[0.3443]	1227	[0.248]	
N° of Females Elected	0.9423	190	0.2205+	337	0.2023	0.2315	190	-0.1695	434	-0.0789	1227	0.1071	
		190	[0.1491]	337	0.2023	[0.2948]	190	[0.4832]	434	[0.3443]	1227	[0.248]	
Prop. of Females in Legislature	10.4701	190	1.1578	349	0.2083	1.6378	190	-3.3085	434	-2.0386	1225	-0.0736	
		190	[1.5571]	349	0.2083	[3.0296]	190	[5.1721]	434	[3.6234]	1225	[2.5686]	
Elected at Least One Female	0.6154	190	0.1172*	383	0.2222	0.1386	190	-0.1136	434	-0.0992	1227	0.0345	
		190	[0.0696]	383	0.2222	[0.1311]	190	[0.232]	434	[0.1522]	1227	[0.1058]	

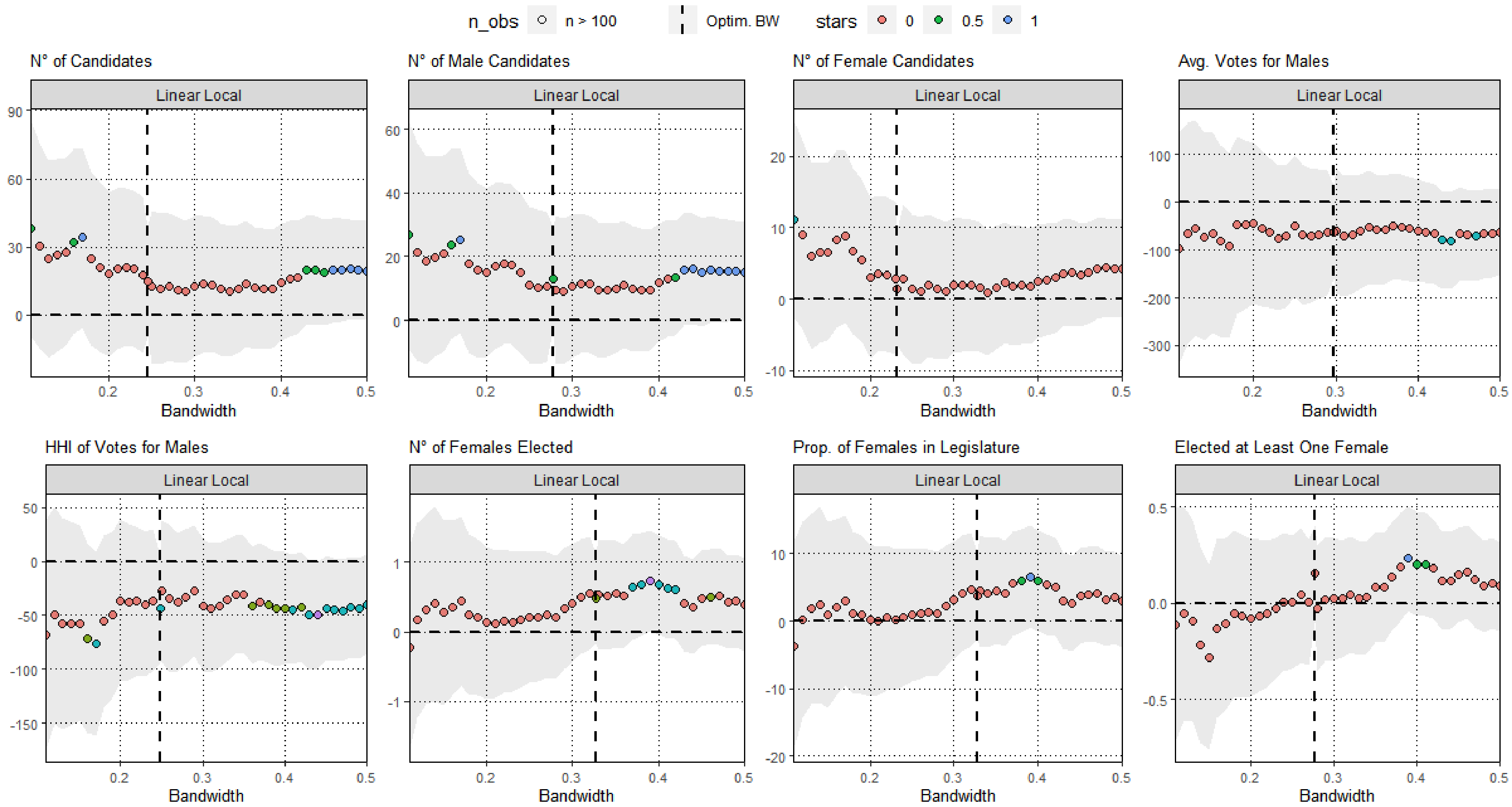
Notes: In this table, we report the RDD estimates for the effect of the council size on political variables associated with the entry of citizens to politics and its electoral results for municipalities that had elected at least one woman in 2000. Columns (1)-(4) reports our estimates. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW*. reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 28 – Sensitivity analysis - Political Variables - Votes cast for women in 2000: below the median



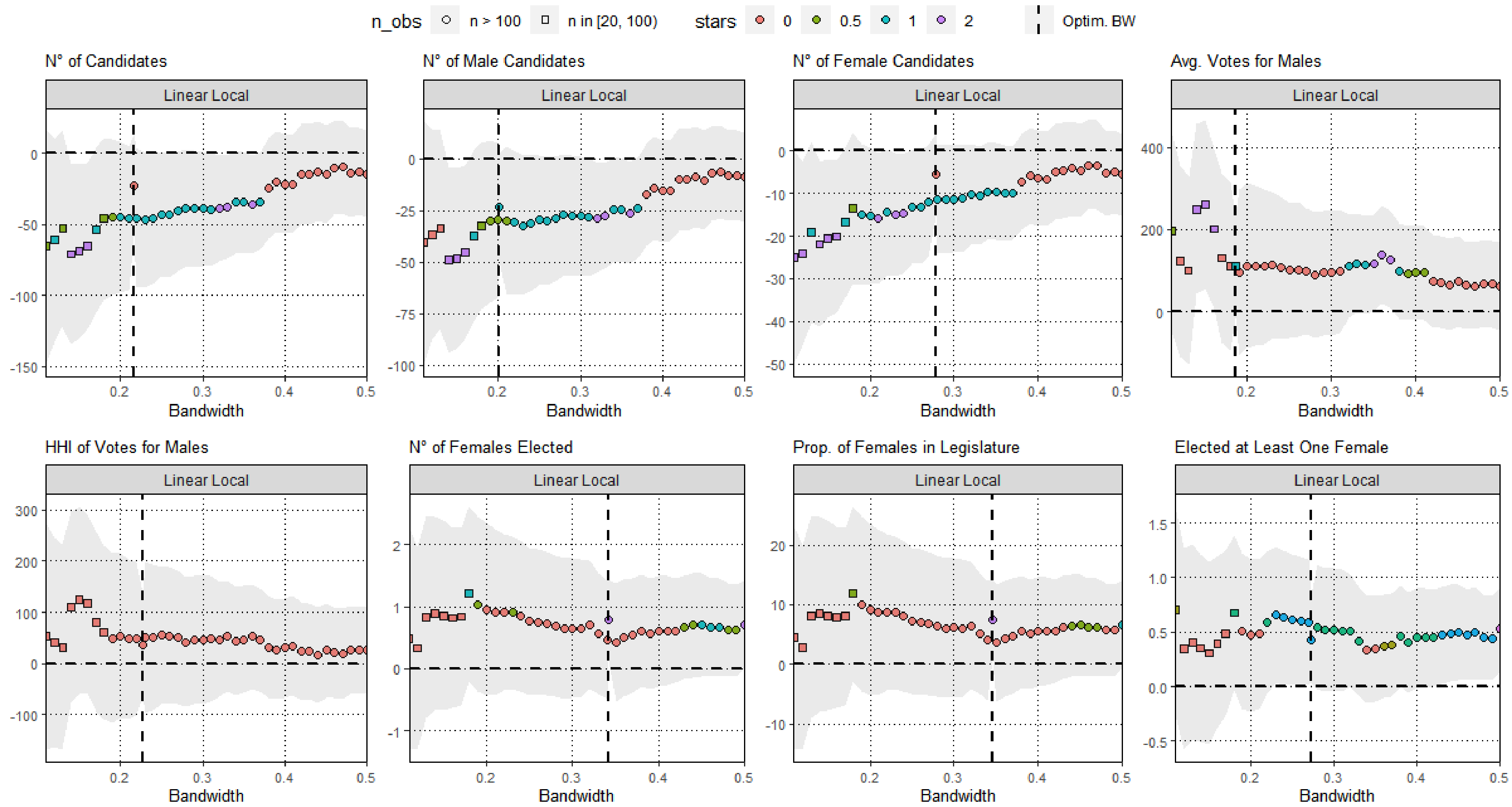
Notes: In this figure, we display a sensitivity analysis for our main RDD specification for political outcomes for municipalities that had a low percentage of votes in women candidates in 2000 (below the median). We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

Figure 29 – Sensitivity analysis - Political Variables - Votes cast for women in 2000: above the median



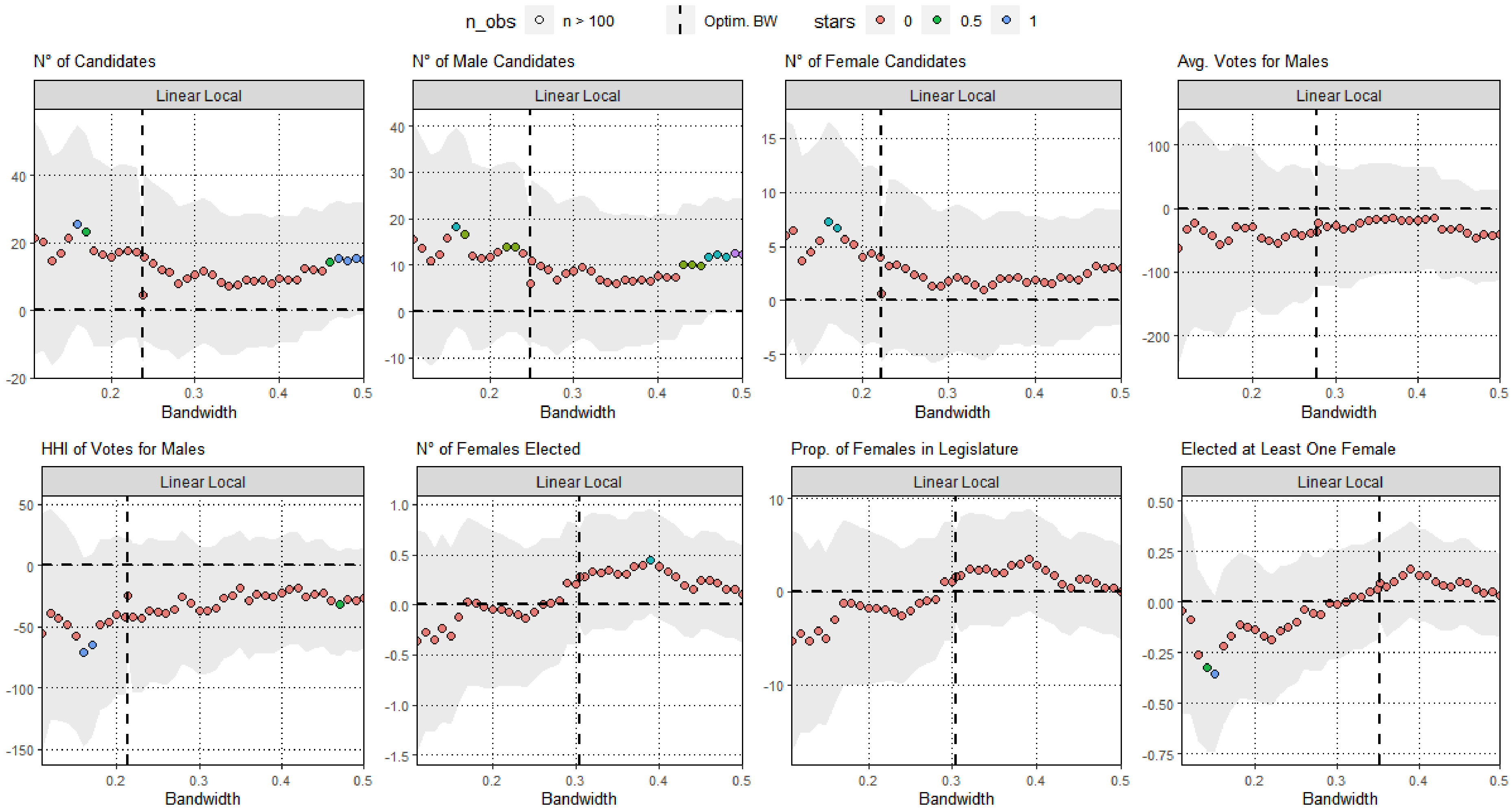
Notes: In this figure, we display a sensitivity analysis for our main RDD specification for political outcomes for municipalities that had a high percentage of votes in women candidates in 2000 (above the median). We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

Figure 30 – Sensitivity analysis - Political Variables - No woman elected in 2000



Notes: In this figure, we display a sensitivity analysis for our main RDD specification for political outcomes for municipalities that did not elect any woman in 2000. We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

Figure 31 – Sensitivity analysis - Political Variables - Had at least one woman elected in 2000



Notes: In this figure, we display a sensitivity analysis for our main RDD specification for political outcomes for municipalities that had elected at least one woman in 2000. We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

1.8 Conclusion

In this paper, we entertained a potential link between the size of the legislature and the representation of women. We examine the quasi-experiment for the 2004 and 2008 municipal elections in which a resolution by the Supreme Electoral Court required legislature sizes to be calculated as a deterministic and discontinuous function of the population of municipalities. We find evidence that indeed increasing the number of seats in municipal councils leads to a phenomenon in which more men enter the competition and compete with each other. Meanwhile, women keep their vote shares more or less intact and this translates into more women being elected. The results of this increase in women's participation lead to sizable improvements in early education, antenatal and infant health care, and assistance spending.

The explanation for this behavior relies only on mild assumptions. First, voters are able to coordinate, but only imperfectly. Second, we assume part of the voters are biased against women. And lastly, the cost for women to run for office is relatively higher for women. Under these assumptions, voters can coordinate and this leads to a higher overall number of both male candidates and male elected candidates, but as more men enter the election, the less able voters are to coordinate their votes. This leads a few votes to be wasted, which means that despite the fact there are more male candidates, it won't necessarily translate into relatively more men being elected at the margin.

If these results are to be convincing, that means the number of legislative seats might be an important variable to understand how the structure of the legislative house might enhance or hinder the participation of women and potentially other underrepresented groups. It is important to keep in mind that this is only one aspect of increasing the size of legislative houses. We only examine a small increase in size. Increasing the number of seats too much might have other consequences. As Pettersson-Lidbom (2012) had found for Finland and Sweden, increasing the council size could have public finance effects, such as a decrease in both spending and revenues.

A second potential issue is the interaction between the increase in the size of legislatures and the level of bias against women candidates. We found that increasing the number of seats had a more visible impact on the electoral success of women in municipalities where they did not do well in the previous elections. Meanwhile, the cannibalization effects were more visible in the set of municipalities that are expected to have less bias against women. Our theoretical framework does not fully explain these results. As such, it may be worth it to further investigate these heterogeneous effects of gender bias and the mechanisms described in this chapter.

Another thing to point out is that we only examined how the size of legislative houses impacts the election of women. However, many of the challenges faced by women might be similar to challenges faced by other under-represented groups. Therefore, it might be possible that our effects are not exclusive to gender. As such, the study of the structure of legislative houses and how it relates to under-represented groups might be an interesting topic for future research.

1.9 Appendix

1.9.1 Exercises with fixed effects of year and state

In this appendix, we include an additional exercise in which we include year and state fixed effects to the main specification. Although it is not necessary for the validity of our Regression Discontinuity Design, it allows us to filter out potential noise that might arise from these fixed effects. In Figures 32, 33, 34 and 35, we filter out the year and state fixed effect from the variables and then plot their values at the cutoff. The results are mostly unchanged, however, some variables display change. The estimates for the number of candidates, male candidates and HHI still display a discontinuity, but it is now more nuanced. Overall, the educational estimates also have worsened, especially for the SAEB score in portuguese and the Ideb score. While they still display their respective discontinuities, the confidence intervals are relatively bigger near the cutoff. There are no noticeable differences for the other variables, however.

In Table 15, 16, 17 and 17 there are more noticeable differences. The number of candidates is now only significant in the mean comparison and the number of male candidates is only significant in the mean comparison and for a 50% bandwidth. Moreover, the point-estimates are lower for the Local Linear specification. The average number of votes for male candidates is now significant for the 50% bandwidth but is still not significant for the other specifications. The HHI of votes for males is now only significant at the mean comparison. As for the variables of female electoral performance, the number of women elected is significant for all fixed bandwidths, and the same is true for the probability of electing at least one woman. However, the proportion of women in the legislature is only significant at the mean comparison and with a large bandwidth of 50%. This result is similar to what we had in the main tables.

Moving on to the policy outcomes, the number of live births, SAEB's score in portuguese and expenditures in child education all lose significance. The magnitude of the effects is also closer to zero, except for the score in portuguese, which is lower overall, but still positive and consistent across the Local Linear specification. On the other hand, the expenditure with social assistance is now significant in all specifications, even though the expenditure with community assistance loses significance in the specification with optimal bandwidth. As for the educational outcomes, before they all were significant for all specifications. Now portuguese loses effect and the other loses significance in some cases but are still significant overall. The magnitude of the effect is also lower in relation to the main tables.

These estimates give us mixed results. Some results lose significance, but most of them are still significant. We further investigate these results in Figures 36, 37 and 38, where we vary the bandwidth to check if there is still a consistency to the point-estimates. The number of candidates is closer to zero, with some negative estimates even when the bandwidth is larger than 10%. This could be explained by the fact the estimates for female candidates are sometimes negative. The effects for the number of male candidates are still positive. Although these estimates for candidates and male candidates are not significant for different bandwidths. Moving on to the votes of male candidates, the average votes for males and HHI of votes or males are still not significant, but they are consistently negative. The variables associated with female electoral performance (number of women elected, proportion of women elected and probability of electing at least one woman) are all positive and they have many estimates that are significant. These results are very similar to the ones without year and state fixed effects.

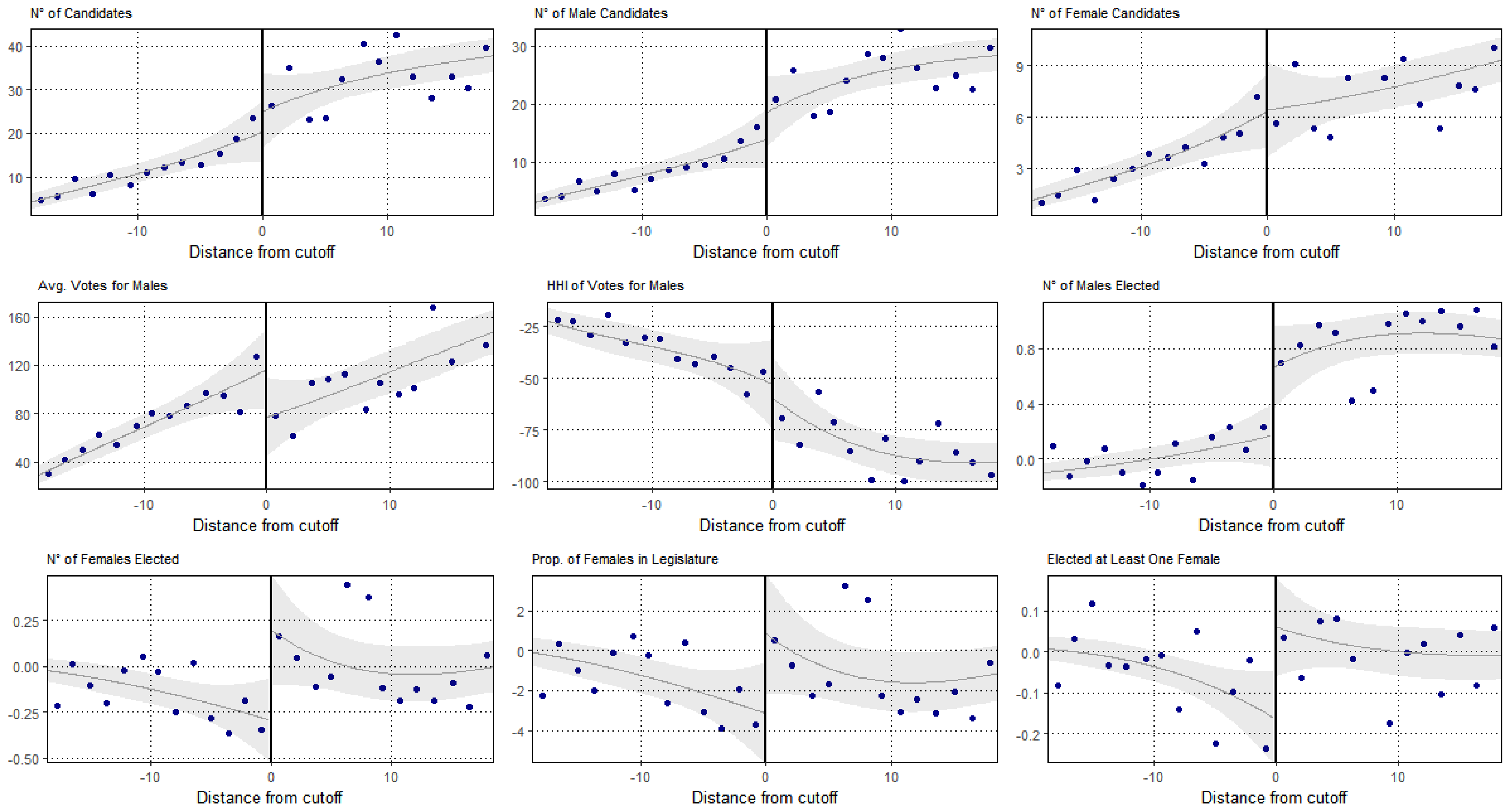
As for policy outcomes, many of them are very similar to the exercise without year and state fixed effects. Some differences are worth noting, however. The number of live births and community assistance expenditures is not significant in most specifications, but the point-estimates of the expenditures with community assistance are still consistent, have the same signs are in the case without controls and display some significant results at higher bandwidths. The number of live births, is less consistent, with values alternating between positive and negative for lower bandwidths. The results for the SAEB score in portuguese indeed lose significance, but the estimates are still overall positive and consistent as the bandwidth varies. The results for child education expenditures, however, are not significant and while most estimates are positive, there are a few negative estimates. As such, although the SAEB score in portuguese lost significance in the main specification, the point estimates a consistency positive and of similar magnitude. However, the number of live births and child education expenditures do not seem to be very robust to the introduction of year and state fixed effects.

1.9.2 Validity Test - FPM and Salary Thresholds

As previously mentioned, the salary of legislators and the thresholds for the coefficients of the Municipal Participation Fund (*Fundo de Participação dos Municípios - FPM*) also establishes thresholds based on population intervals. This may pose a problem for our identification strategy if these thresholds set by other rules are too close to the one being explored in this chapter. In this subsection we try to address this concern by restricting our sample in a way our observations do not cross any of these other thresholds. Our cutoff of interest is the cutoff of 47,620 inhabitants. The FPM has two cutoffs close to this one: one at 44,149 inhabitants and another set at 50,940. The Salary Rule also has a cutoff set at 50,000 inhabitants. As such, we restrict our sample to include only municipalities with populations between 44,149 and 50,000 inhabitants. This ensues that the only relevant cutoff is the one set at 47,620. A caveat of this approach is that we do not have many observations left for such a narrow interval, which leaves us with only 136 observations.

Nevertheless, we estimate the regressions for political entry and electoral outcomes within this interval. The results can be seen at Table 19. As expected, due to the low number of observations, many of the estimates have unusually large standard errors, which leaves most of the estimates not significant at any level. As for the magnitude and sign of the estimates, most of them are in line with the ones in the main table. In particular, the mean comparisons are all in line with our main results. However, when we look at the Local Linear estimates, it's noticeable that the number of candidates, number of male candidates and number of female candidates are all negative, in contrast to what we had found previously. It is important to notice, though, that these estimates all have high standard deviations. Aside from that, the other variables are also

Figure 32 – Main Results - Political Variables while controlling for year and state fixed effects



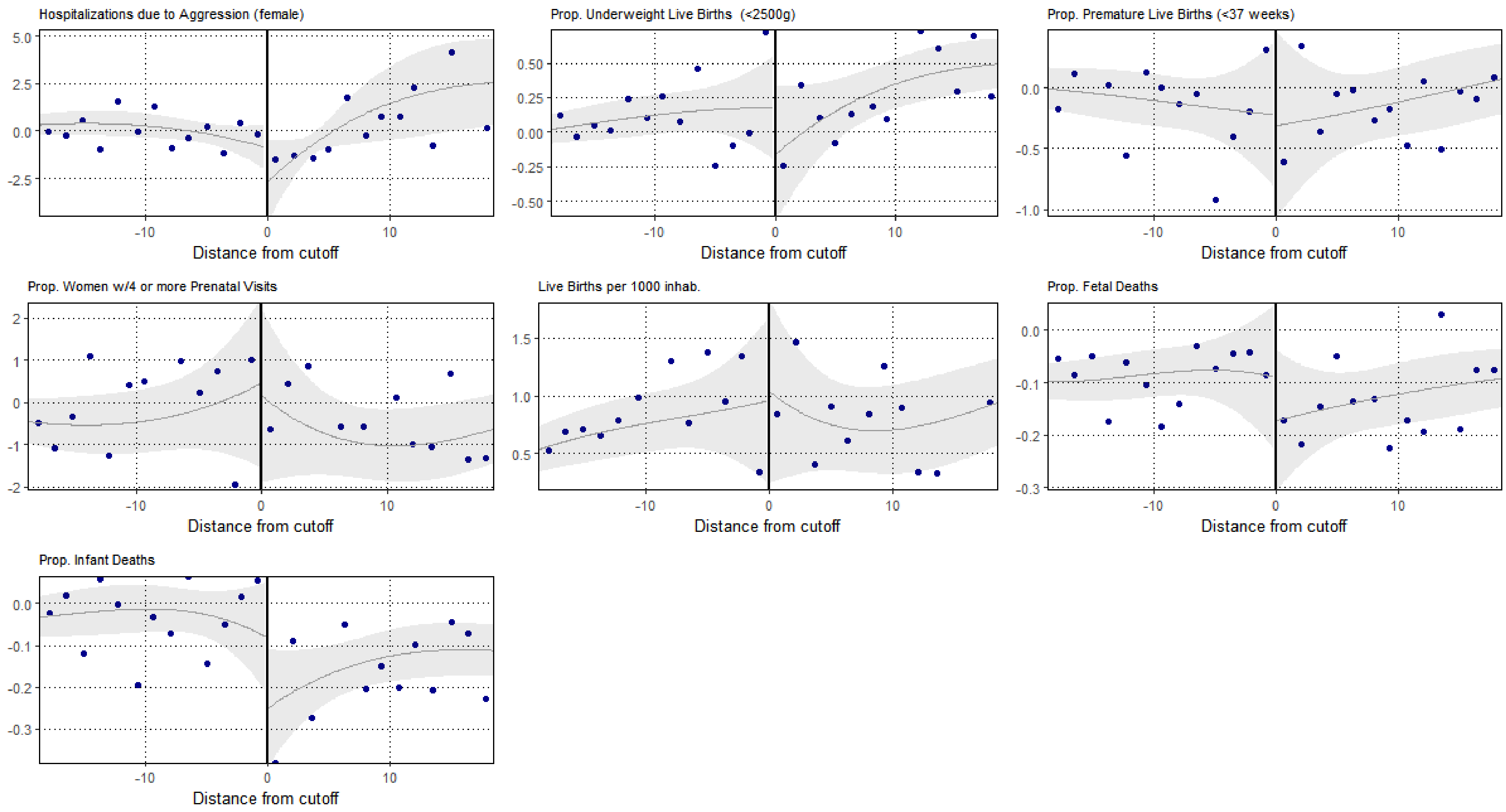
Notes: RD graph for our main results for the political variables controlling for state and year. We first regress each variable against year and state fixed effects and then plot the residuals. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 15 – Effects of Local Council Size on Political Variables while controlling for year and state fixed effects

		Mean Comparison		Local Linear								
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW	
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
N° of	84.012	289	11.2032***	456	0.190	0.4952	289	1.5715	650	2.3309	1802	7.1072
Candidates		289	(3.3645)	456	0.190	(7.0667)	289	(12.8706)	650	(8.5202)	1802	(5.4964)
N° of Male	64.184	289	9.3736***	403	0.173	0.6693	289	2.2726	650	3.3574	1802	6.6441*
Candidates		289	(2.4493)	403	0.173	(5.2844)	289	(8.8862)	650	(6.0527)	1802	(3.9472)
N° of Female	19.828	289	1.8297+	369	0.156	-0.6532	289	-0.7011	650	-1.0265	1802	0.4654
Candidates		289	(1.1141)	369	0.156	(2.615)	289	(4.288)	650	(2.7802)	1802	(1.8189)
Avg. Votes	333.386	289	-11.0929	607	0.236	-21.2665	289	-48.1807	650	-35.5333	1802	-38.0241+
for Males		289	(14.5084)	607	0.236	(25.9048)	289	(49.1963)	650	(35.565)	1802	(23.349)
HHI of Votes	240.962	289	-29.4395***	423	0.178	-18.4584	289	-40.0112	650	-23.1041	1802	-15.0796
for Males		289	(8.9222)	423	0.178	(17.7106)	289	(31.7616)	650	(21.3614)	1802	(14.3162)
N° of Males	8.147	289	0.711***	152	0.073	0.5251+	289	0.4567	650	0.5907**	1802	0.5913***
Elected		289	(0.1254)	152	0.073	(0.3307)	289	(0.3772)	650	(0.2833)	1802	(0.1933)
N° of Females	0.853	289	0.289**	353	0.150	0.3568	289	0.5433+	650	0.4093+	1802	0.4087**
Elected		289	(0.1254)	353	0.150	(0.2667)	289	(0.3772)	650	(0.2833)	1802	(0.1933)
% Legislature Fem.	9.475	289	1.9485+	353	0.150	2.8439	289	4.5853	650	3.2829	1800	3.2319+
		289	(1.3149)	353	0.150	(2.7915)	289	(4.009)	650	(2.9893)	1800	(2.012)
Elected at Least	0.558	289	0.1817***	257	0.115	0.1466	289	0.3333*	650	0.2174*	1802	0.1957**
One Female		289	(0.0593)	257	0.115	(0.1381)	289	(0.1808)	650	(0.1261)	1802	(0.0867)
Polynomial			Zero		One			One		One		One

Notes: In this table, we report our preferred RDD estimates for the effect of the council size on political variables associated with entry of citizens to politics and its electoral results while controlling for year and state fixed effects. We use municipalities as observations. Columns (1)-(4) reports our estimates. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW.* reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 33 – Main Results - Policy Outcomes (health) while controlling for year and state fixed effects



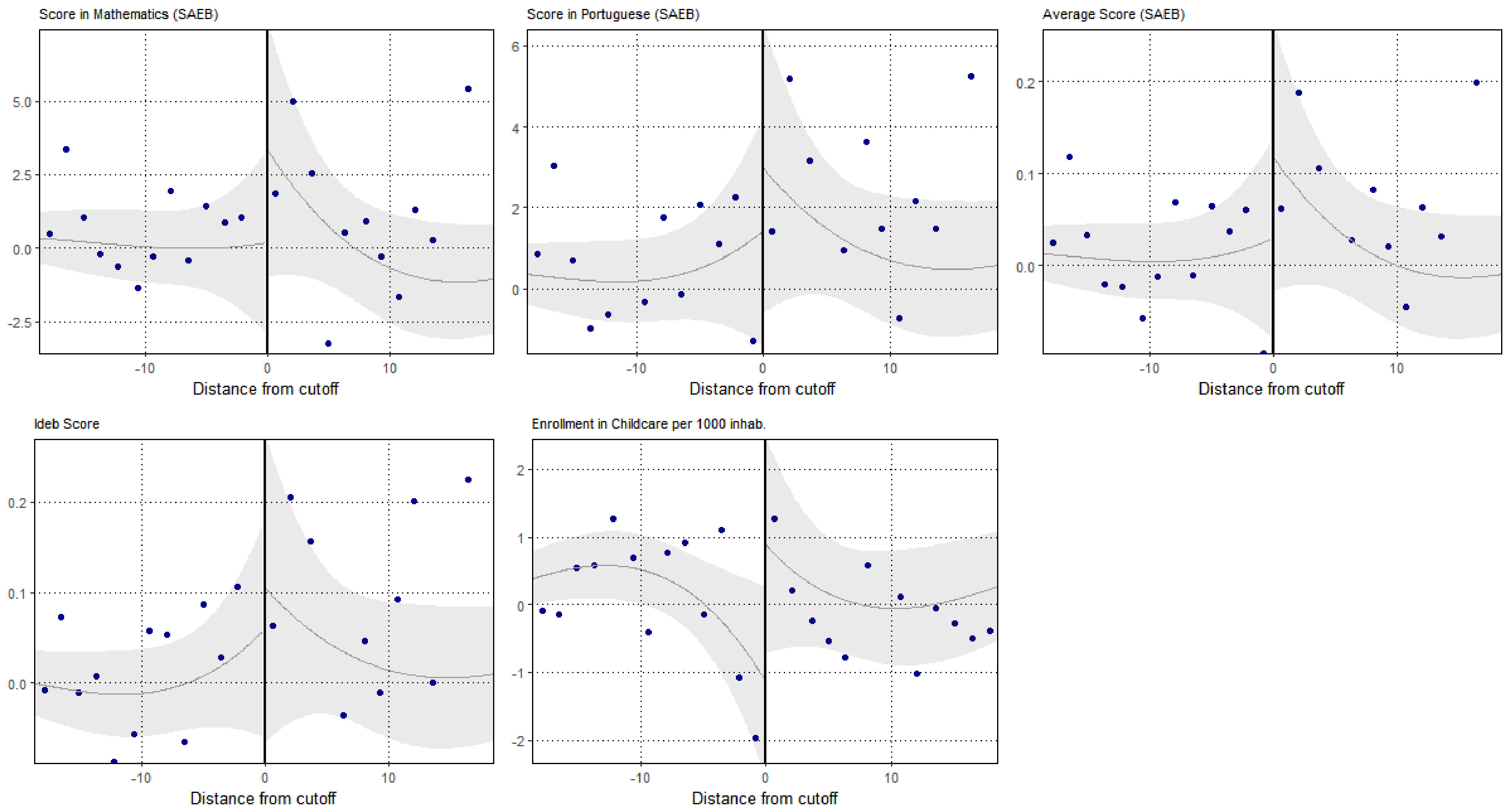
Notes: RD graph for our main results for the health outcome variables while controlling for state and year. We first regress each variable against year and state fixed effects and then plot the residuals. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 16 – Effects of Local Council Size on Political Variables - Health Outcomes while controlling for year and state fixed effects

		Mean Comparison		Local Linear								
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW	
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
Hospitalizations	2.718	178	-0.6829	153	0.111	-0.3097	178	0.3019	368	-0.8795	879	-2.8727***
due to Aggression (fem.)		178	(0.5171)	153	0.111	(1.0167)	178	(1.1558)	368	(0.9992)	879	(1.104)
% Underweight	7.469	292	0.0491	485	0.195	-0.1111	292	-0.5604	655	-0.155	1812	-0.349
Live Births (<2500g)		292	(0.199)	485	0.195	(0.3689)	292	(0.5035)	655	(0.4345)	1812	(0.3184)
% Premature	6.983	292	0.2029	237	0.101	-0.5768	292	-0.4781	655	-0.4004	1812	0.3067
Live Births (<37 weeks)		292	(0.303)	237	0.101	(0.8103)	292	(0.9816)	655	(0.757)	1812	(0.5066)
% Women w/4 or	87.964	292	-0.9719	239	0.104	0.7617	292	-0.8902	655	1.4057	1812	0.1348
more Prenatal Visits		292	(1.0696)	239	0.104	(2.2977)	292	(2.2587)	655	(2.3049)	1812	(1.5056)
N° of Live Births	16.874	292	-0.1014	381	0.161	0.0452	292	0.6429	655	-0.2375	1812	-0.2191
per Capita		292	(0.3442)	381	0.161	(0.6649)	292	(0.8237)	655	(0.707)	1812	(0.5258)
% Fetal Deaths	1.226	292	-0.0881	571	0.221	-0.1865*	292	-0.1934	654	-0.2241+	1799	-0.095
		292	(0.0676)	571	0.221	(0.1133)	292	(0.2184)	654	(0.1466)	1799	(0.0984)
% Infant Deaths	1.616	292	-0.0997+	127	0.064	-0.4908***	292	-0.4139***	654	-0.4236***	1808	-0.2307**
		292	(0.0654)	127	0.064	(0.1426)	292	(0.1497)	654	(0.126)	1808	(0.0957)
Polynomial			Zero		One			One		One		One

Notes: In this table, we report our preferred RDD estimates for the effect of the council size on health outcomes while controlling for year and state fixed effects. We use municipalities as observations. Sample sizes in columns (3) and (4) for Hospitalization due to Aggression (Female), Prop. of Fetal Deaths and Infant Deaths are, respectively, 368 and 879, 655 and 1801 and 655 and 1810. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW*. reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 34 – Main Results - Policy Outcomes (education) while controlling for year and state fixed effects



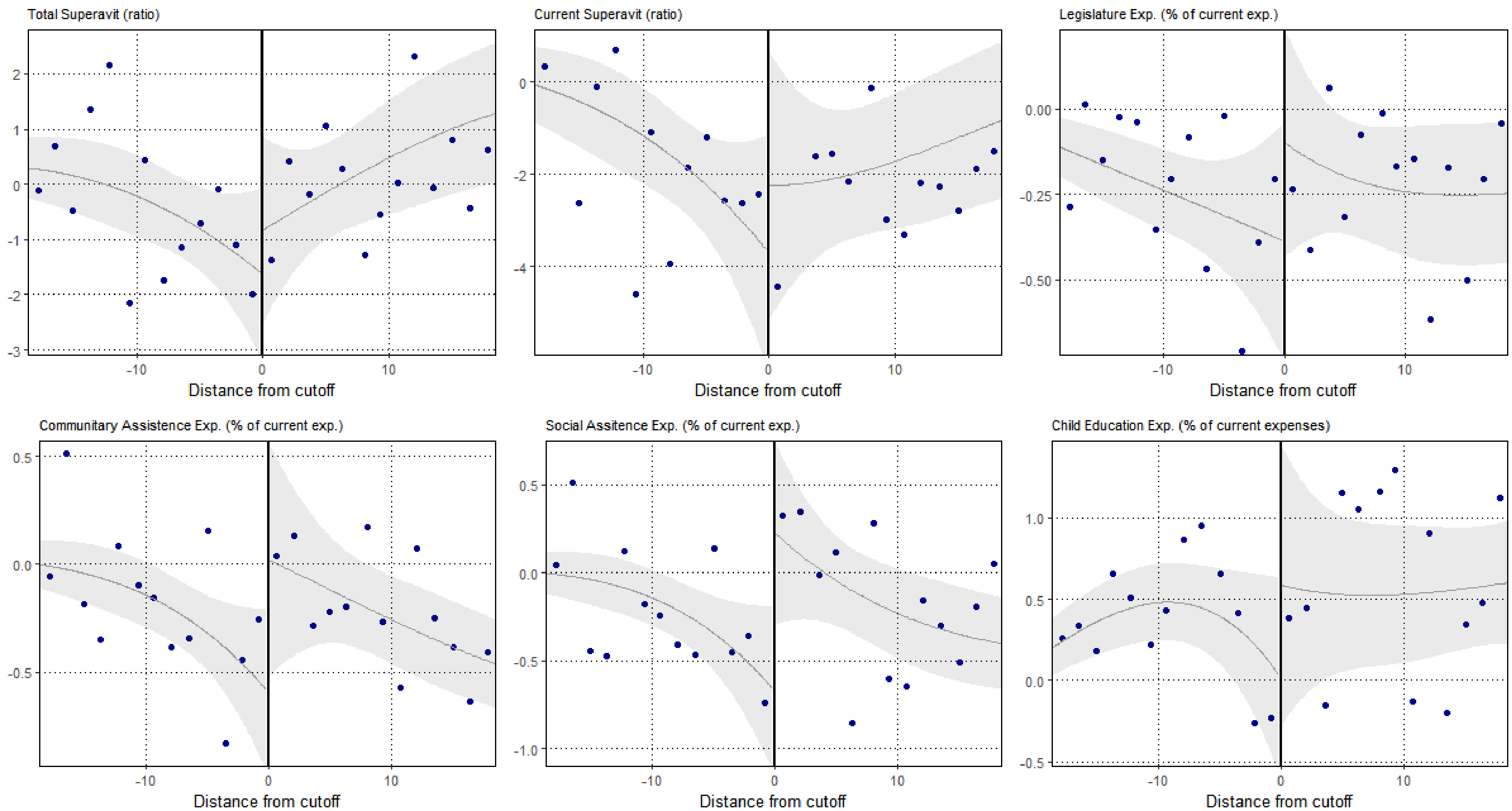
Notes: RD graph for our main results for the educational outcomes while controlling for state and year. We first regress each variable against year and state fixed effects and then plot the residuals. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 17 – Effects of Local Council Size on Political Variables - Educational Outcomes while controlling for year and state fixed effects

		Mean Comparison		Local Linear							
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
Score in	193.281	288	1.9006	134	0.067	8.9327*	288	8.1712+	644	7.7458*	1772	4.8514*
Mathematics (SAEB)		288	(1.8446)	134	0.067	(5.3569)	288	(5.5686)	644	(4.0319)	1772	(2.8089)
Score in	175.540	288	1.2395	203	0.091	4.0679	288	5.3418	644	4.3053	1772	2.5883
Portuguese (SAEB)		288	(1.6168)	203	0.091	(3.9931)	288	(4.7788)	644	(3.4485)	1772	(2.3812)
Avg. Score	4.844	288	0.0588	267	0.118	0.3115**	288	0.2531	644	0.2261+	1772	0.1396+
(SAEB, standarized)		288	(0.0638)	267	0.118	(0.1462)	288	(0.1916)	644	(0.1384)	1772	(0.0958)
Ideb Score	4.207	288	0.0463	177	0.082	0.1369	288	0.2425	644	0.3094**	1772	0.1073
		288	(0.0745)	177	0.082	(0.1836)	288	(0.2148)	644	(0.1498)	1772	(0.11)
Enrollment in Childcare	9.608	292	-0.2645	287	0.124	3.1152**	292	4.0681**	655	2.3168+	1812	0.9457
per Capita		292	(0.7541)	287	0.124	(1.4568)	292	(1.8172)	655	(1.4307)	1812	(1.0794)
Polynomial			Zero		One			One		One		One

Notes: In this table, we report our preferred RDD estimates for the effect of the council size on educational outcomes while controlling for year and state fixed effects. Columns (1)-(4) reports our estimates. Sample sizes in columns (3) and (4) for Enrollment in Childcare per Capita are, respectively, 656 and 1814. We use municipalities as observations. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW.* reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 35 – Main Results - Policy Outcomes (fiscal) while controlling for year and state fixed effects



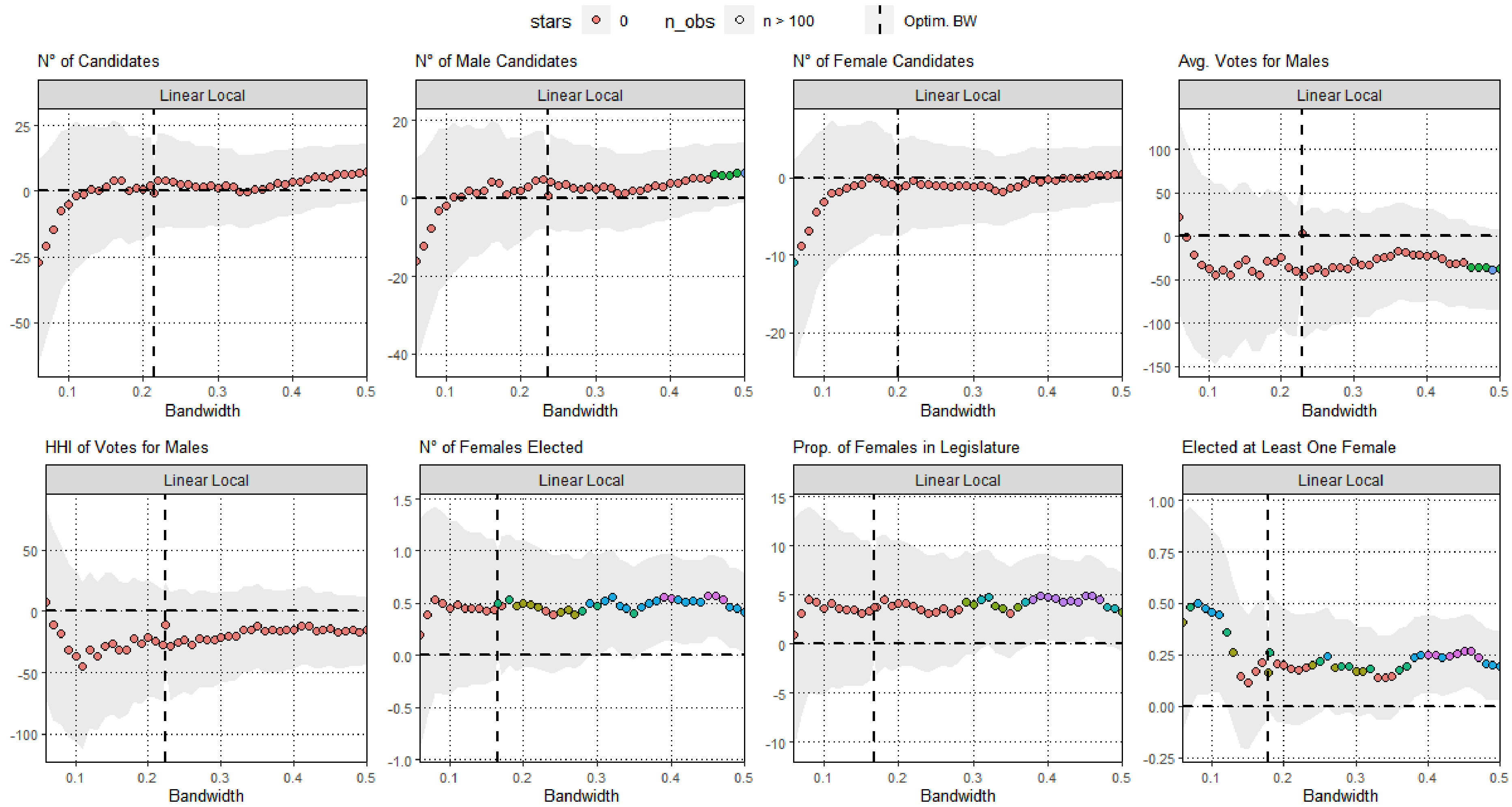
Notes: RD graph for our main results for the fiscal outcome variables while controlling for state and year. We first regress each variable against year and state fixed effects and then plot the residuals. In the vertical line at the middle of the graph, we have the first cutoff defined by the rule that assigns the number of seats to each municipality. The horizontal axis displays the distance from cutoff in terms of thousands of inhabitants and the vertical axis displays the average number of seats in each bin. We divided the assignment variable into 26 bins. The figure also reports a third-order polynomial regression fitted separately on each side of the cutoff. The gray area surrounding the fitted line is the 95% confidence interval of the polynomial. The procedure to construct this graph follows (LEE; LEMIEUX, 2010). We pooled the election of 2004 and 2008.

Table 18 – Effects of Local Council Size on Political Variables - Fiscal Outcomes while controlling for year and state fixed effects

		Mean Comparison		Local Linear									
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW	
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
Total Superavit	103.052	286	0.0277	122	0.062	0.6452	286	0.9857	644	-0.0188	1774	0.6575
(ratio)		286	(0.7898)	122	0.062	(1.7037)	286	(2.0958)	644	(1.4945)	1774	(1.12)
Current Superavit	120.559	286	-0.6339	165	0.077	-3.1759	286	-1.5401	644	-1.961	1773	1.0573
(ratio)		286	(1.2975)	165	0.077	(2.594)	286	(3.1145)	644	(2.3863)	1773	(1.9058)
Legislature Expenditures	2.721	286	0.1567	364	0.157	0.183	286	0.2125	644	0.0841	1773	0.2969
(% of curr. expenses)		286	(0.1676)	364	0.157	(0.2755)	286	(0.4169)	644	(0.3021)	1773	(0.2375)
Communitary Assist. Exp.	1.930	286	0.2696	498	0.204	0.4258	286	0.1869	644	0.4969	1773	0.5621+
(% of curr. expenses)		286	(0.238)	498	0.204	(0.4166)	286	(0.5782)	644	(0.4854)	1773	(0.3494)
Social assistance Exp.	3.946	286	0.399*	268	0.120	1.1**	286	1.0549+	644	0.8772*	1773	0.6623*
(% of curr. expenses)		286	(0.2332)	268	0.120	(0.5137)	286	(0.6425)	644	(0.4891)	1773	(0.355)
Child Education Exp.	3.357	286	-0.1264	276	0.123	0.3199	286	0.2374	644	0.4247	1773	0.0041
(% of curr. expenses)		286	(0.4322)	276	0.123	(0.6321)	286	(0.852)	644	(0.6282)	1773	(0.5206)
Polynomial			Zero		One			One		One		One

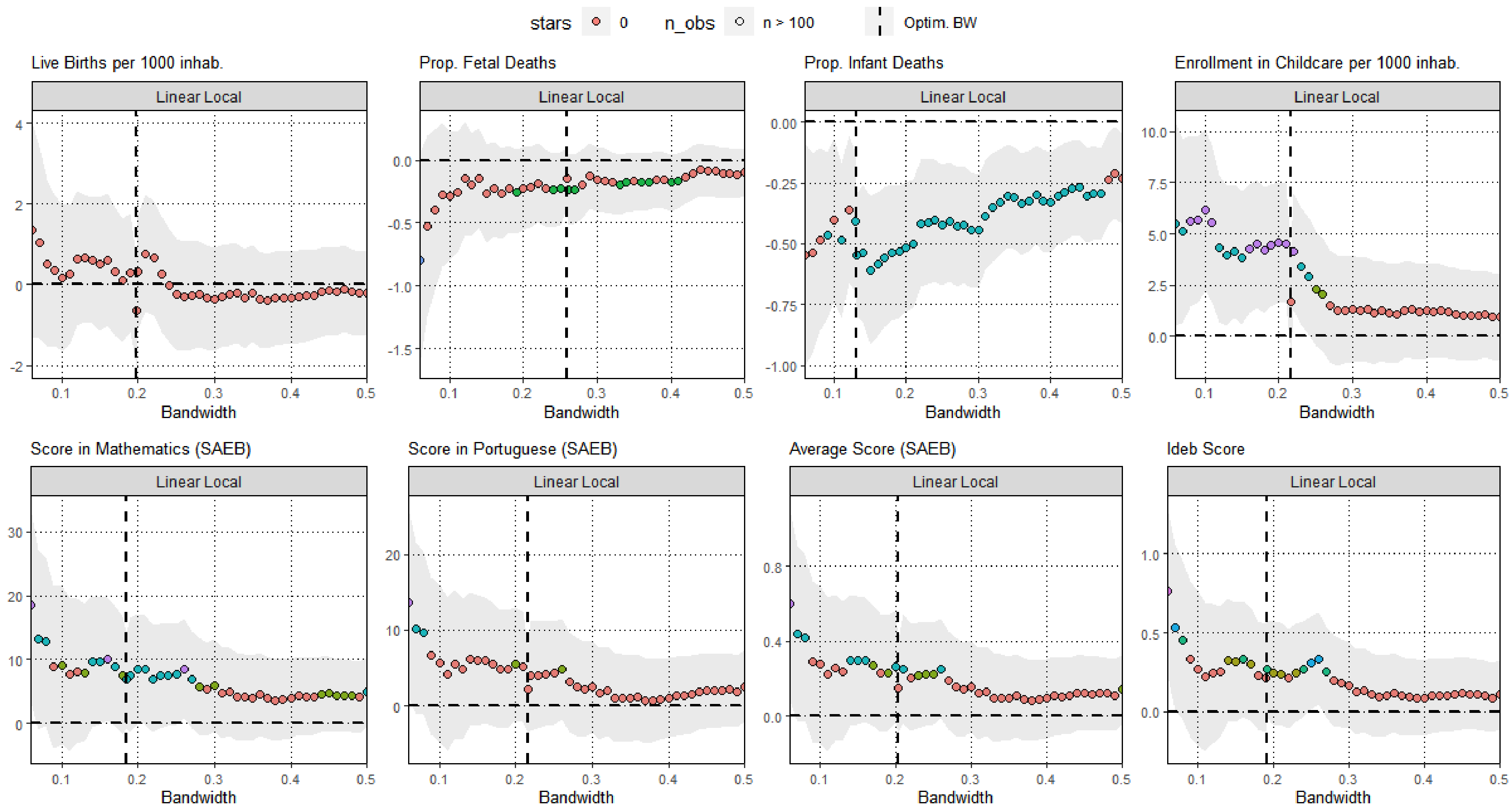
Notes: In this table, we report our preferred RDD estimates for the effect of the council size on fiscal outcome variables associated with entry of citizens to politics and its electoral results while controlling for year and state fixed effects. Columns (1)-(4) reports our estimates. We use municipalities as observations. Sample sizes in columns (3) and (4) for Total Superavit and Current Superavit are, respectively, 451 and 1244, and 637 1764. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW*. reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 36 – Sensitivity analysis - Political Variables while controlling for year and state fixed effects



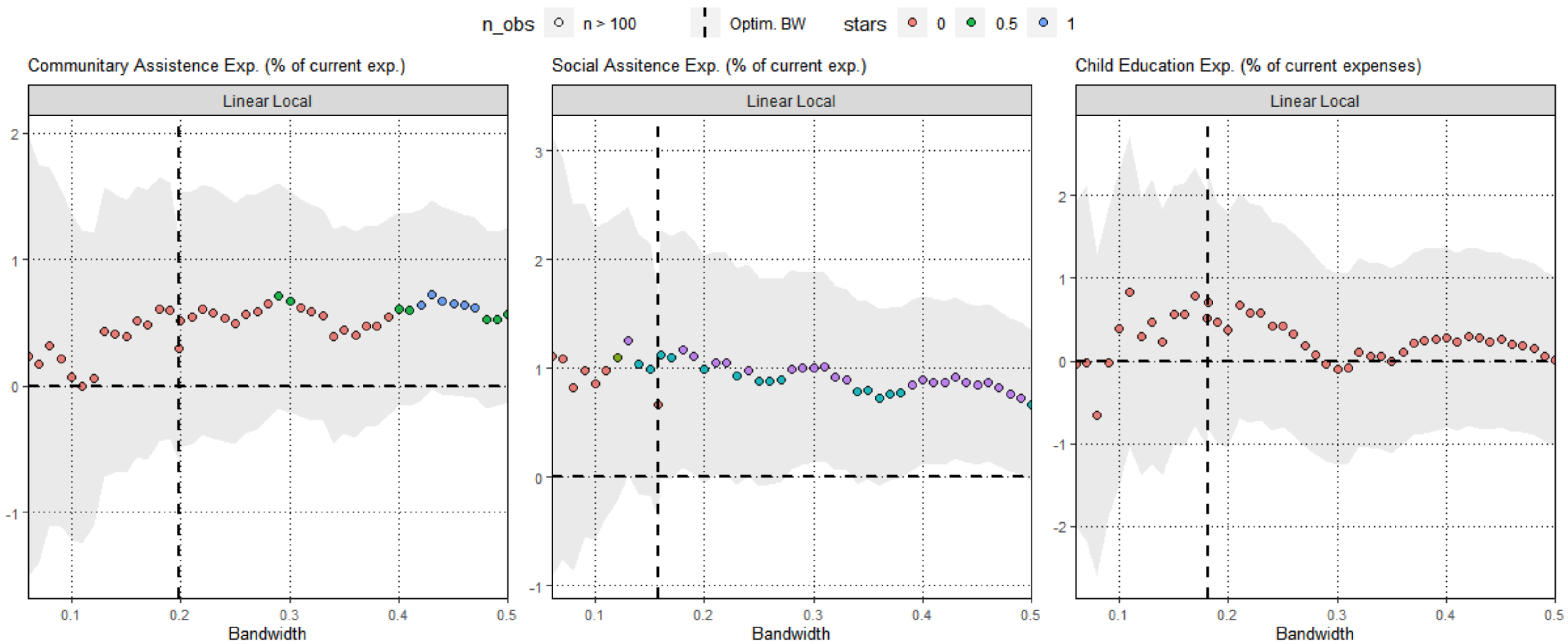
Notes: In this figure, we display a sensitivity analysis for our main RDD specification for political outcomes while controlling for year and state fixed effects. We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

Figure 37 – Sensitivity analysis - Policy Outcomes while controlling for year and state fixed effects (1/2)



Notes: In this figure, we display a sensitivity analysis for our main RDD specification for policy outcomes while controlling for year and state fixed effects. We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

Figure 38 – Sensitivity analysis - Policy Outcomes while controlling for year and state fixed effects (2/2)



Notes: In this figure, we display a sensitivity analysis for our main RDD specification for policy outcomes while controlling for year and state fixed effects. We plot the point-estimates for RD regressions at varying bandwidths, ranging from 1% to 50% at intervals of 1%. For each variable, we estimate a Local Linear Regression with bias-corrected confidence intervals following (CALONICO; CATTANEO; TITIUNIK, 2014) and kernel. 95% intervals are represented by the gray areas. The colors of each point-estimate denote the significance of the point estimates according to the number of stars in the legend. 0.5 stars denotes $p < 0.15$, 1 star denotes $p < 0.1$, 2 stars denotes $p < 0.05$ and 3 stars denotes $p < 0.01$. 0 stars denote results that are not significant at any of the levels mentioned above. The dashed vertical line marks the point-estimate from the regression using a MSE-optimal bandwidth. The shape of each point-estimate represents the number of observations (n).

mostly in line with the main tables in the Linear Local Estimates. As such, many of our results survive this restriction, leading us to believe our results indeed are coming from the cutoff of 47,620 rather than other cutoffs nearby.

Table 19 – Validity Test - Political Variables within the interval set by the salary rule and FPM ($Pop_{t-1} \in (44149, 50000)$)

		Mean Comparison		Local Linear								
		12.5% BW		Optimal BW			12.5% BW		25% BW		50% BW	
Variables	Mean	Obs.	(1)	Obs.	BW	(2)	Obs.	(3)	Obs.	(4)	Obs.	(5)
N° of Candidates	87.5000	134	12.8182*	35	0.0175	-5.9268	134	-14.1829	134	-14.1829	134	-14.1829
		134	(6.6067)	35	0.0175	(47.6722)	134	(20.1152)	134	(20.1152)	134	(20.1152)
N° of Male Candidates	66.6111	134	10.548**	35	0.0171	-5.3059	134	-9.4897	134	-9.4897	134	-9.4897
		134	(4.6831)	35	0.0171	(34.1032)	134	(13.8909)	134	(13.8909)	134	(13.8909)
N° of Female Candidates	20.8889	134	2.2702	24	0.0130	-5.2493	134	-4.6932	134	-4.6932	134	-4.6932
		134	(2.1301)	24	0.0130	(16.9175)	134	(6.6366)	134	(6.6366)	134	(6.6366)
Avg. Votes for Males	329.0517	134	-25.3711	32	0.0168	4.2063	134	-24.9191	134	-24.9191	134	-24.9191
		134	(24.8749)	32	0.0168	(176.623)	134	(69.989)	134	(69.989)	134	(69.989)
HHI of Votes for Males	235.1560	134	-34.2344**	35	0.0182	-66.9758	134	-26.1124	134	-26.1124	134	-26.1124
		134	(15.8137)	35	0.0182	(112.6322)	134	(53.0499)	134	(53.0499)	134	(53.0499)
N° of Males Elected	8.0889	134	0.6611***	15	0.0079	-0.7038	134	0.2891	134	0.2891	134	0.2891
		134	(0.1932)	15	0.0079	(1.212)	134	(0.5286)	134	(0.5286)	134	(0.5286)
N° of Females Elected	0.9111	134	0.3389*	15	0.0079	1.7038	134	0.7109	134	0.7109	134	0.7109
		134	(0.1932)	15	0.0079	(1.212)	134	(0.5286)	134	(0.5286)	134	(0.5286)
Prop. of Females in Legislature	10.1235	134	2.3765	18	0.0088	15.2449+	134	6.4674	134	6.4674	134	6.4674
		134	(1.985)	18	0.0088	(10.0467)	134	(5.6191)	134	(5.6191)	134	(5.6191)
Elected at Least One Female	0.6111	134	0.0707	23	0.0121	1.0286**	134	0.6541**	134	0.6541**	134	0.6541**
		134	(0.087)	23	0.0121	(0.4764)	134	(0.2696)	134	(0.2696)	134	(0.2696)
Polynomial			Zero		One			One		One		One

Notes: In this table, we report our preferred RDD estimates for the effect of the council size on political variables associated with entry of citizens to politics and its electoral results while restricting the sample to be within one of the intervals of the salary rule and FPM thresholds ($Pop_{t-1} \in (44149, 50000)$). We use municipalities as observations. Columns (1)-(4) reports our estimates. The *Mean* is an average before the cutoff point using a bandwidth of 12.5%. Column *BW.* reports the optimal bandwidth for the Linear Local estimates of column (1). We followed the procedure by *Calonico et al. (2014)*. In columns (2)-(4) we, respectively, change the polynomial order to zero, use a fixed bandwidth equal to 25% of the threshold and double the fixed bandwidth previously used. Standard errors are in parenthesis. $^{+}p < 0.15$, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

2 FEMALE POLITICAL REPRESENTATION AND PUBLIC POLICY: IS THERE ANY IMPACT? EVIDENCE FROM BRAZIL'S MUNICIPAL LEGISLATIVE HOUSES

2.1 Introduction

A known stylized fact in representative democracies is that elected representatives are often very different from the ones who elected them, leading certain groups to be systematically underrepresented in the political arena. One of such groups are women. Despite being roughly half of the population, women occupied only 25.6% of seats in national parliaments worldwide in 2020, according to the Inter-Parliamentary Union (IPU). In Brazil, their under-representation is even more noteworthy, with only around 14.6% of seats being held by women also in 2020. This disparity brings up two major questions: why does it happen and what are its impacts on public policy?

As for the first question, there is a myriad of factors that may affect women's access and willingness to participate in the political arena, ranging from cultural and socio-economic factors, such as views on gender-based social roles, to political factors, such as political rights to women and political regime. Tremblay (2007), for example, finds a positive correlation between the voting system and participation of women in parliaments in countries that did not have been democratic for long periods of time. In countries with longer democracies, the cultural aspects of gender roles have a more visible correlation. While none of this is causal in their study, this highlights some potential aspects that may influence and be influenced by a higher percentage of women in politics.

On the note of how the political structure might enable more women to participate in politics, Bhalotra, Clots-Figueras and Iyer (2018) find, using data from state legislature elections in India, that there is a feedback effect in which electing women to state legislatures leads to subsequent political participation of women. They also notice, though, that this increase in subsequent female participation comes mostly from female incumbents re-contesting, and in states with more entrenched gender biases, there is a backlash effect, with a decline in the entry of new female candidates.

Given these different factors that may hinder or enhance female participation in politics, we turn to the second question of how does the identity of representatives influences policymaking. If we go by the classic work of Downs (1957), we should not expect for the identity of the politician to matter, since policy decisions are a reflection of policy preferences of the median voter. However, other models do take into account the politician's preference in their policy decisions, such as is the case in the citizen-candidate models based on the work of Besley and Coate (1997). And this is indeed the case when we take into account the politician's characteristics. Pande (2003) and Chattopadhyay and Duflo (2004) studied how mandated political reservations in India affects policy decisions. In the first study, it was found that political reservations for underrepresented minorities increased transfers for these groups. Likewise, the latter study focused on reservations for women to the position of village council head and found that politicians invested more in infrastructure that had more direct relevance for the needs of their own genders. In another study, Bhalotra and Clots-Figueras (2014) found a positive causal relationship between female political participation in state legislatures and antenatal and childhood health outcomes in India. In Brazil, Brollo and Troiano (2016) find that electing female mayors leads to better health outcomes. They also find that these women tend to be more apt at attracting federal discretionary transfers and are less likely to be involved in cases of corruption or administrative irregularities.

This paper aims to contribute to this second line of investigation that deals with the impacts of electing underrepresented minorities in terms of public policy. More specifically, we are interested in the impacts of electing women to local legislative houses in Brazil in a context of close electoral races. In close elections, whether a candidate barely won or barely lost the election can be viewed as if it were a random event. As such, it is possible to estimate the causal effects of this electoral victory (LEE; MORETTI; BUTLER, 2004; LEE, 2008). In Brazil, this strategy has been employed to study questions related to gender in a context of mayoral elections (BROLLO; TROIANO, 2016; ARVATE; FIRPO; PIERI, 2021). This study, on the other hand, focuses on the legislative branch - we are interested in the effects of electing a female councilperson to a municipal legislative house. However, Brazil utilizes an open-list proportional representation system to choose the members of the legislative at municipal level. That means we have candidates who barely won and barely lost for each party or coalition that participated in the elections, which can then be compared with each other. This approach is similar to the one employed by Boas, Hidalgo and Richardson (2014), which also examines close elections to study the relationship between government contracts and campaign donations in a context of legislative elections (but in their case, to federal-deputy elections). However, this and other studies that exploit close elections also employ a Regression Discontinuity Approach (RDD). Here we employ a similar but simpler approach. Instead, we have selected pairs formed by a man and a woman candidates with the lowest difference in their vote shares for each municipality. Due to the uncertainty on the precise ranking of candidates for those in the margin and the uncertainty in the number of seats each party will receive, we argue the victory margin of the last winner is exogenous for this subsample of candidates.

It is also important to notice that this approach of exploiting close elections allows us to examine the effects of electing a woman to the local legislative without relying on exogenous changes in variables associated with the political structure, such as the number of seats or quotas. This is important because, as we mentioned before, a myriad of factors may influence both women's participation and their ability to influence the political process. As such, both the characteristics and potential ability to impact public policy of female candidates elected under those circumstances might be tied to the interaction between the election of women and the political institutions that influenced their participation.

We find that an electoral victory of a woman has some impact in health outcomes by a decrease in the proportion of underweight live births, as well as a seemly increase in education expenditures - but with no increase in expenditures on child education nor educational outcomes. It is important to notice, however, that we examined many potential policy outcomes. As such, these results could be false positives and we indeed cannot rule out this possibility in the current version of this paper. This lack of robust effects contrasts with other studies in the literature, which find large and significant effects derived from an increase in the participation of women in politics. One possible explanation for this result or lack thereof is that the proportion of votes of a candidate might be correlated to their political influence or other variables that impact their ability to influence policy decisions. Since we focus on barely winners and barely losers, the winners will mostly have a very low vote count. As such, they might be very different characteristics from the most voted candidates.

The remainder of the paper is structured as follows. In section 2 we present an overview of Brazil's municipal elections in regards to how exactly candidates are chosen in the context of open-list proportional-representation in Brazil and also describe the data. In section 3, we present our identification strategy and working sample. In section 4, we examine covariate balance and discuss potential issues that could invalidate our identification strategy. In section 5, we present our main results for our working sample. In section 6 we perform a few robustness checks. Namely, we investigate how our results change when taking into account multiple hypotheses and when we utilize a less restricted sample. Section 7 summarizes our findings.

2.2 Institutions and Data

2.2.1 Brazilian municipal elections

Brazil is a federative republic whose political and administrative organization comprises the Union, the states, the Federal District and the municipalities, each being treated as autonomous units according to Brazil's 1988 Federal Constitution. Our focus in this study is the municipal administration, which are administrative divisions of the states. There are over 5000 municipalities across 26 states. Each municipality has a local government composed of an executive and a legislative body. The executive is led by a mayor and the legislative is run by a municipal chamber. Both mayors and councilpersons are elected by the population of their respective municipality every four years. The size of legislative houses vary according to the population of each municipality, but the exact rule to determine the number of seats has changed over the years [1]. For this paper, we will be mainly concerned with the municipal elections of 2004, 2008 and 2012.

Members of the municipal chambers (*vereadores*) are elected through an open-list proportional representation system. The procedure is detailed in the Electoral Code, Law n°4737/65. Arts. 106 and 107 define, respectively, the Electoral Quotient and the Party Quotient. The former is defined as:

$$Q_e = \frac{V_v}{S} \tag{2.1}$$

Where Q_e is the Electoral Quotient, V_v is the number of total valid votes in the election and S is the number of seats to be filled. Then, for each party, it is calculated a Party Quotient, given by:

$$Q_p = \frac{V_p}{Q_e} \tag{2.2}$$

Where Q_p is the Party Quotient and V_p is the total number of votes obtained by the party. Each party then receives an amount of seats equal to the integer part of the Party Quotient. After the procedure is over, there may still be some seats remaining. In this case, Art. 109 of the electoral law determined that the remainders should be distributed in a similar fashion to the D'Hondt method. Let the quotient of each party be R_p, then:

$$R_p = \frac{Q_p}{(s+1)} \tag{2.3}$$

Where s is the number of seats obtained by the respective party. The party with the highest quotient then receives the remaining seat. The procedure is then repeated until all seats have been allocated to some party or coalition.

As for which candidates will receive the seats, each party or coalition must announce a list of candidates before elections. Voters then can vote for either candidates or parties. Both votes in parties and candidates count towards the total votes of the party, but when a voter chooses a candidate, they are also choosing the placement of that candidate in the party list. Candidates are then ranked by their vote counts with the one with the highest vote count being ranked first. Seats obtained by the party are then distributed according to this ranking[2]. Once elections are over, the elected candidate become city councilors and are then able to influence public policy in various manners. They participate in the elaboration of the municipality's organic law, including the elaboration of the municipality budget; legislate upon matters of local interest and supervise the executive branch's activities.

[1] In the 2004 municipal elections, the Supreme Electoral Court (*Supremo Tribunal Eleitoral* - STE) issued Resolution n°21.702/2004, which defined the number of seats for each municipality according to population thresholds. This rule was valid for the 2004 and 2008 elections but was later changed in 2009 by the Constitutional Amendment n°58/2009 This Amendment defined new population thresholds and instead of setting an exact number of seats for each municipality, it only defined the maximum amount of seats each municipality could have. As such, the legislative house itself had the agency to determine its own size as long as it did not exceed the caps defined by the amendment.

[2] In 2015 an electoral reform established the so-called Electoral threshold (*Cláusula de Barreira*), which determined that candidates from a party or coalition need a minimum vote count equal to 10% of the electoral quotient for them to be able to get elected. In case this minimum vote count isn't reached, the candidate won't be able to hold office even if it party of the coalition was able to obtain enough seats (Art. 108 of Law n° 4737/65). However, our analysis examines only elections that happened before this change was put in place.

2.2.2 Electoral and Municipal Data

For this paper, we focus on three legislative mandates: 2005-2008, 2009-2012 and 2013-2016. Our dataset comes from various publicly available government sources. For these mandates, we collected electoral data from the Superior Electoral Court (*Superior Tribunal Eleitoral - TSE*) for each corresponding election: 2004, 2008 and 2012. This is the main dataset from which we draw information from the male and female candidates.

The TSE stores a wide array of data on electoral information, which allows us to gather information on both elected candidates and non-elected candidates, including their gender, vote count, political affiliation, educational level and so forth. An issue with the datasets from TSE is that they are not perfectly compatible across years, requiring some pre-processing corrections. Because of this, the Center of Politics and Economics of the Public Sector (CEPESP) of the Getulio Vargas Foundation (FGV) routinely collects data from TSE and processes them. These pre-processing corrections consist mostly of standardization of certain variables (e.g. description of the election) so that they are compatible across years[3].

We also collect data on policy outcomes. For that, we examine a series of variables associated with the municipality's prenatal health, infant education and fiscal outcomes. All policy outcomes are from the third year of the mandate of each election. So we gather annual data from 2007, 2011 and 2015. Fiscal data comes from the National Secretary of Treasury (*Secretaria do Tesouro Nacional - STN*). They make this information available through its dataset called FINBRA (*Finanças do Brasil - Dados Contábeis dos Municípios*). This dataset contains information from declarations received by the National Treasury, which includes annual consolidated data on municipal spending and revenues for all Brazilian municipalities. From this dataset we retrieve information on the ratio of total and current revenues and expenditures; expenditures with the legislature, community assistance, social assistance, child education, education as a whole, health and assistance to children[4]. All expenditure variables are calculated as a proportion of the current expenses.

Data on prenatal health outcomes come from the Data Processing Department of the Unified Health System (DATASUS) from the Ministry of Health. We collect data on the proportion of live births with less than 2500g, the proportion of premature live births (less than 37 weeks), the proportion of women with four or less prenatal visits, number of live births per capita, the proportion of fetal death and proportion of infant deaths. All proportions are in relation to the number of live births. The number of live births comes from the Information System on Live Births (*Sistema de Informações sobre Nascidos Vivos - SINASC*), which also makes available the information on the number of prenatal visits, duration of pregnancy and weight. Information on infant and fetal mortality comes from the Information System on Mortality (*Sistema de Informações sobre Mortalidade - SIM*).

[3]For more details on how these pre-processing corrections are made, see https://github.com/Cepesp-Fgv/tse-dados/wiki

[4]For the fiscal exercises of 2007 and 2011, the rubric related to assistance to children was simply called "Child Assistance". This was later changed and in 2015 the rubric started including adolescents too, being called "Child and Adolescent Assistance".

Data on educational outcomes are focused on the early years of education. We gather data from two sources: we use data on educational quality from the Basic Education Assessment System (*Sistema de Avaliação da Educação Básica - SAEB*), which is composed of a set of external standardized examinations in mathematics and portuguese. Results from SAEB and school pass rates obtained from the Annual School Census are used to calculate the Municipal Index of Basic Education Development (*Índice de Desenvolvimento da Educação Básica - Ideb*), which is an index of educational quality. As such, for our measures of educational quality in early years (up to 5^{th} grade) we look at the Ideb and the scores in mathematics and portuguese from the SAEB. Moreover, we also collect data on children's enrollment in daycare centers from the Annual School Census. All this information is collected and processed by the National Institute for Educational Studies and Research "Anísio Teixeira" (*Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - INEP*), a special research agency linked to the Ministry of Education.

Lastly, we also gather data on general characteristics of municipalities from the Brazilian Population Census, which is conducted by the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Etatística - IBGE*). The census occurs every 10 years. As such, unlike other databases previously mentioned, we do not have precise information for the election years we wish to examine. Therefore, for the election years of 2004 and 2008, we use data from the 2000 Census, and for 2012 we use data from the 2010 Census. We collected data on Average Monthly Nominal Salary, Gini, percentage of households in urban areas, percentage of women in the municipality, average age, average years of schooling for both men and women and we also calculated the Herfindahl–Hirschman Indexes (HHI) for race and schooling. All these variables were used as pre-treatment variables for our balance tests.

2.3 Estimation Strategy

2.3.1 Identification Strategy and estimation

The main goal of this paper is to examine the effects on public policy of electing a female to the legislative. However, electoral success is influenced by a myriad of institutional and political factors. As such, we should not discard the possibility that the gender of the candidate is related to some of these factors, thus leading to an endogeneity problem. To circumvent this issue, our approach consists of examining cases in which candidates barely won or barely lost the election. More specifically, we use intra-coalition pairs formed by a man and a woman candidate that barely lost or won the elections.

Brazil adopts an open-list proportional-representation system for its legislative elections. Voters can vote for a specific candidate from a coalition or for the coalition itself. These votes are then summed together to determine the number of seats that should be assigned for that party or coalition following the D'Hondt formula. The placement of each candidate within the list is also determined by their total votes. Candidates with the highest vote counts are ranked first and so forth.

As such, we can expect there to be some element of randomness on who is going to be the least voted candidate in the list who won a seat (last winner) and the most voted candidate that did not get a seat (first loser). Thus, we can treat the electoral victory as random and compare these two candidates. Using this approach we can structure an algorithm in the following manner. There are $j \in \{1, ..., J\}$ coalitions and each j coalition wins s_j seats (which can be zero). For each coalition, there is a set of candidates ranked from 1 to N_j. Their ranking in the list is indexed by i and is determined by their amount of votes, $v_{i,j}$. In this setting, the 'last winner' is a candidate such that $i = s_j$, and the 'first loser' is simply $i = s_j + 1$. For each coalition we rank the candidates according to their votes and select the pair $\{s_j, s_j + 1\}$. Then we discard all pairs that do not contain at least one woman and one man, leaving us with only mixed-gender pairs of candidates. Among these pairs, we select the pair that has the lowest margin of victory (i.e. $min\{v_{s_j,j} - v_{s_j+1,j}\}$) in the municipality. We repeat the procedure for each municipality, thus giving us a sample of candidates equal to two times the number of municipalities.

This is approach is similar to studies utilizing Regression Discontinuity Design (RDD) applied to close elections[5]. While in a RDD study the procedure narrows down the sample to have only municipalities near the cutoff, we also narrow down our sample to select only the pairs with the lowest vote margin for each municipality. It is also important to notice that unlike in electoral systems with plurality voting, in our case there is a lot more uncertainty about the precise ranking of candidates, at last for those who are around the threshold of being elected or not. Not only the number of seats each coalition will get is uncertain to the elections, but also each candidate's ranking within the list is also determined by the elections. As such, we should expect the victory margin between the last winner and first loser to be exogenous for this subsample of candidates.

Our approach of comparing intra-coalition pairs is similar to Boas, Hidalgo and Richardson (2014), except that in their case they do not select the coalitions with the lowest vote margin and instead follow a more traditional RDD approach by narrowing down the sample to only pairs with vote margins sufficiently close to zero.

As mentioned before, with this strategy we should expect the margin of victory to be exogenous for our subsample of candidates. However, some cases might be a source of concern. There are a few cases in which we may fail to have a clean experiment due to how Brazil's electoral system works. Popular candidates' votes also count toward the number of seats their party will get. As such, the top-ranked candidates in each party may not be comparable to the rest. This may be a problem when there are too few elected candidates in a specific party. In this case, we might observe large differences in the votes between the last winner and the first loser. Similarly, whenever there is a draw, age becomes the tiebreaker, with elderly candidates winning the seat. As such, we removed from our sample the following cases: (i) when the party or coalition of the pair with the least difference in votes was able to obtain only a single seat in the legislative house; (ii) when there were draws, where we consider there to be a draw when: (1)

[5]See Lee and Lemieux (2010) for an overview on RDD and Eggers et al. (2015) for a discussion on the validity of RDD for close elections.

two coalitions had the same difference in votes for their respective pairs; and (2) there is more than one candidate that was last winner or first loser; and (iii) when the difference between votes in a pair was zero.

Given this sample and following a potential outcomes framework, let $Y_i(1)$ and $Y_i(0)$ be, respectively, the potential outcome for a policy variable in municipality i when a female is elected at the margin and when a male is elected at the margin. We are then interested in estimating the quantity $\tau = E[Y_i(1) - Y_i(0)]$. But we only observe one of these outcomes for each municipality. More specifically, we only observe $Y_i(1)|T = 1$ and $Y_i(0)|T = 0$, where T is an indicator variable equal to 1 if the winning candidate at the margin was female. Given our identification strategy, we expect T to be exogenous because the chances of a female winning in such circumstances and such a restricted sample should not be much different from playing heads or tails. If this is indeed the case, then $E[Y_i(1)|T = 1] = E[Y_i(1)|T = 0]$ and $E[Y_i(0)|T = 1] = E[Y_i(0)|T = 0]$. This means we can effectively use the subset of municipalities where a woman did not win the close election as our counterfactual. The quantity we are interested in is thus:

$$\hat{\tau} = E[Y_i(1)|T = 1] - E[Y_i(0)|T = 0] \tag{2.4}$$

This is our local average treatment effect since are only working with a subset of the original sample. The most straightforward way to estimate this quantity is simply by running a pooled Ordinary Least Squares (OLS) regression.

A major concern with this strategy regarding the validity of our identification strategy is that we cannot test whether the chances of a woman being the last winner is indeed exogenous. The best we can do is to check if it is plausible to believe so. To do so, we perform a covariate balance test on a wide variety of variables related to both the municipality and candidates. We also check if elected and non-elected female candidates in our sample are similar in observables. Moreover, if the chances of a woman being elected or not is indeed random in our subsample, then we should expect that the chances of a woman winning follows a binomial distribution with a probability of 0.5. So we perform a test to check if the proportion of women elected in our sample indeed falls within the 95% interval of confidence of a binomial distribution, we do so by comparing the proportion of women elected in our sample to its corresponding theoretical binomial distribution. We also repeat this exercise for various vote margins to check if this can affect our results.

2.3.2 Working Sample

Even with a proper identification strategy, the restrictions put in place might not be enough to recover and interpret the parameter of interest. In the subsample of candidates who barely won or barely lost the elections, there is still a lot of heterogeneity: the size of the population, size of legislative houses and even the composition of the legislative house in terms of gender. Therefore, to make sense of the results of our experiment, we consider two cases: one in which there is only zero or one woman elected in the legislative house as a whole and the general case without this restriction. For the former, we made sure that in the cases in which only one woman won, this woman was the last winner. This restriction allows us to better understand what happens when a single woman enters a male-dominated legislature.

Table 20 shows the number of pairs and municipalities for each case in more detail. We start by counting all mixed-gender pairs and then separate them by listing the ones with the lowest vote margin per municipality, and then we count the cases with draws, margin exactly equal to zero and coalitions with only one seat - which are the cases we wish to avoid. As it can be seen there, there are not many zero-margin cases nor cases in which the coalition had only one seat. Nonetheless, we remove these cases from our sample. It is also possible to notice that we do not lose many municipalities by ignoring the cases in which there were draws. As such, our final sample consists of minimum mixed-gender pairs excluding the cases above-mentioned.

Another concern is with the number of total seats in the legislative house. The number of seats not only influences the ranking of the last winner and first loser but also may influence the elected candidate's ability to influence public policy. As such, we decided to restrict our sample further to include only municipalities that had 9 or 10 seats in their legislative house. Table 21 shows the number of pairs and municipalities for each seat. As it can be seen, 9 and 10 are the lowest number of seats we can get and also are the ones with most of the municipalities. That means we will not be losing many observations by focusing solely on them. It is important to notice also that the number of seats is defined according to the electoral legislation, which ties the number of seats to the size of each municipality. That means the municipalities with 9 or 10 seats are small municipalities[6].

It is also important to notice that our data for policy outcomes comes from various sources. As such, the set of municipalities for which we have data on some specific outcomes may not be the same for others. This is undesirable because it may affect our ability to compare results, due to the fact different sample sizes imply in different statistical power for each regression, which in turn may affect our ability to compare results. To avoid that, we defined a work sample that contains only municipalities for which we have information on all the policy outcome variables. In 22 we present each of these policy variables and show the proportion of municipalities without missing information for that particular variable. Columns (1) and (5) display the number of observations, columns (2) and (6) display the proportion of municipalities with information for that variable, columns (3) and (7) display the number of municipalities that have information for all the variables listed and columns (4) and (5) compared the number of municipalities in columns (3) and (7) with the full set of observations in (1) and (5). As it can be seen, we have close to 90% or more observations for almost all variables, except for the ratio of Fetal deaths

[6]For the elections of 2004 and 2008, the number of available seats in municipal chambers was regimented by Resolution n°21,702/2004 from the Supreme Electoral Court (STE). According to this resolution, the number of seats was based solely on the population estimates of 2003 and 2007 for each respective municipality. Municipalities with up to 47,619 inhabitants should have precisely 9 seats. Municipalities between 47,620 and 95,238 should have 10 seats. This rule was later changed through Constitutional Amendment n°58/2009, which imposed a new rule still based on population thresholds. According to this new rule, municipalities with up to 15,000 inhabitants could have at most 9 seats. Municipalities with between more than 15,000 and up to 30,000 were allowed to have at most 11 seats and so forth. As such, we only have municipalities with up to 95,238 inhabitants in 2004/2008 and only up to 30,000 inhabitants in 2012, when we restrict our sample to have only municipalities with 9 or 10 seats.

Table 20 – Subsample for different pairs of candidates.

	Full Sample									
	Mixed Gender Pairs		Min. Mixed Gender Pairs		MM without draws		MM with zero margin		MM whose colligation had 1 seat	
Year	Pairs	Mun.	Pairs	Mun.	Pairs	Mun.	Pairs	Mun.	Pairs	Mun.
2004	8765	3164	6489	3156	6004	3002	128	64	12	6
2008	8727	3156	6418	3133	5994	2997	126	63	8	4
2012	9880	3450	7026	3428	6564	3282	88	44	6	3
Total	27372	9770	19933	9717	18562	9281	342	171	26	13

	Only 0 or 1 woman elected									
	Mixed Gender Pairs		Min. Mixed Gender Pairs		MM without draws		MM with zero margin		MM whose colligation had 1 seat	
Year	Pairs	Mun.	Pairs	Mun.	Pairs	Mun.	Pairs	Mun.	Pairs	Mun.
2004	3043	1229	2514	1223	2330	1165	42	21	4	2
2008	3010	1239	2529	1232	2344	1172	40	20	4	2
2012	2755	1109	2258	1103	2120	1060	26	13	2	1
Total	8808	3577	7301	3558	6794	3397	108	54	10	5

Notes: This table presents the number of pairs our restrictions in our sample. The first panel presents the results for the full sample. The column under Mixed Gender Pairs displays the number of mixed-gender pairs in our sample. The column Min. Mixed Gender Pairs presents the number of pairs whose candidates had the minimum difference of votes in the municipality. The column MM without draws displays the previous sample but without draws. We consider it to be a draw when (1) two colligations had the same difference in votes for their respective pairs or (2) there is more than one candidate that was the last winner or first loser. The column MM with zero margin takes the sample of MM without draws and counts how many pairs had a vote margin of zero. And finally, MM whose colligation had 1 seat counts the number of cases in which the colligation of the pair selected had only 1 seat available for them. For this last one, we consider only minimum mixed-gender pairs without draws and without zero margin. The second panel presents the same information for the subset of the full sample that contains only municipalities where zero or one woman won. When there is one woman elected, she was the last winner.

Table 21 – Number of mixed pairs per number of seats.

Num. of	Full sample		Only 0 or 1 woman	
Seats	Pairs	Municipalities	Pairs	Municipailities
9	12286	6143	4552	2276
10	632	316	262	131
11	1162	581	400	200
12	118	59	32	16
13	482	241	130	65
14	62	31	14	7
15	240	120	56	28
16	32	16	6	3
17	130	65	22	11
18	16	8	2	1
19	40	20	12	6
20	8	4	4	2

Notes: This table presents the number of pairs and municipalities for each number of seats up to 20 seats in the legislative house. The sample used here only contains the minimum mixed-gender pairs and we removed draws, zero-margin cases and single-seat colligations.

and Infant deaths. On column 3-4 and 7-8, we have the number and proportion of municipalities for which we have data on all of the selected variables. This leaves us with about 50% of the full set of observations. Although this restriction is important to make sure we can compare policy effects, we still lose a significant amount of observations and one can argue these exclusions are somewhat *ad hoc*, so we also examine the case without this restriction as a robustness check.

2.4 Covariate Balance and Validity Checks

A direct implication of our identifying assumption is that bare winners and losers, whether male or female, will have the same odds of winning or losing as a coin toss. In Table 23 we include some descriptive statistics for the full sample of candidates without any restriction. In the first panel, there are the total number of candidates, the number of elected and the proportions of each gender. In the second panel, we restrict the sample by including only municipalities that have at least one woman in their legislative house. We can see there are many more male candidates and the proportion of male winners is even higher. However, in Table 24, we focus on the subset of pairs of last winners and first losers (minimum mixed pairs, MM pairs for short), the gender ratio becomes much more even, as it can be seen in Table 24. Of course, the proportion of men and women in this sample is equal to 50% by construction, but it is important to notice that the proportion of elected candidates of each gender is also approximately 50%. We can test if it is too implausible to assume the chances of winning follow a binomial distribution with a chance of success of 50%.

In Table 25 we perform the binomial distribution test to our working sample, which

Table 22 – Number of observations in the full sample and in the working sample.

	Full Sample				Working Sample			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prop. Underweight Live Births ($< 2500g$)	9108	98.1%	5106	55.0%	6312	97.7%	3252	50.3%
Prop. Premature Live Births ($<$ 37 weeks)	9072	97.7%	5106	55.0%	6286	97.3%	3252	50.3%
Prop. Women w/4 or more Prenatal Visits	9281	100.0%	5106	55.0%	6459	100.0%	3252	50.3%
N° of Live Births per Capita	9281	100.0%	5106	55.0%	6459	100.0%	3252	50.3%
Prop. Fetal Deaths	6988	75.3%	5106	55.0%	4518	69.9%	3252	50.3%
Prop. Infant Deaths	7331	79.0%	5106	55.0%	4792	74.2%	3252	50.3%
Score in Mathematics (SAEB)	8476	91.3%	5106	55.0%	5789	89.6%	3252	50.3%
Score in Portuguese (SAEB)	8476	91.3%	5106	55.0%	5789	89.6%	3252	50.3%
Average Score (SAEB)	8476	91.3%	5106	55.0%	5789	89.6%	3252	50.3%
Ideb Score	8476	91.3%	5106	55.0%	5789	89.6%	3252	50.3%
Enrollment in Childcare per Capita	9281	100.0%	5106	55.0%	6459	100.0%	3252	50.3%
Total Superavit (ratio)	9050	97.5%	5106	55.0%	6307	97.6%	3252	50.3%
Current Superavit (ratio)	9050	97.5%	5106	55.0%	6307	97.6%	3252	50.3%
Legislature Exp. (% current exp.)	8615	92.8%	5106	55.0%	6069	94.0%	3252	50.3%
Communitary Assistance Exp. (% current exp.)	8949	96.4%	5106	55.0%	6244	96.7%	3252	50.3%
Social Assitence Exp. (% current exp.)	9033	97.3%	5106	55.0%	6294	97.4%	3252	50.3%
Child Education Exp. (% current expenses)	8748	94.3%	5106	55.0%	6137	95.0%	3252	50.3%
Education Exp. (% current exp.)	9035	97.3%	5106	55.0%	6297	97.5%	3252	50.3%
Health Exp. (% current exp.)	9037	97.4%	5106	55.0%	6297	97.5%	3252	50.3%
Assistance to Children Exp. (% current exp.)	8635	93.0%	5106	55.0%	6050	93.7%	3252	50.3%

Notes: This table presents the number of observations in the full sample and in our working sample. Observations here are the pooled number of municipalities for 2004, 2008 and 2012. Column 1 presents the number of observations for each variable in the full sample. Column 2 presents the proportion of observations of each variable in relation to the full number of observations in our full sample. Column 3 presents the number of observations that are common to all the variables in the table. Column 4 presents the proportion of the observations in column 3 in relation to the full set of observations. Columns 5-8 presents similar information for a subsample that contains only observations of municipalities that (i) had between 9 and 10 seats in their respective legislative house; (ii) without zero-margin cases; and (iii) without single-seat colligations. As such, column 7 presents the number of observations common to all our policy variables. These observations form our working sample.

Table 23 – Averages for the general population of candidates.

Full Sample						
Year	n° cand.	n° elected	% males	% fem.	% males Elected	% fem. Elected
2004	334799	50839	79.89%	20.08%	87.31%	12.67%
2008	325372	51837	79.80%	20.20%	87.41%	12.59%
2012	395756	57434	71.64%	28.34%	86.44%	13.56%
Only 0 or 1 woman elected						
Year	n° cand.	n° elected	% males	% fem.	% males Elected	% fem. Elected
2004	80750	11386	80.33%	19.66%	94.40%	5.59%
2008	76767	11511	80.61%	19.39%	94.77%	5.23%
2012	80380	11225	72.32%	27.67%	95.06%	4.94%

Notes: Table with number of candidates, number of candidates elected and the share of each sex in the general population of candidates. Here we are counting all candidates for the full sample and for the subsample that had 0 or 1 woman elected.

Table 24 – Averages for the subset of min. mixed pairs of candidates.

Subset of minimum mixed pairs						
Year	n° cand.	n° elected	% males	% fem.	% males Elected	% fem. Elected
2004	4920	2460	50.00%	50.00%	47.93%	52.07%
2008	4896	2448	50.00%	50.00%	52.12%	47.88%
2012	5584	2792	50.00%	50.00%	53.01%	46.99%
Only 0 or 1 woman elected						
Year	n° cand.	n° elected	% males	% fem.	% males Elected	% fem. Elected
2004	1892	946	50.00%	50.00%	47.89%	52.11%
2008	1860	930	50.00%	50.00%	52.90%	47.10%
2012	1762	881	50.00%	50.00%	52.10%	47.90%

Notes: Table with number of candidates, number of candidates elected and the share of each sex in the subset of mixed gender pairs of candidates with the least difference in vote share in their respective municipality. The first panel presents the subset in full and the second panel is further restricts this sample to have only pairs of municipalities that had only 0 or 1 woman elected.

includes the restrictions mentioned in the last section. As in the previous case, the proportion of male and female candidates winning an election is also very close to 50%, to the point we cannot reject the hypothesis that the observed values were drawn from a binomial distribution. To show that, we displayed the 95% confidence interval of a theoretical binomial distribution using the number of candidates and elected candidates as input. As it can be seen, the observed proportions all fall within the 95% confidence interval, and this is valid for both the full subset of minimum mixed pairs as well as for the subset with only 0 or 1 woman elected. We also broke down the sample by election year and the results are still the same.

We turn next to the victory margins of each candidate to check the distribution of victory margins. We expect most margins to be small and this is indeed the case. In Figure 39 we plot the histogram for the victory margins in the working sample, and in 40 we restrict the raw vote margins to be lower than 400 votes, to remove outliers and make the graphs easier to see. As it can be seen, most municipalities have a raw victory margin of around 200 votes or less.

In Figure 41 we want to check how the binomial distribution test done in 25 behaves when we restrict the victory margin. That is, we want to check if we can reject the assumption that the probability of electing a man or a woman comes from a binomial distribution for different margins of victory. We plot the on-tailed p-value of the test against different vote margin bandwidths. We performed the test for 20 victory margin bandwidths varying from 10 to 1000 votes around the cutoff. Similar to what we had seen in the case where we do not restrict the victory margins at all, we usually cannot reject the hypothesis that the proportion of elected female candidates comes from a binomial distribution. The only cases in which we reject this null hypothesis is when we consider a victory margin very close to zero. It is important to notice that for such a narrow victory margin, the number of municipalities becomes very small, which might be the reason for the skewed proportion of female elected candidates, thus putting these cases away from the 95% confidence interval.

Now we turn to covariate balance tests for municipality characteristics that could be associated with policy outcomes. As such, we examine both general municipality characteristics as well as electoral variables. Figure 42 displays the p-values of OLS regressions for each covariate. For each regression we calculated three p-values: the 'standard' cluster-robust p-value calculated from the estimated regression model; the 'ri' p-values obtained through the randomized inference procedure using 1000 replications[7]; and the 'rwolf' p-values, which employs the Romano-Wolf correction to control for familywise error rate (FWER) using 1000 bootstrap replications[8].

For the full minimum mixed pairs sample, there are only two variables whose p-values are significant at 10%. One of which, however, is still significant using Romano-Wolf's correction,

[7]What we called randomization inference p-values here are the exact p-values for sharp null hypotheses. It follows Fisher et al. (1937)'s approach to statistical inference. This method calculates regression p-values by using the permutation of the treatment assignment rather than from the sampling strategy. See Young (2019) for an application and Athey and Imbens (2017) for more information about methods related to randomized experiments.

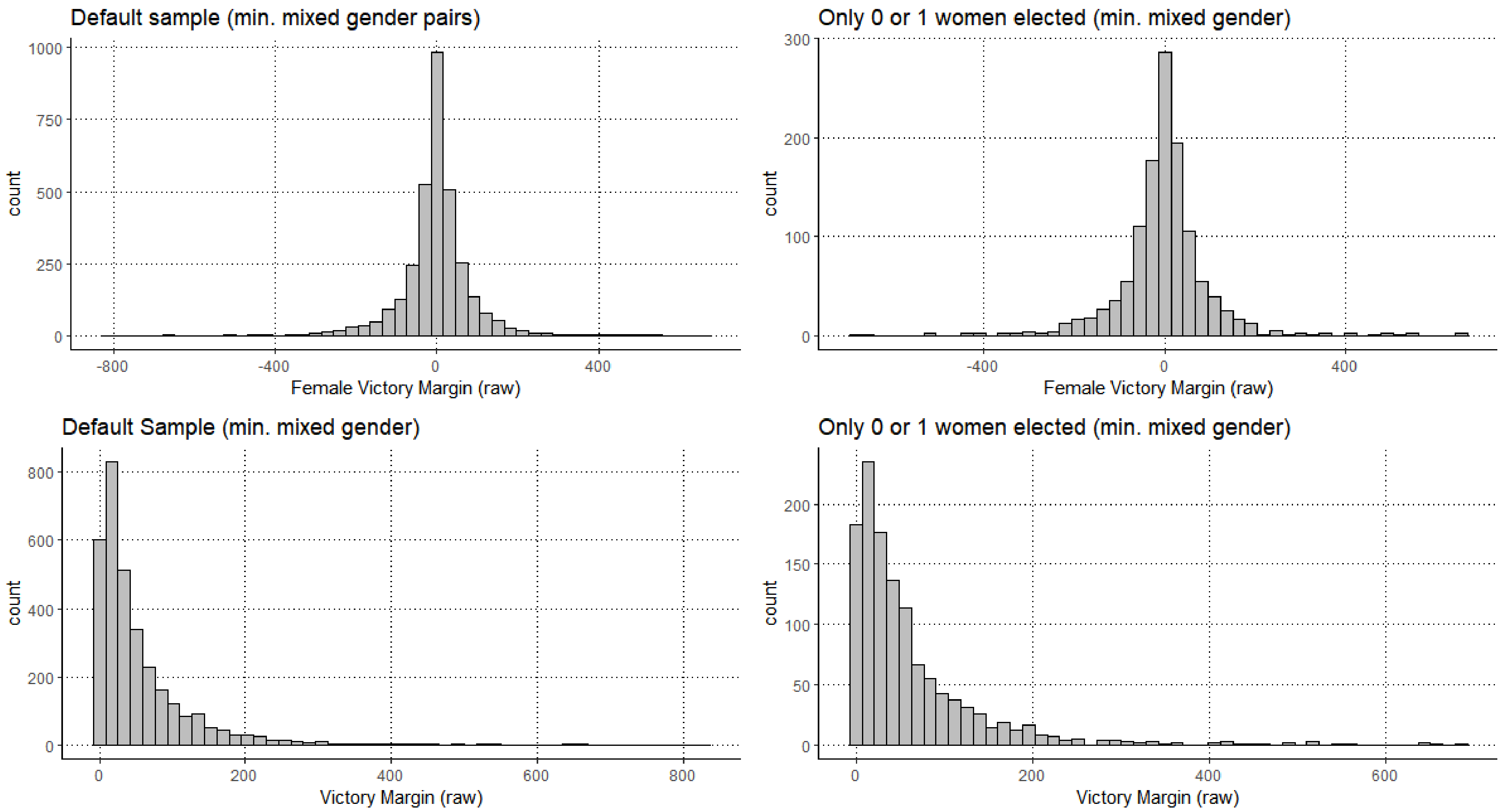
[8]See Romano and Wolf (2005a), Romano and Wolf (2005b) and Romano and Wolf (2016) for the procedure. Clarke, Romano and Wolf (2020) presents a implementation of this procedure for Stata.

Table 25 – Averages for the subset of min. mixed pairs of candidates, but this time restricting the number of seats to 9 or 10 seats.

Subset of minimum mixed pairs											
					% males	% fem.	CI 95%		one-tail	p-value	
Year	Num. cand.	Num. elected	% males	% fem.	Elected	Elected	lwr	upr	p-value	≥ 0.025	Area
2004	2782	1391	50.00%	50.00%	48.60%	51.40%	47.38%	52.62%	0.14175	TRUE	95.28%
2008	2632	1316	50.00%	50.00%	51.29%	48.71%	47.26%	52.74%	0.1815	TRUE	95.59%
2012	1090	545	50.00%	50.00%	49.36%	50.64%	45.87%	54.13%	0.36594	TRUE	95.13%
Total	6504	3252	50.00%	50.00%	50.18%	49.82%	48.28%	51.72%	0.42352	TRUE	95.25%
Only 0 or 1 woman elected											
					% males	% fem.	CI 95%		one-tail	p-value	
Year	Num cand.	Num elected	% males	% fem.	Elected	Elected	lwr	upr	p-value	≥ 0.025	Area
2004	1088	544	50.00%	50.00%	47.06%	52.94%	45.77%	54.23%	0.07852	TRUE	95.62%
2008	984	492	50.00%	50.00%	51.63%	48.37%	45.53%	54.47%	0.24946	TRUE	95.76%
2012	390	195	50.00%	50.00%	45.13%	54.87%	43.08%	56.92%	0.07595	TRUE	95.53%
Total	2462	1231	50.00%	50.00%	48.20%	51.80%	47.20%	52.80%	0.1049	TRUE	95.40%

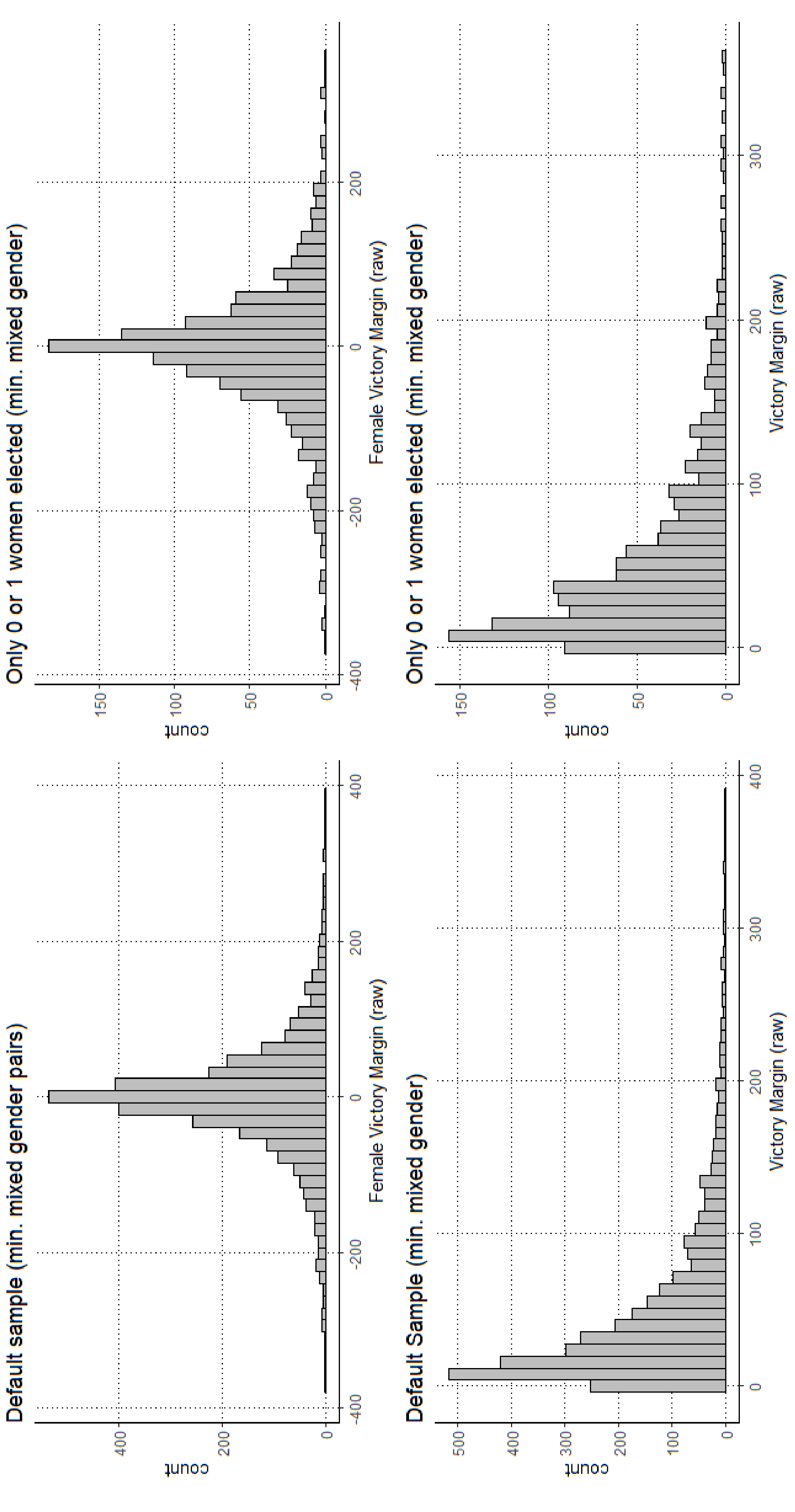
Notes: Table with number of candidates, number of candidates elected and the share of each sex in the subset of mixed gender pairs of candidates with the least difference in vote share in their respective municipality *for our working sample*, counting only the municipalities that had 9 our 10 seats available. The first panel presents the subset in full and the second panel further restricts this sample to have only pairs of municipalities that had only 0 or 1 woman elected. We also calculated the theoretical confidence level of 95% for a binomial distribution with a probability of success of 0.5, where the number of successes is equal to the number of female elected candidates observed in our data. We then compare it to the observed number of females elected. The tail on the one-tail p-value is based on whether the % of females elected is below or above 50%. The last column, Area, displays the area covered by the confidence interval. It is not exactly 95% because the number of successes is a discrete variable.

Figure 39 – Histogram of victory margins (working sample).



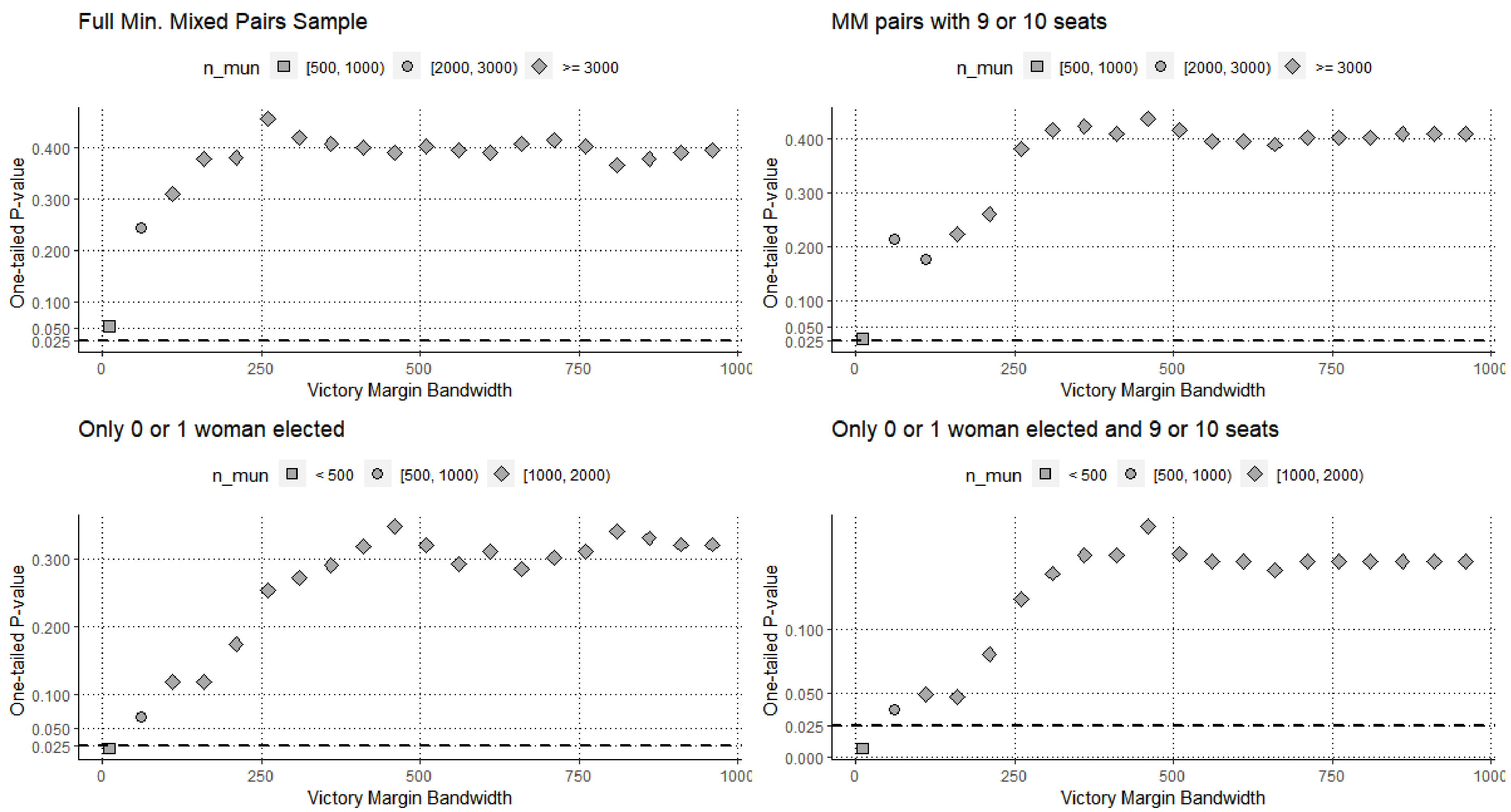
Notes: histogram of victory margins of mixed gender pairs for our sample of minimum mixed gender pairs. The upper panels show the victory margin of women and the lower panels show the victory margin of whoever won.

Figure 40 – Histogram of victory margins (working sample, < 400 raw vote margin).



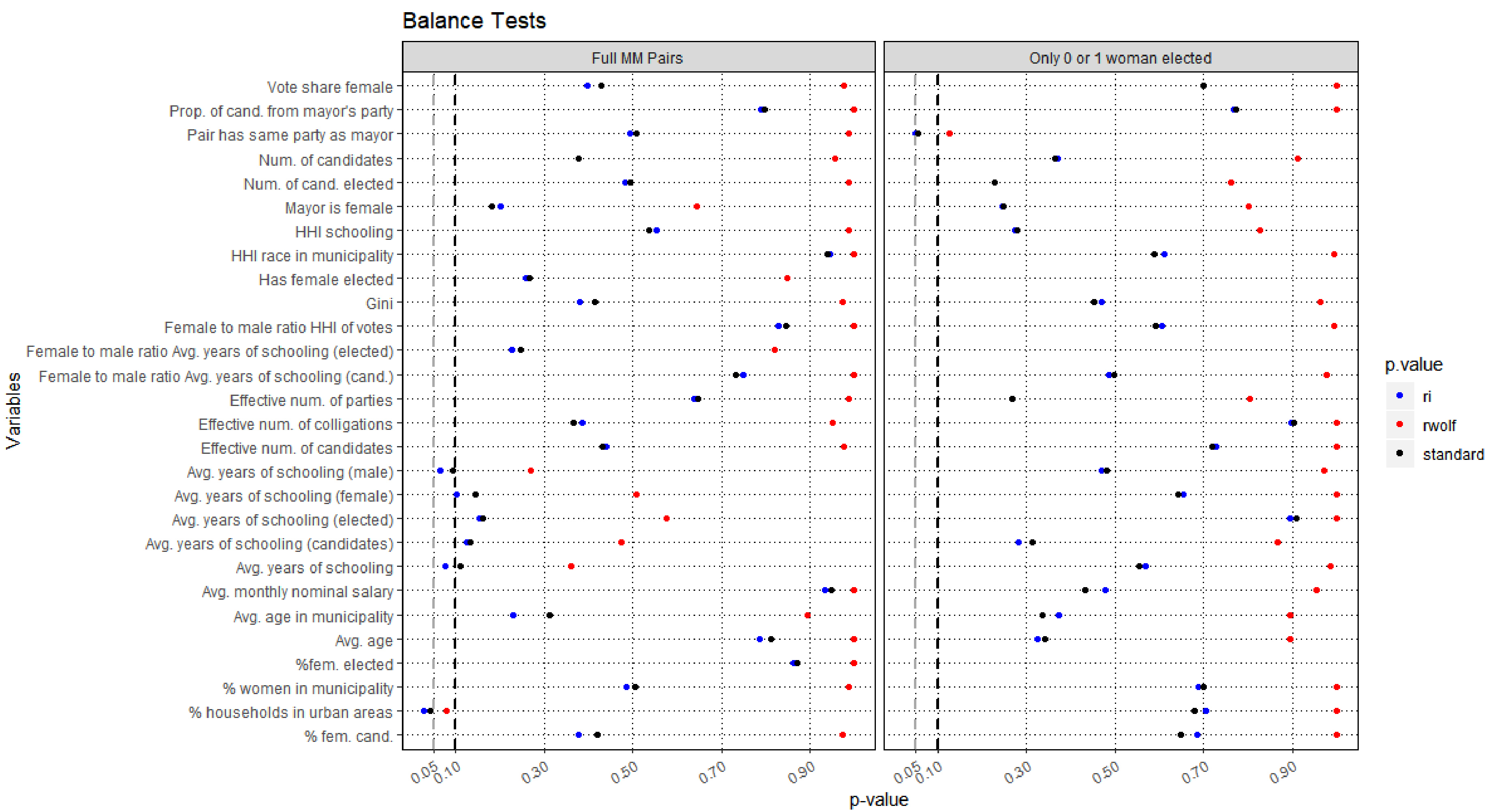
Notes: histogram of victory margins of mixed gender pairs for our sample of minimum mixed-gender pairs, but only including victory margins with less than 400 votes. The upper panels show the victory margin of women and the lower panels show the victory margin of whoever won.

Figure 41 – P-value of the binomial distribution test for both the full working sample and with zero or one woman elected.



Notes: For each panel, we perform various binomial distribution tests where we compare the observed proportion of elected female candidates with its corresponding theoretical binomial distribution with a probability of success of 0.5. The idea here is to test if we can reject the hypothesis that the proportion of elected female candidates comes from a binomial distribution with a probability of success equal to 0.5. In order to evalue that, we then calculate the corresponding one-tailed p-value, where the tail is based on whether the % of females elected is below or above 50%. We perform this test for different victory margins and plot the results in a graph. The dashed horizontal line marks the point from which we can reject the null hypothesis that the observed value comes from a binomial distribution with 95% of confidence. The victory margin is considered in terms of raw votes. We consider a total of 20 bandwidths ranging from 10 to 1000 votes for each subset analyzed. The shape of each plot point varies according to the number of municipalities considered. The first graph uses the full working sample of minimum mixed gender pairs. We then restrict this sample to only municipalities with 9 and 10 seats in panel 2. In panels 3 and 4 we repeat the same procedure but for the restricted sample that considers only municipalities that elected only zero or one woman.

Figure 42 – Balance Statistics for Municipalities in the working sample



Notes: In this figure, we display the balance statistics for municipality characteristics and candidate characteristics. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate three p-values. The standard p-value is calculated from the estimated regression model; the ri p-values are obtained through the randomized inference procedure using 1000 replications; the rwolf p-values employ the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variables are the variables in the y-axis and the explanatory variable is a dummy that is equal to 1 when a woman from the minimum mixed pair won the election. We use our working sample with 9 or 10 seats for all regressions.

meaning there is a slight imbalance even after controlling for multiple hypotheses. That said, we find no covariate imbalance using Romano-Wolf's correction for the sample with only 0 or 1 women elected, and even when we consider the standard and ri p-values, only one covariate shows imbalance (whether the pair has someone from the same party as the mayor).

We further examine the covariate balance by giving a closer inspection to the balance between candidates of each gender. If our treatment is indeed random, we should expect for elected and non-elected female candidates to be very similar, and the same for male candidates. This covariate balance check between elected and non-elected candidates by gender are displayed in Figures 43 and 44. For this exercise, we checked mainly for imbalances in terms of party representation, education, age and region. Overall most variables seem to be balanced between elected and non-elected. However, there seem to be a few imbalances. Most of those lose significance when we correct for FWER, though. In the full MM sample, it seems there is an imbalance in the proportion of municipalities from south and north as well an imbalance in terms of age and education. These imbalances are not present for the restricted sample with only 0 or 1 woman elected, which only displays a lack of balance for one party after we correct for FWER (Party number 45, the Brazilian Social Democracy Party - PSDB).

Given these imbalances, we perform a F-test to test for joint orthogonality[9] and also compare the size of differences of each variable between the treated and controls. Tables 26, 27 and 28 presents our results. Following Imbens and Rubin (2015) we calculated the normalized difference between treated and controls for each sample, the results are displayed in the column "Nd."[10]. Although the % households in urban areas was found to not be balanced even when we accounted for FWER in the full MM pairs sample, it seems the size of this difference is quite small, with only 0.072 standard deviation difference in means. Moreover, the F-test was not significant. As for the sample with 0 or 1 woman elected, we reject the null of F-test, meaning there are imbalances in this sample, with the dummy variable of the candidate being in the same party as the mayor being the culprit[11]. However, just like in the previous case, this difference does not seem to be sizable, with a standard deviation difference of 0.109.

Next, we turn to the balance between elected and non-elected female candidates. As it can be seen in Table 43, for both samples we reject the null of the F-test that all coefficients are jointly zero. This is in line with the exercise using p-values, which showed there to be a few imbalances. The same is also true for the sample with only elected and non-elected male candidates, again we reject the null of the F-test for both subsamples. However, in both cases the magnitude of the differences are small. The highest difference in Table 27 is of 0.116, while the highest difference in Table 28 is of 0.157.

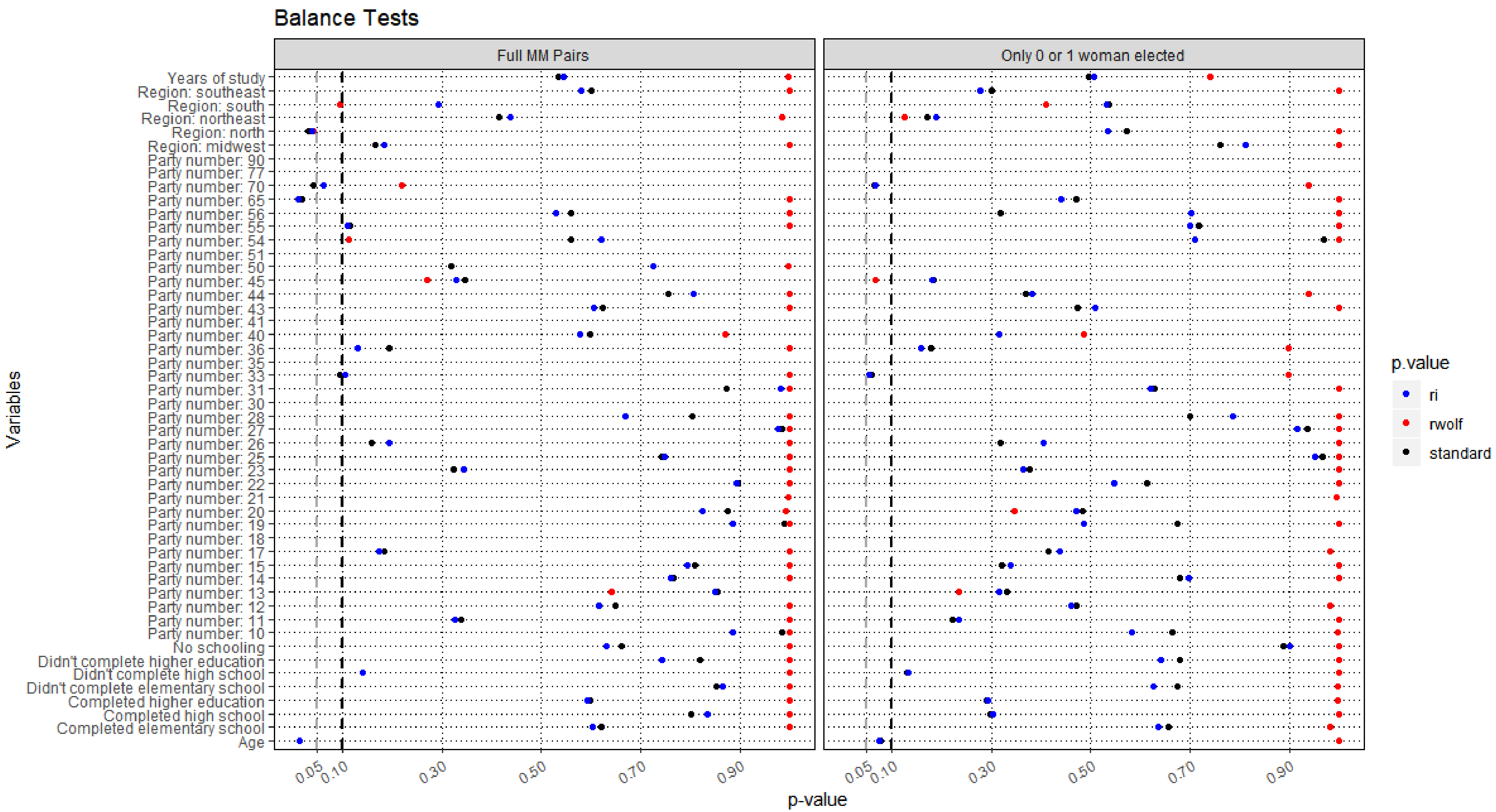
Given all the exercises performed, we believe the balance is reasonably similar to what we

[9] McKenzie (2015) suggests this simple procedure as a complement to usual covariance balance tests.

[10] Normalized difference here is calculated as $(\mu_t - \mu_c)/((\sigma_t^2 + \sigma_c^2)/2)^{1/2})$, where μ_c and σ_c are the mean and standard deviation of the control and μ_t and σ_t are the mean and standard deviation of the treated.

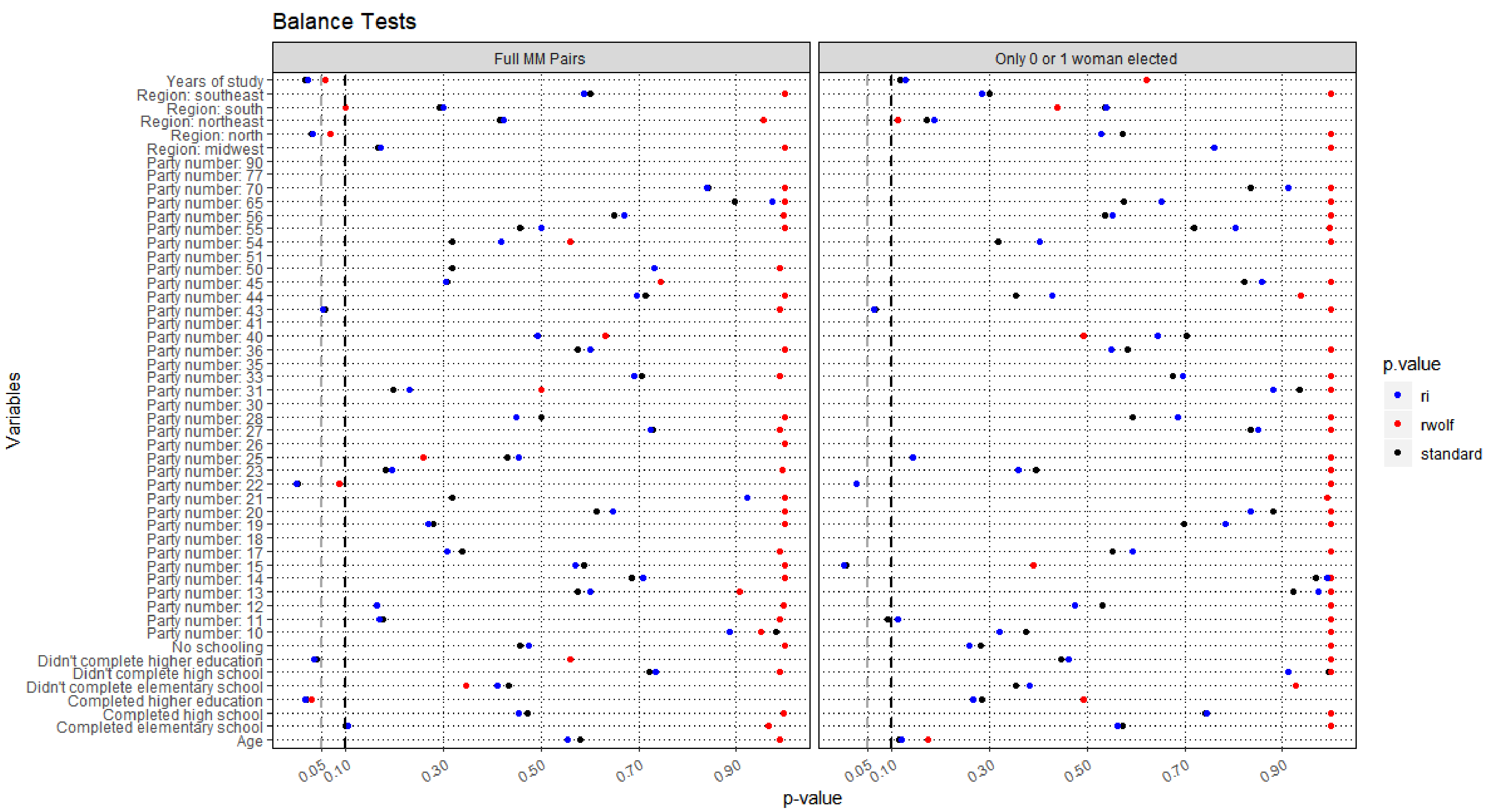
[11] If we remove this variable and perform the F-test again, the p-value becomes 0.1674 and we no longer reject the null hypothesis.

Figure 43 – Balance Statistics for Female Candidates (elected/non-elected)



Notes: In this figure, we display the balance statistics for elected and non-elected female candidates. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate three p-values. The standard p-value is calculated from the estimated regression model; the ri p-values are obtained through the randomized inference procedure using 1000 replications; the rwolf p-values employ the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variables are the variables in the y-axis and the explanatory variable is a dummy that is equal to 1 when a candidate from the minimum mixed pair won the election. We use our working sample with 9 or 10 seats for all regressions and restrict the sample to have only female candidates.

Figure 44 – Balance Statistics for Male Candidates (elected/non-elected)



Notes: In this figure, we display the balance statistics for elected and non-elected male candidates. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate three p-values. The standard p-value is calculated from the estimated regression model; the ri p-values are obtained through the randomized inference procedure using 1000 replications; the rwolf p-values employ the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variables are the variables in the y-axis and the explanatory variable is a dummy that is equal to 1 when a candidate from the minimum mixed pair won the election. We use our working sample with 9 or 10 seats for all regressions and restrict the sample to have only male candidates.

Table 26 – Balance between Treated and Controls for the working sample

	Full MM Pairs					0 or 1 woman elected				
	Treated (N = 1632)		Controls (N = 1620)			Treated (N = 633)		Controls (N = 598)		
Variables	Mean	Sd.	Mean	Sd.	Nd.	Mean	Sd.	Mean	Sd.	Nd.
Avg. years of schooling	4.930	1.165	4.995	1.123	-0.057	5.077	1.194	5.115	1.090	-0.033
Avg. years of schooling (female)	5.179	1.135	5.237	1.091	-0.052	5.314	1.161	5.343	1.064	-0.026
Avg. years of schooling (male)	4.680	1.215	4.752	1.178	-0.060	4.840	1.248	4.888	1.138	-0.040
Avg. age	42.910	2.398	42.930	2.333	-0.008	43.002	2.413	43.130	2.328	-0.054
Avg. years of schooling (candidates)	8.768	1.481	8.843	1.409	-0.052	8.721	1.510	8.805	1.424	-0.057
Elected at east one female	0.612	0.487	0.631	0.483	-0.039					
Female to male ratio Avg. years of schooling (cand.)	1.220	0.260	1.217	0.250	0.012	1.196	0.285	1.186	0.249	0.039
Avg. years of schooling (elected)	9.586	1.975	9.681	1.944	-0.048	9.359	2.005	9.372	1.939	-0.006
Female to male ratio Avg. years of schooling (elected)	1.315	0.532	1.288	0.507	0.052					
Gini	0.558	0.066	0.556	0.066	0.028	0.554	0.066	0.557	0.065	-0.042
HHI schooling	0.192	0.044	0.193	0.043	-0.022	0.193	0.044	0.196	0.043	-0.062
HHI race in municipality	0.537	0.119	0.537	0.119	0.003	0.542	0.126	0.546	0.127	-0.031
Female to male ratio HHI of votes	0.169	0.167	0.170	0.172	-0.007	0.055	0.038	0.053	0.039	0.031
Avg. age in municipality	28.049	3.142	28.160	3.103	-0.036	28.244	3.227	28.416	3.166	-0.054
Num. of candidates	60.334	32.201	59.341	30.961	0.031	64.806	35.313	63.072	31.678	0.052
Num. of cand. elected	8.094	0.301	8.087	0.299	0.024	8.107	0.320	8.085	0.313	0.070
Effective num. of candidates	30.846	13.354	30.473	13.421	0.028	32.417	14.425	32.132	13.733	0.020
Effective num. of colligations	3.955	1.597	3.904	1.590	0.032	4.189	1.675	4.177	1.577	0.007
Effective num. of parties	6.781	2.535	6.821	2.408	-0.016	6.941	2.507	7.093	2.404	-0.062
% households in urban areas	0.628	0.220	0.644	0.214	-0.072	0.644	0.229	0.649	0.216	-0.024
Mayor is female	0.111	0.314	0.097	0.295	0.047	0.098	0.298	0.079	0.270	0.067
% fem. cand.	0.219	0.068	0.221	0.069	-0.028	0.204	0.071	0.202	0.071	0.026
%fem. elected	0.116	0.118	0.117	0.116	-0.006					
% women in municipality	0.495	0.014	0.495	0.014	-0.024	0.495	0.013	0.495	0.014	0.022
Prop. of cand. from mayor's party	0.168	0.095	0.167	0.094	0.009	0.160	0.087	0.159	0.086	0.017
Avg. monthly nominal salary	445.086	290.449	444.444	263.995	0.002	468.478	297.053	456.230	252.673	0.044
Pair has same party as mayor	0.247	0.431	0.237	0.425	0.023	0.278	0.448	0.231	0.422	0.109
Vote share female	0.151	0.071	0.153	0.072	-0.028	0.099	0.044	0.098	0.046	0.022
F-Test for Joint Orthogonality of All Variables (p-value)					0.422					0.077

Notes: This table displays the means and standard deviations (Sd.) for balance covariates for our working sample. The first panel presents the full minimum mixed pairs sample and the second panel presents the selected statistics only for the municipalities that elected 0 or 1 woman. Our treatment is whether the female candidate won or not. We also calculated the normalized difference (Nd.) between treated and controls. This difference is defined as $(\mu_t - \mu_c)/((\sigma_t^2 + \sigma_c^2)/2)^{1/2})$, where μ_c and σ_c are the mean and standard deviation of the control and μ_t and σ_t are the mean and standard deviation of the treated. In the last row we also perform an F-test based on an OLS regression using the variables listed. In the Full MM Pairs sample we didn't include the variable "Elected at least one female" in order to avoid Perfect multicollinearity.

Table 27 – Balance between Treated and Controls for the working sample (female candidates only)

	Full MM Pairs					0 or 1 woman elected				
	Treated (N = 1632)		Controls (N = 1620)			Treated (N = 633)		Controls (N = 598)		
Variables	Mean	Sd.	Mean	Sd.	Nd.	Mean	Sd.	Mean	Sd.	Nd.
No schooling	0.017	0.128	0.019	0.135	-0.015	0.017	0.131	0.018	0.134	-0.008
Years of study	11.497	3.598	11.418	3.617	0.022	11.396	3.631	11.535	3.644	-0.038
Completed high school	0.344	0.475	0.340	0.474	0.009	0.348	0.477	0.319	0.467	0.060
Didn't complete high school	0.034	0.181	0.044	0.205	-0.052	0.033	0.179	0.050	0.218	-0.085
Completed elementary school	0.088	0.284	0.093	0.291	-0.017	0.101	0.302	0.094	0.292	0.025
Didn't complete elementary school	0.086	0.280	0.084	0.277	0.007	0.087	0.282	0.080	0.272	0.024
Age	43.545	9.592	44.378	9.809	-0.086	43.926	9.840	44.875	9.307	-0.099
Party number: 10	0.010	0.099	0.010	0.099	-0.001	0.013	0.112	0.010	0.100	0.025
Party number: 11	0.085	0.279	0.076	0.265	0.034	0.093	0.291	0.074	0.261	0.071
Party number: 12	0.063	0.243	0.059	0.236	0.016	0.060	0.238	0.070	0.256	-0.041
Party number: 13	0.091	0.288	0.093	0.291	-0.007	0.092	0.289	0.109	0.312	-0.057
Party number: 14	0.066	0.249	0.064	0.244	0.011	0.070	0.255	0.064	0.244	0.024
Party number: 15	0.157	0.364	0.154	0.361	0.009	0.163	0.369	0.142	0.349	0.057
Party number: 17	0.017	0.130	0.012	0.108	0.045	0.013	0.112	0.018	0.134	-0.047
Party number: 19	0.005	0.070	0.005	0.070	-0.001	0.006	0.079	0.008	0.091	-0.024
Party number: 20	0.018	0.134	0.019	0.137	-0.006	0.021	0.142	0.015	0.122	0.041
Party number: 22	0.066	0.248	0.067	0.250	-0.004	0.055	0.229	0.062	0.241	-0.028
Party number: 23	0.031	0.172	0.037	0.189	-0.035	0.024	0.152	0.032	0.176	-0.049
Party number: 25	0.108	0.310	0.104	0.306	0.011	0.098	0.297	0.099	0.298	-0.002
Party number: 26	0.000	0.000	0.001	0.035	-0.050	0.000	0.000	0.002	0.041	-0.058
Party number: 27	0.010	0.099	0.010	0.099	-0.001	0.006	0.079	0.007	0.082	-0.005
Party number: 28	0.006	0.078	0.007	0.082	-0.008	0.005	0.069	0.003	0.058	0.022
Party number: 31	0.010	0.099	0.009	0.096	0.006	0.011	0.105	0.008	0.091	0.027
Party number: 33	0.010	0.099	0.017	0.128	-0.060	0.005	0.069	0.017	0.128	-0.116
Party number: 36	0.005	0.070	0.009	0.093	-0.046	0.005	0.069	0.012	0.108	-0.077
Party number: 40	0.051	0.221	0.056	0.229	-0.018	0.049	0.216	0.062	0.241	-0.056
Party number: 43	0.022	0.147	0.025	0.155	-0.017	0.021	0.142	0.027	0.162	-0.041
Party number: 44	0.012	0.107	0.010	0.102	0.011	0.016	0.125	0.010	0.100	0.051
Party number: 45	0.126	0.332	0.115	0.320	0.033	0.147	0.354	0.120	0.326	0.078
Party number: 50	0.000	0.000	0.001	0.025	-0.035					
Party number: 54	0.001	0.025	0.001	0.035	-0.020	0.002	0.040	0.002	0.041	-0.002
Party number: 55	0.015	0.123	0.023	0.149	-0.055	0.013	0.112	0.015	0.122	-0.021
Party number: 56	0.001	0.025	0.001	0.035	-0.020	0.000	0.000	0.002	0.041	-0.058
Party number: 65	0.004	0.065	0.012	0.108	-0.084	0.006	0.079	0.010	0.100	-0.041
Party number: 70	0.010	0.102	0.004	0.066	0.071	0.009	0.097	0.002	0.041	0.105
Region: midwest	0.077	0.266	0.091	0.287	-0.051	0.081	0.272	0.085	0.280	-0.017
Region: northeast	0.415	0.493	0.401	0.490	0.029	0.365	0.482	0.328	0.470	0.078
Region: north	0.099	0.299	0.078	0.268	0.076	0.084	0.277	0.075	0.264	0.031
Region: southeast	0.257	0.437	0.265	0.441	-0.018	0.294	0.456	0.321	0.467	-0.059
Region: south	0.152	0.359	0.165	0.372	-0.037	0.177	0.382	0.191	0.393	-0.035
Completed higher education	0.365	0.482	0.356	0.479	0.019	0.360	0.480	0.390	0.488	-0.061
Didn't complete higher education	0.067	0.250	0.065	0.246	0.008	0.054	0.226	0.048	0.215	0.024
F-Test for Joint Orthogonality of All Variables (p-value)					0.000					0.000

Notes: This table displays the means and standard deviations (Sd.) for balance covariates for our working sample. The first panel presents the full minimum mixed pairs sample and the second panel presents the selected statistics only for the municipalities that elected 0 or 1 woman. Our treatment is whether the female candidate won or not. We also calculated the normalized difference (Nd.) between treated and controls. This difference is defined as $(\mu_t - \mu_c)/((\sigma_t^2 + \sigma_c^2)/2)^{1/2})$, where μ_c and σ_c are the mean and standard deviation of the control and μ_t and σ_t are the mean and standard deviation of the treated. In the last row we also perform an F-test based on an OLS regression using the variables listed. In the Full MM Pairs sample we didn't include the party number 70 in order to avoid perfect multicollinearity. Likewise, we removed party number 56 from the sample with 0 or 1 woman elected.

Table 28 – Balance between Treated and Controls for the working sample (male candidates only)

	Full MM Pairs					0 or 1 woman elected				
	Treated (N = 1632)		Controls (N = 1620)			Treated (N = 633)		Controls (N = 598)		
Variables	Mean	Sd.	Mean	Sd.	Nd.	Mean	Sd.	Mean	Sd.	Nd.
No schooling	0.054	0.226	0.048	0.214	0.026	0.057	0.232	0.043	0.204	0.061
Years of study	8.990	4.177	9.340	4.225	-0.083	9.171	4.230	9.544	4.119	-0.089
Completed high school	0.327	0.469	0.315	0.465	0.025	0.335	0.472	0.326	0.469	0.019
Didn't complete high school	0.055	0.228	0.058	0.234	-0.012	0.058	0.235	0.059	0.235	0.000
Completed elementary school	0.149	0.356	0.129	0.335	0.057	0.133	0.340	0.144	0.351	-0.032
Didn't complete elementary school	0.224	0.417	0.212	0.409	0.027	0.209	0.407	0.187	0.390	0.053
Age	43.815	10.745	43.609	10.269	0.020	44.336	10.397	43.413	10.150	0.090
Party number: 10	0.013	0.115	0.014	0.116	-0.001	0.009	0.097	0.015	0.122	-0.051
Party number: 11	0.090	0.286	0.077	0.267	0.047	0.103	0.304	0.075	0.264	0.096
Party number: 12	0.069	0.253	0.057	0.232	0.049	0.060	0.238	0.052	0.222	0.036
Party number: 13	0.080	0.271	0.085	0.279	-0.020	0.084	0.277	0.085	0.280	-0.006
Party number: 14	0.064	0.245	0.068	0.252	-0.014	0.076	0.265	0.075	0.264	0.002
Party number: 15	0.156	0.363	0.149	0.357	0.019	0.180	0.385	0.124	0.330	0.157
Party number: 17	0.014	0.118	0.010	0.102	0.033	0.013	0.112	0.017	0.128	-0.034
Party number: 19	0.008	0.089	0.005	0.070	0.038	0.005	0.069	0.003	0.058	0.022
Party number: 20	0.018	0.134	0.016	0.126	0.018	0.013	0.112	0.012	0.108	0.008
Party number: 21	0.001	0.025	0.000	0.000	0.035					
Party number: 22	0.054	0.226	0.082	0.275	-0.112	0.054	0.226	0.085	0.280	-0.124
Party number: 23	0.034	0.182	0.043	0.203	-0.046	0.030	0.171	0.038	0.192	-0.046
Party number: 25	0.112	0.316	0.104	0.305	0.027	0.096	0.295	0.122	0.328	-0.082
Party number: 27	0.010	0.099	0.009	0.093	0.012	0.009	0.097	0.008	0.091	0.012
Party number: 28	0.007	0.085	0.006	0.074	0.022	0.003	0.056	0.002	0.041	0.030
Party number: 31	0.011	0.104	0.007	0.082	0.045	0.006	0.079	0.007	0.082	-0.005
Party number: 33	0.010	0.099	0.011	0.105	-0.013	0.006	0.079	0.008	0.091	-0.024
Party number: 36	0.009	0.095	0.007	0.086	0.020	0.009	0.097	0.007	0.082	0.031
Party number: 40	0.050	0.217	0.055	0.228	-0.024	0.055	0.229	0.060	0.238	-0.021
Party number: 43	0.021	0.145	0.032	0.176	-0.066	0.016	0.125	0.032	0.176	-0.105
Party number: 44	0.010	0.099	0.011	0.105	-0.013	0.009	0.097	0.005	0.071	0.053
Party number: 45	0.122	0.327	0.110	0.314	0.036	0.126	0.333	0.122	0.328	0.013
Party number: 50	0.000	0.000	0.001	0.025	-0.035					
Party number: 54	0.000	0.000	0.001	0.025	-0.035	0.000	0.000	0.002	0.041	-0.058
Party number: 55	0.013	0.115	0.017	0.128	-0.026	0.013	0.112	0.015	0.122	-0.021
Party number: 56	0.001	0.035	0.002	0.043	-0.016	0.002	0.040	0.003	0.058	-0.036
Party number: 65	0.013	0.113	0.012	0.110	0.005	0.013	0.112	0.017	0.128	-0.034
Party number: 70	0.009	0.092	0.009	0.096	-0.007	0.009	0.097	0.008	0.091	0.012
Region: midwest	0.077	0.266	0.091	0.287	-0.051	0.081	0.272	0.085	0.280	-0.017
Region: northeast	0.415	0.493	0.401	0.490	0.029	0.365	0.482	0.328	0.470	0.078
Region: north	0.099	0.299	0.078	0.268	0.076	0.084	0.277	0.075	0.264	0.031
Region: southeast	0.257	0.437	0.265	0.441	-0.018	0.294	0.456	0.321	0.467	-0.059
Region: south	0.152	0.359	0.165	0.372	-0.037	0.177	0.382	0.191	0.393	-0.035
Completed higher education	0.147	0.354	0.177	0.382	-0.082	0.163	0.369	0.186	0.389	-0.060
Didn't complete higher education	0.045	0.207	0.060	0.238	-0.071	0.046	0.209	0.055	0.229	-0.043
F-Test for Joint Orthogonality of All Variables (p-value)					0.000					0.000

Notes: This table displays the means and standard deviations (Sd.) for balance covariates for our working sample. The first panel presents the full minimum mixed pairs sample and the second panel presents the selected statistics only for the municipalities that elected 0 or 1 woman. Our treatment is whether the female candidate won or not. We also calculated the normalized difference (Nd.) between treated and controls. This difference is defined as $(\mu_t - \mu_c)/((\sigma_t^2 + \sigma_c^2)/2)^{1/2})$, where μ_c and σ_c are the mean and standard deviation of the control and μ_t and σ_t are the mean and standard deviation of the treated. In the last row we also perform an F-test based on an OLS regression using the variables listed. In the Full MM Pairs sample we didn't include the party number 56 in order to avoid perfect multicollinearity.

should expect if the treatment was randomly assigned for municipality characteristics. Although there are a few imbalances between treated and controls in a small number of variables, the magnitude of the difference is reasonably small. As such, we believe it might not impact our ability to correctly identify the parameter of interest in the outcome regressions. However, for the comparison between candidates of each gender, the lack of balance seems to be more prevalent. As such, for our policy results, we include specifications with year and region fixed effects.

2.5 Main Results

This section examines whether electing a woman to the legislative has any effect in terms of policy outcomes. All policy outcomes are from 3 years after the respective election. That is, for 2004, 2008 and 2012, we are looking for effects in policy variables in, respectively, 2007, 2011 and 2015. We focus our analysis on policy impacts in three areas: health, education and fiscal outcomes. For our purposes, we are mainly interested in health and educational outcomes associated with women's preferences according to the economic literature. This leads us to examine mainly health outcomes associated with prenatal care and child education. As for fiscal outcomes, they are there mainly for us to examine whether the policy outcomes are associated with more or less spending. The results can be found in Tables 29, 30 and 31 respectively. Columns (1) and (3) present our OLS regression results for each sample of MM pairs. We also include year and region effects to these regressions to this set of regressions in columns (2) and (4). For context, we also display the means for each variable in municipalities where a man won the elections, which are our control group.

Out of all these groups, we have only found consistent effects in health outcomes. We found evidence that electing a woman to the legislative decreases the proportion of underweight live births (infants born with less than 2500g) by around 0.3%, but only for municipalities where the woman elected was the sole woman in the legislative house. This is the only effect that is still consistent when we include year and region fixed effects. Although not significant, it is also interesting to notice the point-estimates of the proportion of premature live births are also negative (except for column (4)), suggesting improvement in prenatal health care services. However, it is important to notice the proportion of women with 4 or more prenatal displays a negative point-estimate, which might be counterintuitive given our previous result. This result, however, is only significant for the sample of full MM pairs and the effect is no longer significant when we include year and region fixed effects. As for the other health outcome, the number of live birth displays positive point-estimates and is even significant for the subsample with only 0 or 1 woman elected, but it loses significance once we include fixed effects. Lastly, there does not seem to be any effects on the proportion of both fetal and infant deaths. The point-estimates are small and signs are inconsistent.

In regards to the educational outcomes, we examine the results from the National Basic Education Assessment System (SAEB) for the early years of elementary school (up to 5^{th} grade), the Index of Basic Education Development (IDEB) and the enrollment in childcare. There seems to not be any relevant effects on educational outcomes, point-estimates are even negative, but they are not very consistent. Not only they are not significant, but the magnitude of estimates are also quite different when we include fixed effects.

That said, when we turn to fiscal variables, there seems evidence to suggest municipalities where the female candidate won spend more on education. Education expenditure displays a positive effect in all regressions, but they lose significance when we include year and region fixed effects. It is worth noting that these effects are quite small and reasonably consistent across specifications and sample. It draws attention that this effect is absent when we look only

Table 29 – Effects of Electing a Councilwoman on Health Outcomes

	Full MM Pairs			Only 0 or 1 woman elected		
Variable	Mean	(1)	(2)	Mean	(3)	(4)
Prop. Underweight Live Births (< 2500g, %)	7.7495	-0.1453	-0.0958	7.9385	-0.3233**	-0.2546*
	(0.0681)	(0.0921)	(0.087)	(0.1172)	(0.1558)	(0.1485)
Prop. Premature Live Births (< 37 weeks, %)	8.1737	-0.0923	-0.0149	8.1075	-0.019	0.0493
	(0.1046)	(0.1563)	(0.1365)	(0.1668)	(0.2545)	(0.2245)
Prop. Women w/4 or more Prenatal Visits (%)	89.7817	-0.5935*	-0.1946	90.001	-0.5483	-0.1855
	(0.2558)	(0.3468)	(0.302)	(0.4041)	(0.535)	(0.475)
N° of Live Births per Capita	1.5178	0.0217	0.0055	1.4819	0.0433*	0.027
	(0.01)	(0.0142)	(0.0122)	(0.0154)	(0.0223)	(0.0194)
Prop. Fetal Deaths (%)	1.4429	0.0025	-0.001	1.4355	-0.0191	-0.0286
	(0.024)	(0.0327)	(0.0324)	(0.0434)	(0.0559)	(0.0557)
Prop. Infant Deaths (%)	1.7475	0.0455	0.0327	1.74	0.0674	0.0507
	(0.0279)	(0.0386)	(0.0387)	(0.0452)	(0.0646)	(0.0647)
Observations:		3245	3245		1224	1224
Year and Region Fixed Effects:		No	Yes		No	Yes

Notes: In this table, we report our estimates of the OLS regression for the effect of the elected a female candidate to the legislative house on policy variables associated with health outcomes. Our units of observation are municipalities. We use our working sample restricted to have only municipalities with 9 or 10 seats in their respective legislative houses. On the column Full MM Pairs, we report the results for all minimum mixed-gender pairs on our sample. On the column Only 0 or 1 woman elected we report our effects for only the municipalities that had elected 0 or 1 woman, where the woman elected is the one from the minimum mixed-gender pair. Columns (1) and (3) report the OLS regression for each sample. We also display the same set of regressions but with year and region fixed effects in columns (2) and (4). Columns 'mean' in each sample is the mean for each variable in municipalities where a man won the election. Standard errors are in parenthesis. $*p < 0.1, **p < 0.05, *** < 0.01$.

Table 30 – Effects of Electing a Councilwoman on Educational Outcomes

	Full MM Pairs			Only 0 or 1 woman elected		
Variable	Mean	(1)	(2)	Mean	(3)	(4)
Score in Mathematics (SAEB)	197.548	-1.3856	-0.1806	200.01	-1.94	-0.2838
	(0.6593)	(0.8722)	(0.5542)	(1.0487)	(1.4133)	(0.8897)
Score in Portuguese (SAEB)	179.507	-1.1747	-0.1338	181.579	-1.9332	-0.695
	(0.5946)	(0.7949)	(0.4695)	(0.9377)	(1.2934)	(0.767)
Average Score (SAEB)	4.9978	-0.0478	-0.0059	5.0825	-0.0722	-0.0181
	(0.0232)	(0.0308)	(0.0187)	(0.0367)	(0.0499)	(0.0302)
Ideb Score	4.4131	-0.0639	-0.0046	4.5122	-0.0787	-0.0045
	(0.0296)	(0.0389)	(0.0218)	(0.0469)	(0.0629)	(0.0354)
Enrollment in Childcare per Capita	1.096	0.0122	0.0291	1.0832	0.0424	0.0489
	(0.0223)	(0.0292)	(0.0273)	(0.0338)	(0.045)	(0.0425)
Observations:		3245	3245		1224	1224
Year and Region Fixed Effects:		No	Yes		No	Yes

Notes: In this table, we report our estimates of the OLS regression for the effect of the elected a female candidate to the legislative house on policy variables associated with educational outcomes. Our units of observation are municipalities. We use our working sample restricted to have only municipalities with 9 or 10 seats in their respective legislative houses. On the column Full MM Pairs, we report the results for all minimum mixed-gender pairs on our sample. On the column Only 0 or 1 woman elected we report our effects for only the municipalities that had elected 0 or 1 woman, where the woman elected is the one from the minimum mixed-gender pair. Columns (1) and (3) report the OLS regression for each sample. We also display the same set of regressions but with year and region fixed effects in columns (2) and (4). Columns 'mean' in each sample is the mean for each variable in municipalities where a man won the election. Standard errors are in parenthesis. $*p < 0.1, **p < 0.05, *** < 0.01$.

Table 31 – Effects of Electing a Councilwoman on Fiscal Outcomes

	Full MM Pairs			Only 0 or 1 woman elected		
Variable	Mean	(1)	(2)	Mean	(3)	(4)
Total Superavit (ratio, %)	105.885	0.1645	0.3166	106.052	0.1363	0.2122
	(0.2472)	(0.3724)	(0.3274)	(0.4052)	(0.5829)	(0.5084)
Current Superavit (ratio, %)	122.448	0.2792	0.5104	123.161	-0.1621	0.3136
	(0.3119)	(0.4539)	(0.4332)	(0.506)	(0.7567)	(0.7318)
Legislature Exp. (% current exp.)	0.0342	0.0001	0	0.0337	-0.0002	-0.0007
	(8e-04)	(9e-04)	(9e-04)	(7e-04)	(0.001)	(0.001)
Communitary Assistance Exp. (% current exp.)	0.0251	-0.0001	0.0001	0.0252	-0.0002	-0.0002
	(5e-04)	(7e-04)	(7e-04)	(9e-04)	(0.0012)	(0.0012)
Social Assitence Exp. (% current exp.)	0.0447	-0.0005	-0.0004	0.0439	0.0007	0.0005
	(6e-04)	(8e-04)	(8e-04)	(9e-04)	(0.0013)	(0.0013)
Child Education Exp. (% current expenses)	0.0326	-0.0006	0.0004	0.0336	-0.0002	0.0012
	(0.001)	(0.0013)	(0.0012)	(0.0015)	(0.002)	(0.0018)
Education Exp. (% current exp.)	0.3646	0.0071**	0.0035	0.3522	0.0113**	0.0075
	(0.0027)	(0.0035)	(0.0029)	(0.0041)	(0.0054)	(0.0046)
Health Exp. (% current exp.)	0.2498	0.0009	0.0018	0.2497	0.0025	0.0036
	(0.0016)	(0.0021)	(0.0021)	(0.0026)	(0.0035)	(0.0034)
Assistence to Children* Exp. (% current exp.)	0.0084	-0.0001	-0.0001	0.0082	0.0007	0.0008
	(2e-04)	(3e-04)	(3e-04)	(4e-04)	(6e-04)	(6e-04)
Observations:		3245	3245		1224	1224
Year and Region Fixed Effects:		No	Yes		No	Yes

Notes: In this table, we report our estimates of the OLS regression for the effect of the elected a female candidate to the legislative house on policy variables associated with fiscal outcomes. Our units of observation are municipalities. We use our working sample restricted to have only municipalities with 9 or 10 seats in their respective legislative houses. On the column Full MM Pairs, we report the results for all minimum mixed-gender pairs on our sample. On the column Only 0 or 1 woman elected we report our effects for only the municipalities that had elected 0 or 1 woman, where the woman elected is the one from the minimum mixed-gender pair. Columns (1) and (3) report the OLS regression for each sample. We also display the same set of regressions but with year and region fixed effects in columns (2) and (4). Columns 'mean' in each sample is the mean for each variable in municipalities where a man won the election. Standard errors are in parenthesis. $*p < 0.1, **p < 0.05, *** < 0.01$.

at expenditures directed at child education. Other fiscal variables are not significant in any specification, many with point-estimate close to zero. From here we can infer that despite some moderate effects in health outcomes, in particular in terms of underweight live births, this effect is not associated with higher spending on health as a whole.

2.6 Robustness checks

For this experiment, we analyzed 20 policy outcomes that could have been affected by introducing a female legislator instead of a male one to the legislature. Similar to our balance tests, one major possible concern is that one of these policy outcomes could have been significant by chance, by virtue of the fact we performed the exercise for so many variables. As such, we also perform the Romano-Wolf procedure to calculate p-values that take into account the familywise error rate (FWER). Figure 45 displays these p-values alongside the p-values obtained from the OLS regressions and the p-values obtained from the randomized inference procedure. As in the previous cases, we performed 1000 replications for both the randomized inference estimates and Romano-Wolf ones. Despite a number of variables having had displayed p-values near the 10% threshold for the uncontrolled OLS regressions in the Full MM Pairs sample, only expenditures with education were significant when we control for the FWER. When we include year and region dummies, all variables lose significance even without correcting for multiple hypothesis testing. As for the sample with 0 or 1 woman elected, we find that the results for underweight births, number of live births and education expenditures are all significant after we correct for multiple hypothesis testing. However, none of them are significant after we include year and region dummies, despite the fact education expenditures and proportion of underweight births were significant according to the standard cluster-robust p-values from OLS regressions. As such, although our results lose significance when we include year and region fixed effects, we still find an impact of electing a woman on underweight births and education expenditures in the sample with zero or one woman elected.

Another concern we may have is regarding the fact we conducted our estimation with a working sample that discarded observations for which we did not have information for the full set of policy outcomes. This restriction is important to assure our results do not differ between outcomes because of sample size. However, such restriction can be seen as *ad hoc*. For this reason, we estimate the treatment effect and balance tests for the full sample with 9 or 10 seats. That is, we still use the full sample, but no longer restricting policy outcome regressions to have the same sample size.

Table 32 displays some descriptive statistics and the binomial distribution test for the full sample with 9 or 10 seats. Figure 46 displays the above mentioned test for different vote margins. For the most part, we cannot reject that the proportion of male and female elected candidates did not come from a binomial distribution with a probability of 0.5, as none of the observed values fall outside of the 95% confidence interval of its corresponding binomial distribution. However, it is worth keeping in mind that the p-values are very close to the 0.025

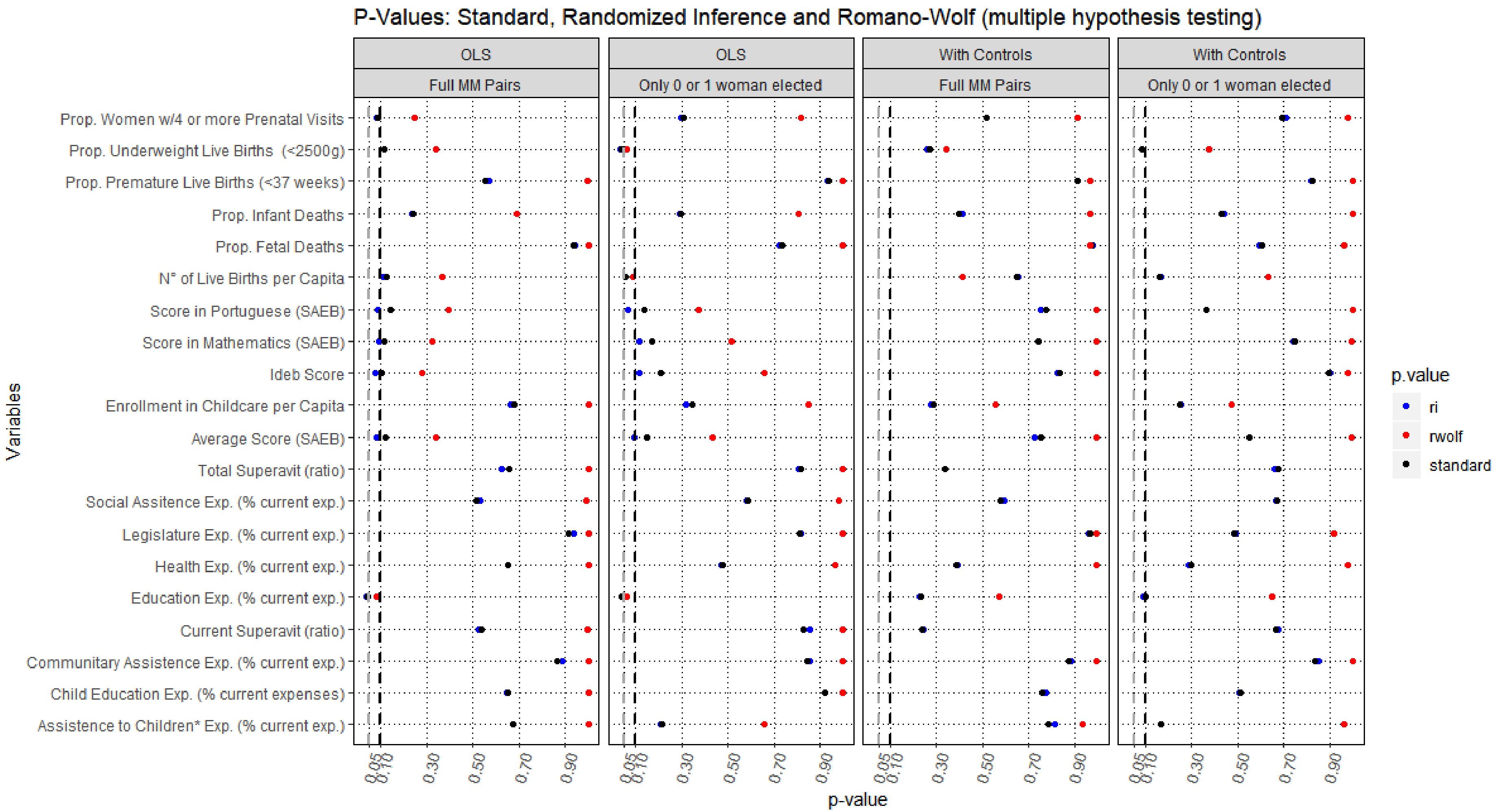
threshold on the full MM pairs sample, particularly so when we do not restrict the number of seats to 9 or 10.

As for the municipal characteristics, figures 47, 48 and 49 display the balance statistics for, respectively, characteristics of the municipalities, female candidates and male candidates. The balance tests are also worse in the full MM Pairs sample when it comes to municipal characteristics balance, with a few variables displaying imbalance, most of which are still significant even when correcting for multiple hypotheses. The sample with only 0 or 1 woman elected fares much better, but there is still an imbalance in terms of the average age in the municipalities. For candidate characteristics, there are slightly more imbalances in the full MM pairs sample too. As for the sample with only 0 and 1 women elected, there is an imbalance in three variables in the women candidate's characteristics, but only party number 45 persists with the Romano-Wolf p-values. As for male candidates, there is a regional imbalance, with the region dummy variable of being in the northeast having a p-value of 0.1059, very close to 10%.

In Tables 33, 34 and 35 we can check the magnitude of the differences as well as the F-test for joint orthogonality. Despite the significant results encountered for the full MM pairs sample, the F-test does not reject the null hypothesis of joint orthogonality. However, the 0 or 1 woman elected sample displays a p-value very close to 10% for the F-test. That said, the magnitude of differences between treated and controls is not big in any of the samples. None of the normalized differences are above 0.1 standard deviation. However, in Tables 34 and 35 the F-tests reject that there are no imbalances, which is in line with the p-values showing some variables significant even with Romano-Wolf p-values. But as in the previous case, the magnitude of differences is also small, never above 0.1 standard deviation. Overall, the balance tests and binomial distribution test fare worse in the full MM pairs sample but it is not as bad for the 0 or 1 woman elected sample. In the latter, there are regional imbalances and the avg. age in the control group is slightly higher. However, the magnitude of differences is rather small.

Now moving on to the policy outcomes, Tables 36, 37 and 38 presents the estimates for the full sample with 9 or 10 seats. Figure 50 compares the p-values from these regressions to the ones obtained through randomized inference and the Romano-Wolf procedure. Our results are mostly unchanged. There are more significant policy outcomes, but most lose significance when we include year and region fixed effects. Moreover, the Proportion of underweight live births is significant in all regressions. Live births and Education expenditures are only significant when we do not include year and region fixed effects, however. That said, all variables lose significance when we use Romano-Wolf p-values. As for the results here, the regression for the full MM Pairs might not be very reliable due to the sample not being well balanced. But even in the 0 or 1 woman elected sample, which has a better balance, we still fail to find results that are significant when we consider multiple hypothesis testing. This is in line with our findings for the working sample. As in the working sample, we found significant results for underweight births and education expenditures, but they are not significant in all specifications.

Figure 45 – P-values for policy variables in the working sample



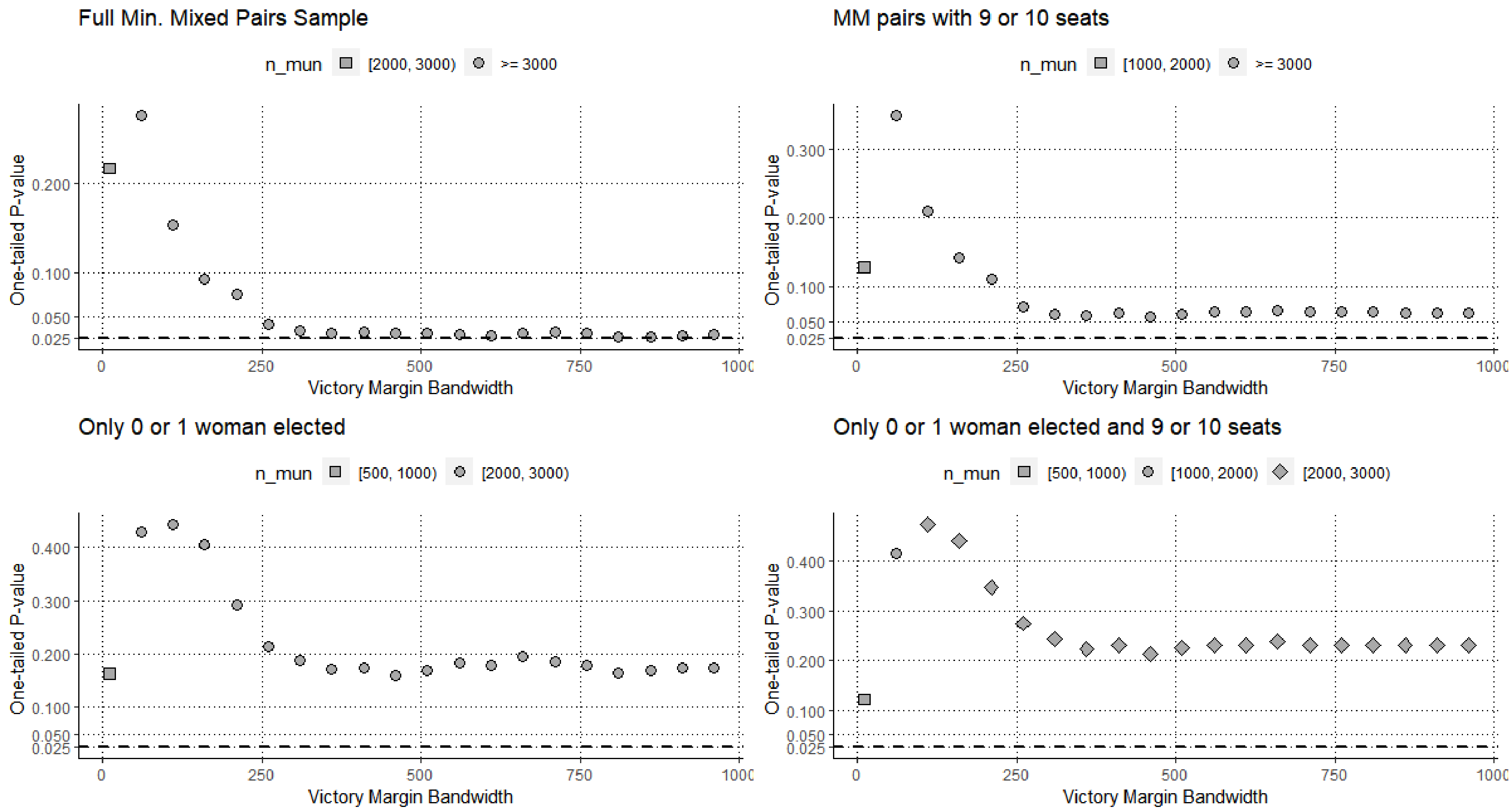
Notes: In this figure, we display the p-values for policy variables in the working sample using different methods. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate three p-values. The standard p-value is calculated from the estimated regression model; the ri p-values are obtained through the randomized inference procedure using 1000 replications; the rwolf p-values employ the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variable are the variables in the y-axis and the explanatory variable is a dummy that is equal to 1 when a candidate from the minimum mixed pair won the election. We use our working sample with 9 or 10 seats for all regressions.

Table 32 – Averages for the full sample of min. mixed pairs of candidates, but this time restricting the number of seats to 9 or 10 seats.

Subset of minimum mixed pairs											
					% males	% fem.	CI 95%		one-tail	p-value	
Year	Num. cand.	Num. elected	% males	% fem.	Elected	Elected	lwr	upr	p-value	≥ 0.025	Area
2004	4614	2307	50.00%	50.00%	48.03%	51.97%	47.94%	52.06%	0.0277	TRUE	95.44%
2008	4616	2308	50.00%	50.00%	52.04%	47.96%	47.96%	52.04%	0.0264	TRUE	95.20%
2012	3688	1844	50.00%	50.00%	53.31%	46.69%	47.72%	52.28%	0.0024	FALSE	95.23%
Total	12918	6459	50.00%	50.00%	51.32%	48.68%	48.78%	51.22%	0.0172	FALSE	95.07%
Only 0 or 1 woman elected											
					% males	% fem.	CI 95%		one-tail	p-value	
Year	Num cand.	Num elected	% males	% fem.	Elected	Elected	lwr	upr	p-value	≥ 0.025	Area
2004	1802	901	50.00%	50.00%	47.84%	52.16%	46.73%	53.27%	0.0913	TRUE	95.44%
2008	1768	884	50.00%	50.00%	52.49%	47.51%	46.72%	53.28%	0.0740	TRUE	95.28%
2012	1244	622	50.00%	50.00%	52.57%	47.43%	46.14%	53.86%	0.1069	TRUE	95.06%
Total	4814	2407	50.00%	50.00%	50.45%	49.55%	47.99%	52.01%	0.3269	TRUE	95.43%

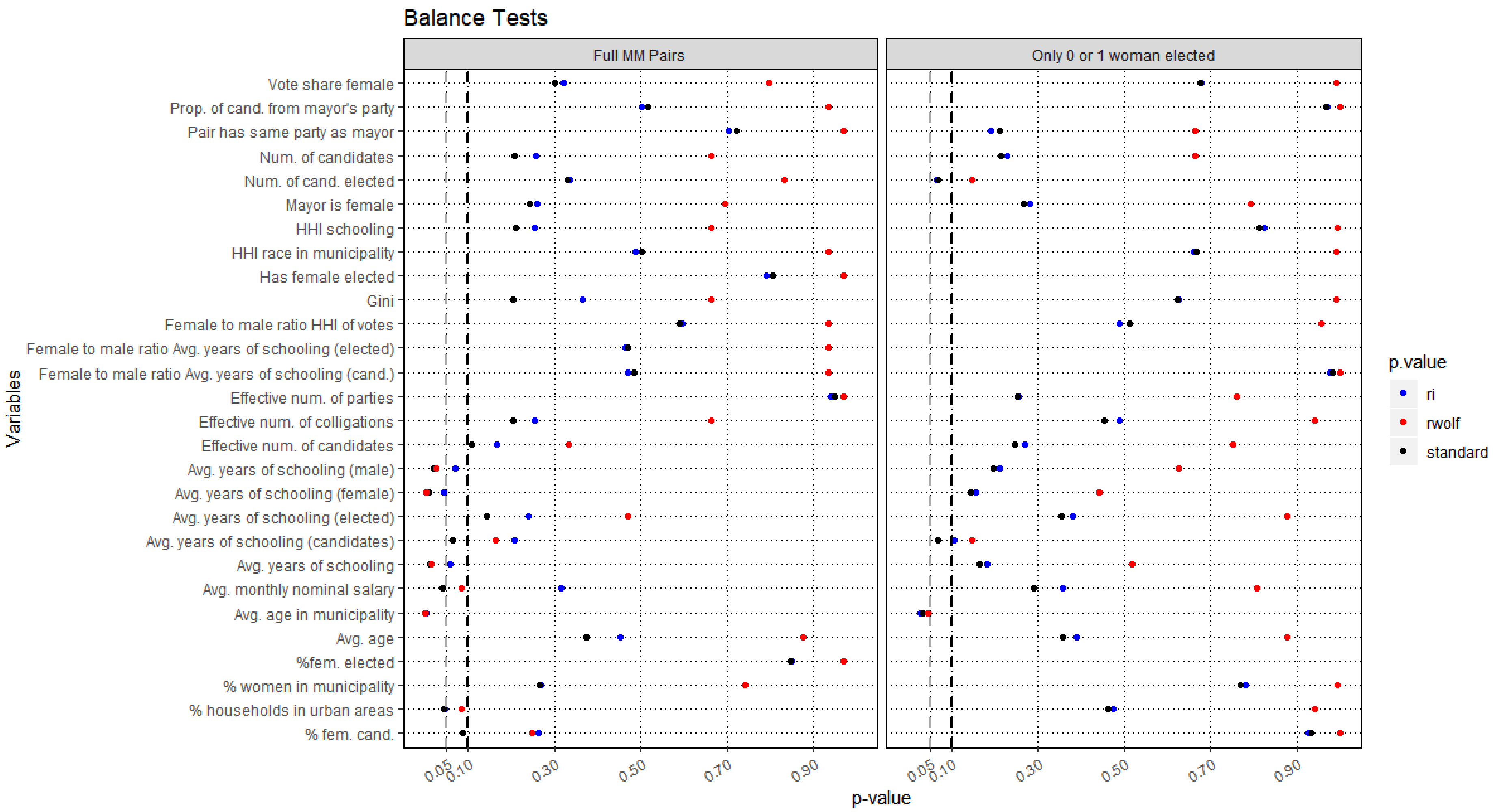
Notes: Table with number of candidates, number of candidates elected and the share of each sex in the subset of mixed gender pairs of candidates with the least difference in vote share in their respective municipality *for the full sample*, but counting only the municipalities that had 9 our 10 seats available. The first panel presents the subset in full and the second panel further restricts this sample to have only pairs of municipalities that had only 0 or 1 woman elected. We also calculated the theoretical confidence level of 95% for a binomial distribution with a probability of success of 0.5, where the number of successes is equal to the number of female elected candidates observed in our data. We then compare it to the observed number of females elected. The tail on the one-tail p-value is based on whether the % of females elected is below or above 50%. The last column, Area, displays the area covered by the confidence interval. It is not exactly 95% because the number of successes is a discrete variable.

Figure 46 – P-value of the binomial distribution test for both the full sample and with zero or one woman elected.



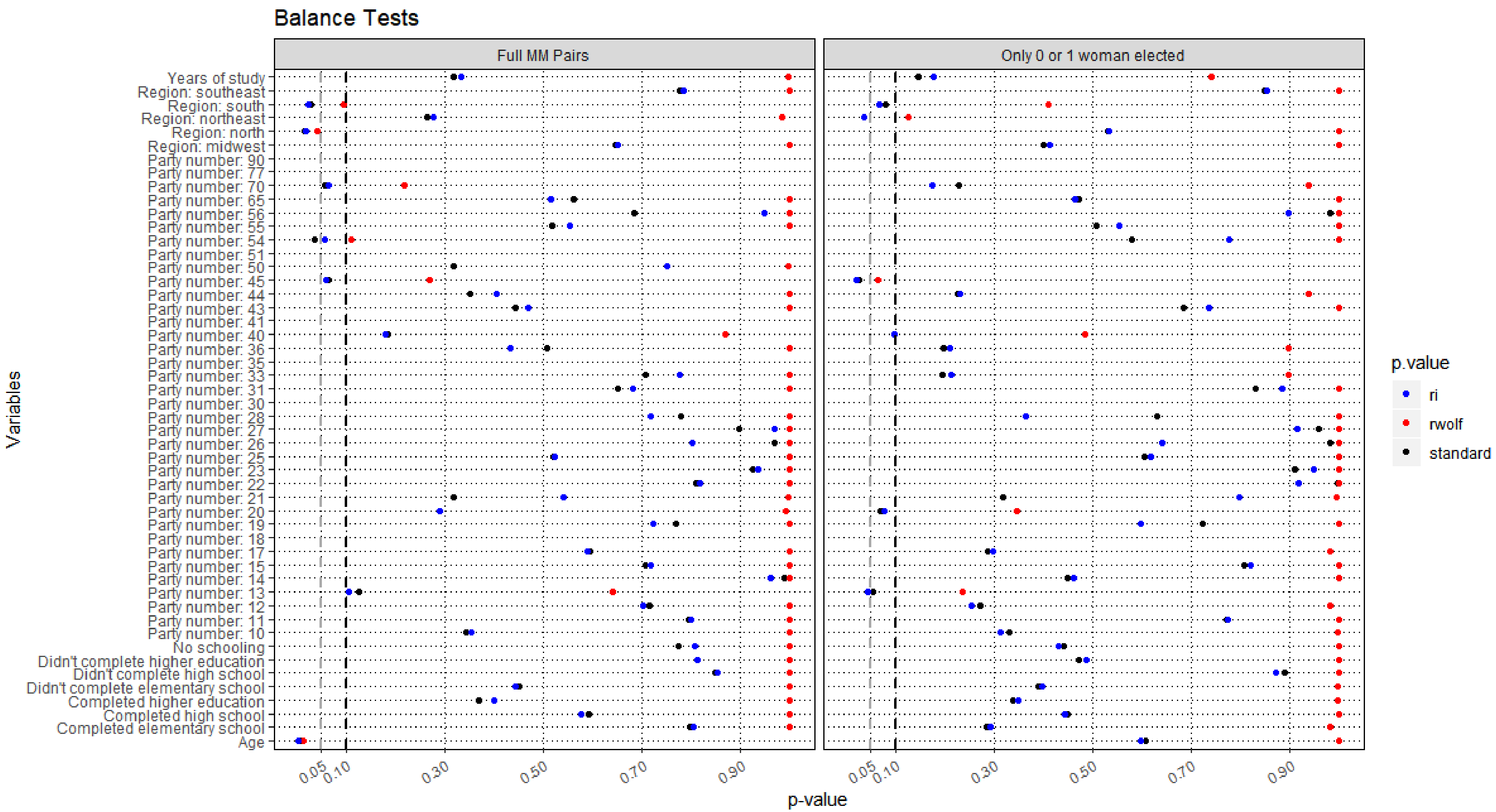
Notes: For each panel, we perform various binomial distribution tests where we compare the observed proportion of elected female candidates with its corresponding theoretical binomial distribution with a probability of success of 0.5. The idea here is to test if we can reject the hypothesis that the proportion of elected female candidates comes from a binomial distribution with a probability of success equal to 0.5. To evaluate that, we then calculate the corresponding one-tailed p-value, where the tail is based on whether the % of females elected is below or above 50%. We perform this test for different victory margins and plot the results in a graph. The dashed horizontal line marks the point from which we can reject the null hypothesis that the observed value comes from a binomial distribution with 95% of confidence. The victory margin is considered in terms of raw votes. We consider a total of 20 bandwidths ranging from 10 to 1000 votes for each subset analyzed. The shape of each plot point varies according to the number of municipalities considered. The first graph uses the full sample of minimum mixed gender pairs. We then restrict this sample to only municipalities with 9 and 10 seats in panel 2. In panels 3 and 4 we repeat the same procedure but for the restricted sample that considers only municipalities that elected only zero or one woman.

Figure 47 – Balance Statistics for Municipalities (full sample with 9 or 10 seats)



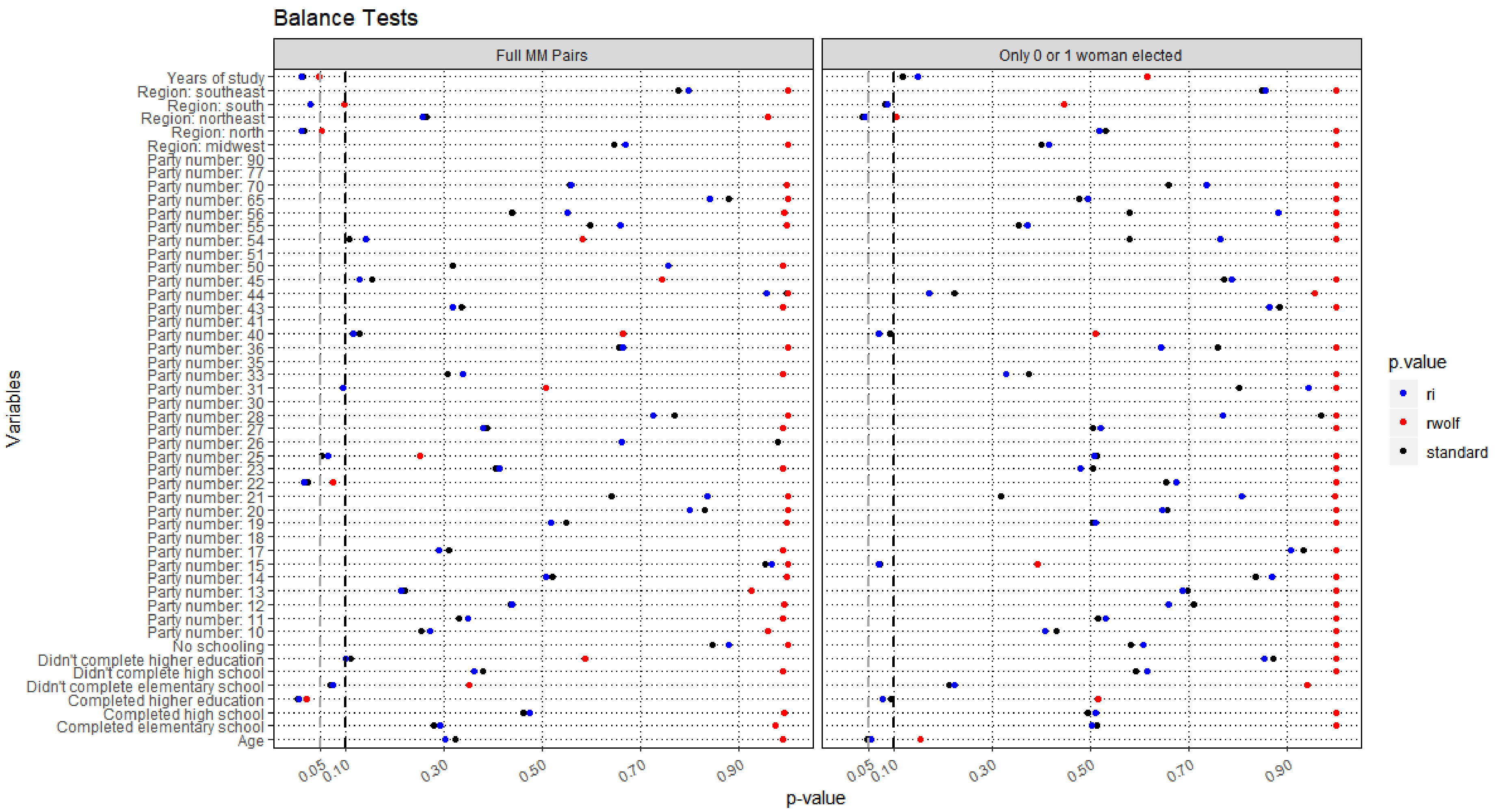
<u>Notes</u>: In this figure, we display the balance statistics municipalities in the full sample with 9 or 10 seats. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate three p-values. The standard p-value is calculated from the estimated regression model; the ri p-values are obtained through the randomized inference procedure using 1000 replications; the rwolf p-values employ the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variables are the variables in the y-axis and the explanatory variable is a dummy that is equal to 1 when a candidate from the minimum mixed pair won the election. We use our full sample with 9 or 10 seats for all regressions.

Figure 48 – Balance Statistics for Female Candidates (elected/non-elected, full sample with 9 or 10 seats)



Notes: In this figure, we display the balance statistics for elected and non-elected female candidates in the full sample with 9 or 10 seats. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate three p-values. The standard p-value is calculated from the estimated regression model; the ri p-values are obtained through the randomized inference procedure using 1000 replications; the rwolf p-values employ the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variables are the variables in the y-axis and the explanatory variable is a dummy that is equal to 1 when a candidate from the minimum mixed pair won the election. We use our full sample with 9 or 10 seats for all regressions and restrict the sample to have only female candidates.

Figure 49 – Balance Statistics for Male Candidates (elected/non-elected, full sample with 9 or 10 seats)



Notes: In this figure, we display the balance statistics for elected and non-elected male candidates in the full sample with 9 or 10 seats. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate three p-values. The standard p-value is calculated from the estimated regression model; the ri p-values are obtained through the randomized inference procedure using 1000 replications; the rwolf p-values employ the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variables are the variables in the y-axis and the explanatory variable is a dummy that is equal to 1 when a candidate from the minimum mixed pair won the election. We use our full sample with 9 or 10 seats for all regressions and restrict the sample to have only male candidates.

Table 33 – Balance between Treated and Controls for the full sample with 9 and 10 seats

	Full MM Pairs					0 or 1 woman elected				
	Treated (N = 3167)		Controls (N = 3292)			Treated (N = 1185)		Controls (N = 1222)		
Variables	Mean	Sd.	Mean	Sd.	Nd.	Mean	Sd.	Mean	Sd.	Nd.
Avg. years of schooling	5.096	1.107	5.165	1.067	-0.063	5.162	1.088	5.221	1.018	-0.057
Avg. years of schooling (female)	5.345	1.086	5.416	1.046	-0.066	5.397	1.063	5.459	1.002	-0.060
Avg. years of schooling (male)	4.850	1.156	4.916	1.118	-0.058	4.930	1.140	4.987	1.063	-0.052
Has female elected	0.626	0.484	0.629	0.483	-0.006					
Avg. age	42.824	2.610	42.881	2.580	-0.022	42.964	2.567	43.059	2.506	-0.037
Avg. years of schooling (candidates)	8.566	1.484	8.633	1.438	-0.046	8.444	1.529	8.556	1.469	-0.075
Female to male ratio Avg. years of schooling (cand.)	1.239	0.286	1.234	0.289	0.017	1.214	0.310	1.214	0.309	0.001
Avg. years of schooling (elected)	9.271	1.962	9.342	1.960	-0.036	8.954	2.019	9.029	1.965	-0.038
Female to male ratio Avg. years of schooling (elected)	1.326	0.535	1.315	0.508	0.023					
Gini	0.539	0.073	0.537	0.074	0.031	0.537	0.071	0.538	0.072	-0.020
HHI schooling	0.204	0.050	0.205	0.048	-0.031	0.207	0.051	0.207	0.049	-0.010
HHI race in municipality	0.559	0.137	0.561	0.138	-0.017	0.562	0.141	0.564	0.140	-0.017
Female to male ratio HHI of votes	0.171	0.166	0.173	0.174	-0.013	0.052	0.038	0.051	0.039	0.027
Avg. age in municipality	29.135	3.545	29.463	3.599	-0.092	29.225	3.429	29.522	3.450	-0.086
Num. of candidates	49.250	28.683	48.357	27.596	0.032	52.643	31.390	51.103	28.772	0.051
Num. of cand. elected	8.049	0.224	8.043	0.221	0.025	8.059	0.243	8.041	0.232	0.076
Effective num. of candidates	26.666	11.937	26.189	11.895	0.040	27.886	12.840	27.292	12.356	0.047
Effective num. of colligations	3.458	1.501	3.411	1.487	0.032	3.681	1.578	3.634	1.522	0.031
Effective num. of parties	6.166	2.433	6.170	2.398	-0.002	6.292	2.425	6.404	2.418	-0.046
% households in urban areas	0.596	0.215	0.607	0.215	-0.051	0.601	0.222	0.607	0.217	-0.030
Mayor is female	0.102	0.303	0.093	0.291	0.029	0.086	0.281	0.074	0.262	0.045
% fem. cand.	0.220	0.074	0.223	0.077	-0.043	0.199	0.076	0.198	0.078	0.003
%fem. elected	0.118	0.116	0.118	0.117	-0.005					
% women in municipality	0.492	0.014	0.492	0.014	-0.028	0.492	0.014	0.492	0.014	-0.012
Prop. of cand. from mayor's party	0.190	0.108	0.192	0.110	-0.016	0.184	0.101	0.184	0.106	-0.002
Avg. monthly nominal salary	481.396	293.495	496.388	287.156	-0.052	483.258	282.690	495.166	270.056	-0.043
Pair has same party as mayor	0.244	0.429	0.240	0.427	0.009	0.268	0.443	0.246	0.431	0.050
Vote share female	0.153	0.075	0.155	0.077	-0.026	0.094	0.046	0.094	0.047	0.017
F-Test for Joint Orthogonality of All Variables (p-value)					0.808					0.117

Notes: This table displays the means and standard deviations (Sd.) for balance covariates for our full sample with 9 and 10 seats. The first panel presents the full minimum mixed pairs sample and the second panel presents the selected statistics only for the municipalities that elected 0 or 1 woman. Our treatment is whether the female candidate won or not. We also calculated the normalized difference (Nd.) between treated and controls. This difference is defined as $(\mu_t - \mu_c)/((\sigma_t^2 + \sigma_c^2)/2)^{1/2})$, where μ_c and σ_c are the mean and standard deviation of the control and μ_t and σ_t are the mean and standard deviation of the treated. In the last row we also perform an F-test based on an OLS regression using the variables listed. In the Full MM Pairs sample we didn't include the variable "Elected at least one female" in order to avoid Perfect multicollinearity.

Table 34 – Balance between Treated and Controls for the full sample with 9 and 10 seats (female candidates only)

	Full MM Pairs					0 or 1 woman elected				
	Treated (N = 3167)		Controls (N = 3292)			Treated (N = 1185)		Controls (N = 1222)		
Variables	Mean	Sd.	Mean	Sd.	Nd.	Mean	Sd.	Mean	Sd.	Nd.
Party number: 10	0.008	0.087	0.010	0.098	-0.023	0.008	0.092	0.012	0.110	-0.038
Party number: 11	0.099	0.298	0.097	0.296	0.006	0.097	0.296	0.101	0.301	-0.012
Party number: 12	0.062	0.242	0.064	0.245	-0.009	0.062	0.241	0.073	0.260	-0.045
Party number: 13	0.083	0.276	0.094	0.292	-0.038	0.086	0.281	0.110	0.313	-0.079
Party number: 14	0.070	0.255	0.070	0.255	0.000	0.078	0.268	0.070	0.255	0.031
Party number: 15	0.169	0.375	0.173	0.378	-0.010	0.181	0.385	0.177	0.382	0.010
Party number: 17	0.011	0.106	0.010	0.100	0.013	0.009	0.096	0.014	0.117	-0.043
Party number: 19	0.004	0.061	0.004	0.065	-0.007	0.005	0.071	0.004	0.064	0.014
Party number: 20	0.020	0.139	0.016	0.126	0.026	0.020	0.141	0.011	0.103	0.078
Party number: 21	0.000	0.000	0.000	0.017	-0.025	0.000	0.000	0.001	0.029	-0.040
Party number: 22	0.059	0.236	0.061	0.239	-0.006	0.053	0.224	0.053	0.225	0.000
Party number: 23	0.035	0.185	0.035	0.184	0.002	0.029	0.167	0.029	0.169	-0.005
Party number: 25	0.102	0.302	0.097	0.296	0.016	0.096	0.295	0.090	0.286	0.021
Party number: 26	0.001	0.025	0.001	0.025	0.001	0.001	0.029	0.001	0.029	0.001
Party number: 27	0.007	0.085	0.007	0.083	0.003	0.005	0.071	0.005	0.070	0.002
Party number: 28	0.004	0.066	0.005	0.070	-0.006	0.003	0.050	0.002	0.040	0.020
Party number: 31	0.007	0.085	0.006	0.080	0.011	0.006	0.077	0.007	0.081	-0.008
Party number: 33	0.009	0.092	0.009	0.097	-0.009	0.005	0.071	0.010	0.099	-0.055
Party number: 36	0.006	0.075	0.007	0.083	-0.016	0.006	0.077	0.011	0.103	-0.052
Party number: 40	0.043	0.202	0.050	0.217	-0.033	0.038	0.191	0.052	0.221	-0.066
Party number: 43	0.018	0.134	0.021	0.143	-0.019	0.019	0.138	0.017	0.130	0.017
Party number: 44	0.010	0.098	0.008	0.087	0.024	0.014	0.115	0.008	0.090	0.051
Party number: 45	0.129	0.336	0.115	0.318	0.046	0.142	0.349	0.111	0.315	0.092
Party number: 50	0.000	0.000	0.000	0.017	-0.025					
Party number: 54	0.000	0.018	0.002	0.046	-0.052	0.001	0.029	0.002	0.040	-0.023
Party number: 55	0.027	0.162	0.024	0.154	0.016	0.024	0.155	0.020	0.142	0.027
Party number: 56	0.001	0.025	0.001	0.030	-0.010	0.001	0.029	0.001	0.029	0.001
Party number: 65	0.007	0.083	0.008	0.090	-0.014	0.005	0.071	0.007	0.086	-0.029
Party number: 70	0.009	0.095	0.005	0.072	0.047	0.007	0.082	0.003	0.057	0.049
No schooling	0.016	0.126	0.017	0.129	-0.007	0.018	0.132	0.014	0.117	0.031
Didn't complete elementary school	0.100	0.300	0.106	0.308	-0.018	0.114	0.318	0.103	0.304	0.035
Completed elementary school	0.098	0.297	0.096	0.294	0.006	0.108	0.311	0.095	0.293	0.043
Didn't complete high school	0.040	0.195	0.041	0.198	-0.005	0.043	0.203	0.044	0.206	-0.006
Completed high school	0.352	0.478	0.358	0.480	-0.013	0.348	0.476	0.363	0.481	-0.031
Didn't complete higher education	0.061	0.240	0.060	0.237	0.006	0.057	0.233	0.051	0.220	0.029
Completed higher education	0.333	0.471	0.323	0.468	0.022	0.312	0.464	0.331	0.471	-0.039
Years of study	11.225	3.654	11.135	3.685	0.025	10.983	3.746	11.198	3.627	-0.058
Age	43.236	9.763	43.923	10.038	-0.069	43.798	10.017	44.005	9.796	-0.021
Region: midwest	0.089	0.284	0.092	0.289	-0.012	0.084	0.277	0.093	0.291	-0.034
Region: northeast	0.330	0.470	0.317	0.466	0.027	0.285	0.452	0.248	0.432	0.084
Region: north	0.086	0.280	0.070	0.255	0.060	0.071	0.257	0.065	0.246	0.025
Region: southeast	0.283	0.450	0.286	0.452	-0.007	0.338	0.473	0.341	0.474	-0.008
Region: south	0.212	0.409	0.235	0.424	-0.054	0.223	0.416	0.253	0.435	-0.071
F-Test for Joint Orthogonality of All Variables (p-value)					0.000					0.000

Notes: This table displays the means and standard deviations (Sd.) for balance covariates for our full sample with 9 and 10 seats. The first panel presents the full minimum mixed pairs sample and the second panel presents the selected statistics only for the municipalities that elected 0 or 1 woman. Our treatment is whether the female candidate won or not. We also calculated the normalized difference (Nd.) between treated and controls. This difference is defined as $(\mu_t - \mu_c)/((\sigma_t^2 + \sigma_c^2)/2)^{1/2})$, where μ_c and σ_c are the mean and standard deviation of the control and μ_t and σ_t are the mean and standard deviation of the treated. In the last row we also perform an F-test based on an OLS regression using the variables listed. In the Full MM Pairs sample we didn't include the party number 70 in order to avoid perfect multicollinearity. Likewise, we removed party number 56 from the sample with 0 or 1 woman elected.

Table 35 – Balance between Treated and Controls for the full sample with 9 and 10 seats (male candidates only)

	Full MM Pairs					0 or 1 woman elected				
	Treated (N = 3167)		Controls (N = 3292)			Treated (N = 1185)		Controls (N = 1222)		
Variables	Mean	Sd.	Mean	Sd.	Nd.	Mean	Sd.	Mean	Sd.	Nd.
Party number: 10	0.009	0.097	0.012	0.111	-0.029	0.008	0.087	0.011	0.103	-0.032
Party number: 11	0.101	0.301	0.094	0.292	0.024	0.105	0.307	0.097	0.297	0.027
Party number: 12	0.068	0.252	0.063	0.244	0.019	0.063	0.244	0.060	0.237	0.015
Party number: 13	0.081	0.273	0.090	0.286	-0.030	0.091	0.288	0.096	0.294	-0.016
Party number: 14	0.068	0.252	0.072	0.259	-0.016	0.077	0.266	0.074	0.263	0.009
Party number: 15	0.166	0.372	0.166	0.372	0.001	0.183	0.387	0.155	0.363	0.074
Party number: 17	0.012	0.110	0.010	0.098	0.025	0.013	0.112	0.012	0.110	0.003
Party number: 19	0.004	0.066	0.005	0.074	-0.015	0.003	0.050	0.004	0.064	-0.027
Party number: 20	0.019	0.136	0.018	0.134	0.005	0.014	0.119	0.012	0.110	0.018
Party number: 21	0.001	0.025	0.000	0.017	0.015	0.000	0.000	0.001	0.029	-0.040
Party number: 22	0.057	0.231	0.070	0.255	-0.056	0.053	0.224	0.057	0.232	-0.018
Party number: 23	0.034	0.182	0.038	0.192	-0.021	0.031	0.174	0.036	0.186	-0.027
Party number: 25	0.106	0.308	0.092	0.289	0.048	0.092	0.289	0.100	0.300	-0.027
Party number: 26	0.000	0.018	0.000	0.017	0.001					
Party number: 27	0.008	0.090	0.006	0.080	0.022	0.006	0.077	0.008	0.090	-0.027
Party number: 28	0.005	0.073	0.005	0.070	0.007	0.003	0.050	0.002	0.050	0.002
Party number: 31	0.008	0.090	0.005	0.070	0.042	0.004	0.065	0.005	0.070	-0.010
Party number: 33	0.008	0.090	0.006	0.078	0.025	0.007	0.082	0.004	0.064	0.036
Party number: 36	0.006	0.079	0.005	0.074	0.011	0.008	0.087	0.007	0.081	0.013
Party number: 40	0.044	0.205	0.052	0.222	-0.038	0.041	0.197	0.055	0.228	-0.067
Party number: 43	0.021	0.142	0.024	0.153	-0.024	0.024	0.152	0.025	0.155	-0.006
Party number: 44	0.008	0.090	0.008	0.090	0.000	0.010	0.100	0.006	0.075	0.050
Party number: 45	0.125	0.330	0.113	0.317	0.035	0.132	0.338	0.128	0.334	0.012
Party number: 50	0.000	0.000	0.000	0.017	-0.025					
Party number: 54	0.000	0.018	0.002	0.039	-0.040	0.001	0.029	0.002	0.040	-0.023
Party number: 55	0.021	0.144	0.023	0.150	-0.013	0.017	0.129	0.022	0.147	-0.038
Party number: 56	0.001	0.025	0.001	0.035	-0.019	0.001	0.029	0.002	0.040	-0.023
Party number: 65	0.010	0.098	0.009	0.097	0.004	0.009	0.096	0.012	0.110	-0.029
Party number: 70	0.007	0.081	0.008	0.089	-0.015	0.006	0.077	0.007	0.086	-0.018
No schooling	0.045	0.208	0.046	0.210	-0.005	0.049	0.216	0.044	0.206	0.023
Didn't complete elementary school	0.262	0.440	0.243	0.429	0.045	0.259	0.438	0.237	0.426	0.050
Completed elementary school	0.159	0.366	0.150	0.357	0.027	0.146	0.353	0.155	0.363	-0.027
Didn't complete high school	0.053	0.224	0.058	0.234	-0.022	0.052	0.223	0.057	0.232	-0.022
Completed high school	0.319	0.466	0.311	0.463	0.018	0.322	0.467	0.309	0.462	0.028
Didn't complete higher education	0.037	0.189	0.045	0.207	-0.039	0.039	0.193	0.040	0.196	-0.007
Completed higher education	0.124	0.329	0.148	0.355	-0.070	0.133	0.340	0.157	0.364	-0.068
Years of study	8.680	4.076	8.937	4.153	-0.062	8.743	4.154	9.009	4.151	-0.064
Age	43.542	10.647	43.285	10.468	0.024	43.981	10.433	43.157	10.151	0.080
Region: midwest	0.089	0.284	0.092	0.289	-0.012	0.084	0.277	0.093	0.291	-0.034
Region: northeast	0.330	0.470	0.317	0.466	0.027	0.285	0.452	0.248	0.432	0.084
Region: north	0.086	0.280	0.070	0.255	0.060	0.071	0.257	0.065	0.246	0.025
Region: southeast	0.283	0.450	0.286	0.452	-0.007	0.338	0.473	0.341	0.474	-0.008
Region: south	0.212	0.409	0.235	0.424	-0.054	0.223	0.416	0.253	0.435	-0.071
F-Test for Joint Orthogonality of All Variables (p-value)					0.005					0.000

Notes: This table displays the means and standard deviations (Sd.) for balance covariates for our full sample with 9 and 10 seats. The first panel presents the full minimum mixed pairs sample and the second panel presents the selected statistics only for the municipalities that elected 0 or 1 woman. Our treatment is whether the female candidate won or not. We also calculated the normalized difference (Nd.) between treated and controls. This difference is defined as $(\mu_t - \mu_c)/((\sigma_t^2 + \sigma_c^2)/2)^{1/2}$, where μ_c and σ_c are the mean and standard deviation of the control and μ_t and σ_t are the mean and standard deviation of the treated. In the last row we also perform an F-test based on an OLS regression using the variables listed. In the Full MM Pairs sample we didn't include the party number 56 in order to avoid perfect multicollinearity.

Table 36 – Effects of Electing a Councilwoman on Health Outcomes (full sample with 9 or 10 seats)

	Full MM Pairs					Only 0 or 1 woman elected				
Variable	Mean	(1)	N	(2)	N	Mean	(3)	N	(4)	N
Prop. Underweight	8.0136	-0.3011***	6312	-0.2449***	6284	8.2037	-0.4321***	2355	-0.3728***	2341
Live Births ($< 2500g$)	(0.0633)	(0.0841)	6312	(0.0804)	6284	(0.1027)	(0.1409)	2355	(0.1355)	2341
Prop. Premature	8.7916	-0.2789**	6286	-0.1803	6255	8.69	-0.1986	2348	-0.0708	2334
Live Births (< 37 weeks)	(0.0814)	(0.1181)	6286	(0.1126)	6255	(0.1297)	(0.1955)	2348	(0.1876)	2334
Prop. Women w/4	91.4057	-0.59***	6459	-0.0958	6422	91.4637	-0.2872	2407	0.2185	2392
or more Prenatal Visits	(0.1695)	(0.2263)	6459	(0.1824)	6422	(0.2637)	(0.3543)	2407	(0.2879)	2392
N° of Live Births	1.3773	0.0375***	6459	0.0131*	6422	1.3511	0.043***	2407	0.0164	2392
per Capita	(0.0078)	(0.0104)	6459	(0.0075)	6422	(0.0118)	(0.0163)	2407	(0.0118)	2392
Prop. Fetal Deaths	1.6226	0.0185	4518	0.0288	4501	1.6416	-0.0423	1695	0.0175	1683
	(0.0248)	(0.0344)	4518	(0.0318)	4501	(0.0443)	(0.0592)	1695	(0.0527)	1683
Prop. Infant Deaths	1.945	0.0313	4792	0.0228	4780	1.97	0.0038	1831	0.0241	1822
	(0.0277)	(0.0371)	4792	(0.0356)	4780	(0.0459)	(0.0607)	1831	(0.0566)	1822
Year and Region Fixed Effects:		No		Yes			No		Yes	

Notes: In this table we report our estimates of the OLS regression for the effect of the elected a female candidate to the legislative house on policy variables associated with health outcomes. Our units of observation are municipalities. We use our sample restricted to have only municipalities with 9 or 10 seats in their respective legislative houses, but this time allowing the regressions to have different number of observations. On the column Full MM Pairs, we report the results for all minimum mixed-gender pairs on our sample. On the column Only 0 or 1 woman elected we report our effects for only the municipalities that had elected 0 or 1 woman, where the woman elected is the one from the minimum mixed-gender pair. Columns (1) and (3) report the OLS regression without controls for each sample. Columns (2) and (4) report the same regressions but with controls. We use as controls the variables that didn't pass the balance tests. These are avg. monthly nominal salary, % households in urban areas, avg. age in the municipality, avg. years of study in the municipality, av. number of children per woman, % households with no access to electricity, % households with access to the garbage collection system, % households with access to piped water, % households with access to bathroom and water supply, avg. years of schooling (candidates), Herfindahl-Hirschman index (HHI) of male votes and avg. age of candidates. We also include regional dummy variables. We did not include dummy variables for each party. Columns 'mean' in each sample is the mean for each variable in municipalities where a man won the election. Standard errors are in parenthesis. $*p < 0.1, **p < 0.05, *** < 0.01$.

Table 37 – Effects of Electing a Councilwoman on Educational Outcomes (full sample with 9 or 10 seats)

	Full MM Pairs					Only 0 or 1 woman elected				
Variable	Mean	(1)	N	(2)	N	Mean	(3)	N	(4)	N
Score in Mathematics (SAEB)	202.2125	-2.0527***	5789	-0.7571	5765	204.2204	-2.5129**	2136	-0.7853	2123
	(0.5372)	(0.6898)	5789	(0.4708)	5765	(0.8187)	(1.107)	2136	(0.7732)	2123
Score in Portuguese (SAEB)	184.1175	-1.861***	5789	-0.6246	5765	185.7191	-2.3905**	2136	-0.7595	2123
	(0.4815)	(0.6344)	5789	(0.4071)	5765	(0.7278)	(1.0088)	2136	(0.6673)	2123
Average Score (SAEB)	5.1707	-0.073***	5789	-0.0258	5765	5.2381	-0.0914**	2136	-0.0288	2123
	(0.0188)	(0.0244)	5789	(0.016)	5765	(0.0285)	(0.0389)	2136	(0.0262)	2123
Ideb Score	4.6491	-0.0977***	5789	-0.0374*	5765	4.7331	-0.1179**	2136	-0.0344	2123
	(0.0235)	(0.0302)	5789	(0.0192)	5765	(0.0358)	(0.0483)	2136	(0.0313)	2123
Enrollment in Childcare	1.1357	0.0092	6459	0.0206	6422	1.0705	0.0392	2407	0.039	2392
per capita	(0.0175)	(0.0227)	6459	(0.0212)	6422	(0.0252)	(0.0346)	2407	(0.0321)	2392
Year and Region Fixed Effects:		No		Yes			No		Yes	

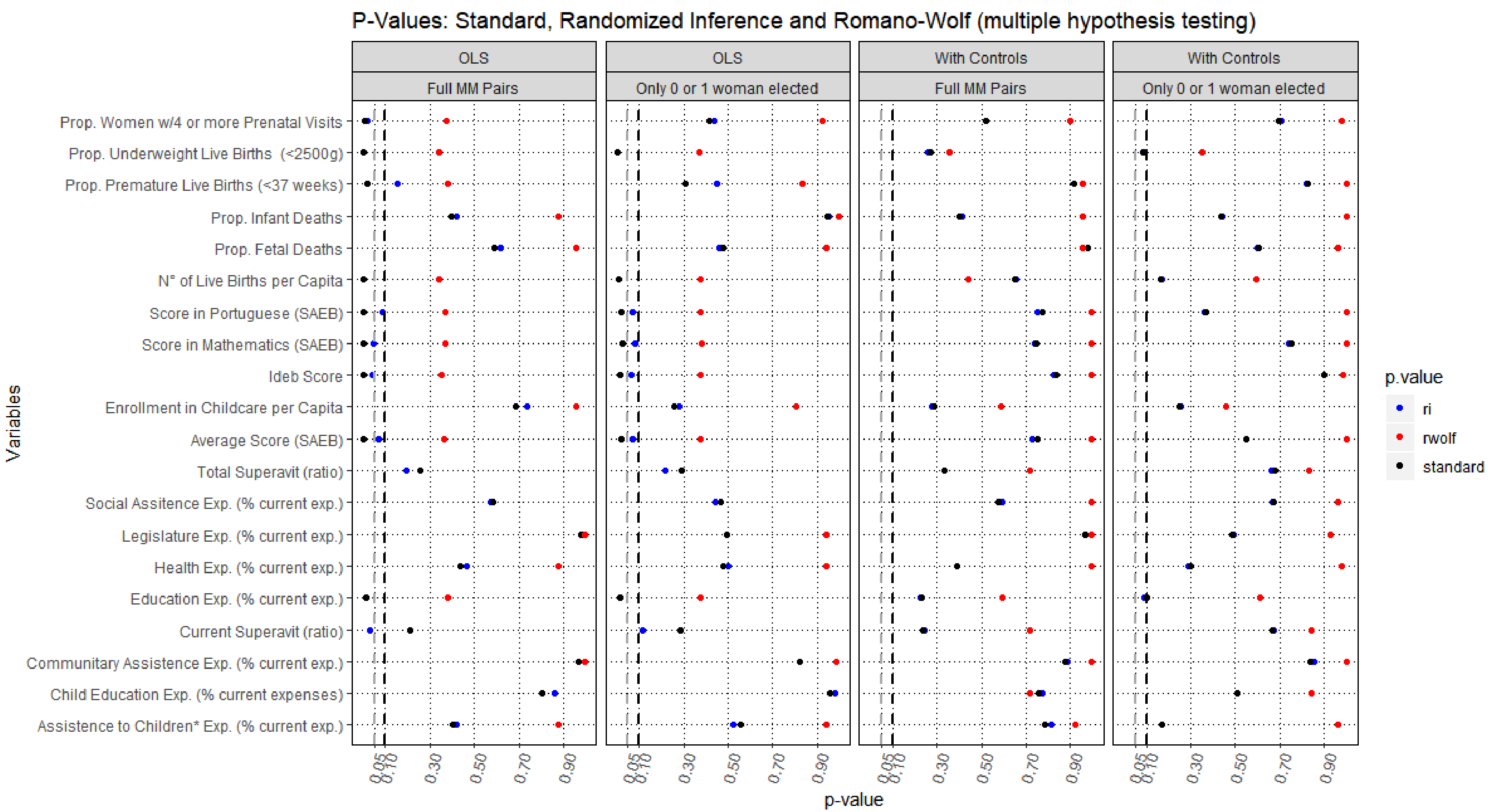
Notes: In this table, we report our estimates of the OLS regression for the effect of the elected a female candidate to the legislative house on policy variables associated with educational outcomes. Our units of observation are municipalities. We use our sample restricted to have only municipalities with 9 or 10 seats in their respective legislative houses, but this time allowing the regressions to have a different number of observations. On the column Full MM Pairs, we report the results for all minimum mixed-gender pairs on our sample. On the column Only 0 or 1 woman elected we report our effects for only the municipalities that had elected 0 or 1 woman, where the woman elected is the one from the minimum mixed-gender pair. Columns (1) and (3) report the OLS regression without controls for each sample. Columns (2) and (4) report the same regressions but with controls. We use as controls the variables that didn't pass the balance tests. These are avg. monthly nominal salary, % households in urban areas, avg. age in the municipality, avg. years of study in the municipality, av. number of children per woman, % households with no access to electricity, % households with access to the garbage collection system, % households with access to piped water, % households with access to bathroom and water supply, avg. years of schooling (candidates), Herfindahl-Hirschman index (HHI) of male votes and avg. age of candidates. We also include regional dummy variables. We did not include dummy variables for each party. Columns 'mean' in each sample is the mean for each variable in municipalities where a man won the election. Standard errors are in parenthesis. $*p < 0.1, **p < 0.05, *** < 0.01$.

Table 38 – Effects of Electing a Councilwoman on Fiscal Outcomes (full sample with 9 or 10 seats)

	Full MM Pairs					Only 0 or 1 woman elected				
Variable	Mean	(1)	N	(2)	N	Mean	(3)	N	(4)	N
Total Superavit	107.7616	3.7581	6307	4.2752	6270	107.4961	9.2277	2352	9.9651	2337
(ratio)	(0.1892)	(3.3036)	6307	(3.4244)	6270	(0.314)	(8.7585)	2352	(9.0669)	2337
Current Superavit	123.2073	4.5272	6307	4.8155	6270	123.6723	10.1987	2352	10.8995	2337
(ratio)	(0.2243)	(3.5997)	6307	(3.7355)	6270	(0.3684)	(9.5558)	2352	(9.8915)	2337
Legislature Exp.	0.0381	0	6069	0.0002	6033	0.0379	-0.0005	2273	-0.0001	2258
(% current exp.)	(6e-04)	(7e-04)	6069	(6e-04)	6033	(5e-04)	(7e-04)	2273	(7e-04)	2258
Communitary Assistence Exp.	0.0282	0	6244	0.0002	6208	0.028	-0.0002	2334	0	2319
(% current exp.)	(4e-04)	(5e-04)	6244	(5e-04)	6208	(7e-04)	(9e-04)	2334	(9e-04)	2319
Social Assitence Exp.	0.0462	0.0003	6294	0.0002	6257	0.0455	0.0007	2349	0.0007	2334
(% current exp.)	(4e-04)	(6e-04)	6294	(6e-04)	6257	(7e-04)	(0.001)	2349	(0.001)	2334
Child Education Exp.	0.0335	-0.0002	6137	0.0005	6101	0.0341	-0.0001	2292	0.0005	2277
(% current expenses)	(7e-04)	(9e-04)	6137	(8e-04)	6101	(0.0011)	(0.0015)	2292	(0.0013)	2277
Education Exp.	0.338	0.0063**	6297	0.001	6260	0.3276	0.0095**	2350	0.0023	2335
(% current exp.)	(0.002)	(0.0025)	6297	(0.0019)	6260	(0.0028)	(0.0038)	2350	(0.003)	2335
Health Exp.	0.2508	-0.0011	6297	-0.0006	6260	0.2525	-0.0017	2349	-0.0017	2334
(% current exp.)	(0.0011)	(0.0014)	6297	(0.0014)	6260	(0.0017)	(0.0024)	2349	(0.0023)	2334
Assistence to Children* Exp.	0.0086	-0.0002	6050	-0.0002	6013	0.0085	0.0002	2274	0.0003	2259
(% current exp.)	(2e-04)	(2e-04)	6050	(2e-04)	6013	(3e-04)	(4e-04)	2274	(4e-04)	2259
Year and Region Fixed Effects:		No		Yes			No		Yes	

Notes: In this table, we report our estimates of the OLS regression for the effect of the elected a female candidate to the legislative house on policy variables associated with fiscal outcomes. Our units of observation are municipalities. We use our sample restricted to have only municipalities with 9 or 10 seats in their respective legislative houses, but this time allowing the regressions to have different number of observations. On the column Full MM Pairs, we report the results for all minimum mixed-gender pairs on our sample. On the column Only 0 or 1 woman elected we report our effects for only the municipalities that had elected 0 or 1 woman, where the woman elected is the one from the minimum mixed-gender pair. Columns (1) and (3) report the OLS regression without controls for each sample. Columns (2) and (4) report the same regressions but with controls. We use as controls the variables that didn't pass the balance tests. These are avg. monthly nominal salary, % households in urban areas, avg. age in the municipality, avg. years of study in the municipality, av. number of children per woman, % households with no access to electricity, % households with access to the garbage collection system, % households with access to piped water, % households with access to bathroom and water supply, avg. years of schooling (candidates), Herfindahl-Hirschman index (HHI) of male votes and avg. age of candidates. We also include regional dummy variables. We did not include dummy variables for each party. Columns 'mean' in each sample is the mean for each variable in municipalities where a man won the election. Standard errors are in parenthesis. $*p < 0.1, **p < 0.05, *** < 0.01$.

Figure 50 – P-values for policy variables in the full sample with 9 or 10 seats



Notes: In this figure, we display the p-values for policy variables in the full sample 9 or 10 seats using different methods. Each point in the graph represents the p-value from an OLS regression with cluster-robust standard errors. We calculate three p-values. The standard p-value is calculated from the estimated regression model; the ri p-values are obtained through the randomized inference procedure using 1000 replications; the rwolf p-values employ the Romano-Wolf correction to control for the familywise error rate (FWER) using 1000 bootstrap repetitions. The dependent variables are the variables in the y-axis and the explanatory variable is a dummy that is equal to 1 when a candidate from the minimum mixed pair won the election. We use our full sample with 9 or 10 seats for all regressions.

2.7 Conclusion

In this paper, we were unable to find very robust evidence that an electoral victory of a woman leads to significant and positive impacts in prenatal and child health care nor educational outcomes. We did find a significant effect of electing women in decreasing the proportion of underweight live births, however, these results are sensitive to the inclusion of year and region fixed effects and to the fact we tested multiple hypotheses in this paper. The same is true for educational outcomes, we were able to find an increase in education expenditures as a proportion of the current expenditures, and these results are still significant when we control for the familywise error rate. However, they lose significance when we include year and region fixed effects. Moreover, we cannot rule out the possibility that these results are false positives when considering a less restricted sample. We do not find, however, an improvement in early childhood educational outcomes nor in expenditures related to infant education.

These results are in contrast with part of the literature, that finds significant and often sizable effects of electing women on health and educational outcomes (Brollo and Troiano (2016) and Bhalotra and Clots-Figueras (2014)). At first, these results would seem to corroborate with the median voter view of the political process implied by Downs (1957). However, we believe there are a number of reasons to believe this might not be the case. It is important to notice that we differ from previous studies for focusing on close elections in municipal elections for the legislative. As such, there are two important differences to emphasize. First, by examining close elections in this context we removed from our sample the most competitive and popular candidates, and those characteristics might be associated with more political influence to push their political agendas. Moreover, in this experiment, we do not consider how changes in political institutions could affect female participation. It could be the case that there are other prerequisites for elected women at the margin to truly be able to influence policy choices, such as the presence of other women in the legislative house or having a female mayor.

These possible interactions and caveats in our strategy lead us to believe were not fully explored in this version of the paper and might be an exciting direction for future research. As such, the precise causal mechanisms remain an open question. Knowing when we should expect an increase in female participation to matter or not is an important step into getting a better grasp of this question.

Bibliography

ALESINA, A.; FERRARA, E. L. Preferences for redistribution in the land of opportunities. *Journal of public Economics*, Elsevier, v. 89, n. 5-6, p. 897–931, 2005. Citado 2 vezes nas páginas 17 and 45.

ARVATE, P.; FIRPO, S.; PIERI, R. Can women's performance in elections determine the engagement of adolescent girls in politics? *European Journal of Political Economy*, Elsevier, p. 102045, 2021. Citado 2 vezes nas páginas 18 and 100.

ATHEY, S.; IMBENS, G. W. The econometrics of randomized experiments. In: *Handbook of economic field experiments*. [S.l.]: Elsevier, 2017. v. 1, p. 73–140. Citado na página 112.

BESLEY, T.; CASE, A. Political institutions and policy choices: Evidence from the United States. *Journal of Economic Literature*, v. 41, n. 1, p. 7–73, 2003. Citado na página 45.

BESLEY, T.; COATE, S. An economic model of representative democracy. *The Quarterly Journal of Economics*, v. 112, n. 1, p. 85–114, 1997. Citado 2 vezes nas páginas 17 and 100.

BHALOTRA, S.; CLOTS-FIGUERAS, I. "Health and the Political Agency of Women". *American Economic Journal: Economic Policy*, v. 6, n. 2, p. 164–197, 2014. Citado 5 vezes nas páginas 17, 42, 45, 100, and 144.

BHALOTRA, S.; CLOTS-FIGUERAS, I.; IYER, L. Pathbreakers? women's electoral success and future political participation. *The Economic Journal*, Oxford University Press Oxford, UK, v. 128, n. 613, p. 1844–1878, 2018. Citado 3 vezes nas páginas 18, 42, and 99.

BOAS, T. C.; HIDALGO, F. D.; RICHARDSON, N. P. The spoils of victory: campaign donations and government contracts in brazil. *The Journal of Politics*, Cambridge University Press New York, USA, v. 76, n. 2, p. 415–429, 2014. Citado 2 vezes nas páginas 100 and 105.

BROLLO, F.; TROIANO, U. What happens when a woman wins an election? evidence from close races in brazil. *Journal of Development Economics*, Elsevier, v. 122, p. 28–45, 2016. Citado 4 vezes nas páginas 17, 42, 100, and 144.

CALONICO, S.; CATTANEO, M. D.; TITIUNIK, R. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, Wiley Online Library, v. 82, n. 6, p. 2295–2326, 2014. Citado 23 vezes nas páginas 33, 40, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 78, 79, 80, 81, 93, 94, and 95.

CALONICO, S.; CATTANEO, M. D.; TITIUNIK, R. rdrobust: An r package for robust nonparametric inference in regression-discontinuity designs. *R J.*, Citeseer, v. 7, n. 1, p. 38, 2015. Citado na página 33.

CAMPBELL, D. E.; WOLBRECHT, C. See jane run: Women politicians as role models for adolescents. *The Journal of Politics*, Cambridge University Press New York, USA, v. 68, n. 2, p. 233–247, 2006. Citado na página 18.

CASAS-ARCE, P.; SAIZ, A. *"Women and Power: Unwilling, Ineffective, or Held Back?"*. [S.l.], 2011. Citado na página 42.

CHATTOPADHYAY, R.; DUFLO, E. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India". *Econometrica*, v. 72, n. 5, p. 1409–1443, 2004. Citado 4 vezes nas páginas 17, 21, 45, and 100.

CLARKE, D.; ROMANO, J. P.; WOLF, M. The romano–wolf multiple-hypothesis correction in stata. *The Stata Journal*, SAGE Publications Sage CA: Los Angeles, CA, v. 20, n. 4, p. 812–843, 2020. Citado 2 vezes nas páginas 36 and 112.

CLOTS-FIGUERAS, I. "Are Female Leaders Good for Education? Evidence from India". *American Economic Journal: Applied Economics*, v. 4, n. 1, p. 212–244, 2012. Citado na página 45.

CORREA, G.; MADEIRA, R. A. *The Size of Local Legislatures and Women's Political Representation: Evidence from Brazil*. [S.l.], 2014. Citado na página 18.

COX, G. Electoral rules and electoral coordination. *Annual Review of Political Science*, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 2, n. 1, p. 145–161, 1999. Citado na página 20.

DOWNS, A. An economic theory of political action in a democracy. *Journal of Political Economy*, v. 65, n. 2, p. 135–150, 1957. Citado 2 vezes nas páginas 99 and 144.

DUFLO, E. "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa". *The World Bank Economic Review*, v. 17, n. 1, p. 1–25, 2003. Citado 2 vezes nas páginas 17 and 42.

EGGERS, A. C. et al. On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science*, Wiley Online Library, v. 59, n. 1, p. 259–274, 2015. Citado na página 105.

EGGERS, A. C. et al. Regression discontinuity designs based on population thresholds: Pitfalls and solutions. *American Journal of Political Science*, Wiley Online Library, v. 62, n. 1, p. 210–229, 2018. Citado na página 34.

FISHER, R. A. et al. *The design of experiments*. Edinburgh: Oliver & Boyd, 1937. (2nd Ed). Citado na página 112.

HANSEN, B. E. *Econometrics*. University of Wisconsin: Princeton University Press, forthcoming, 2021. Citado na página 33.

IMBENS, G. W.; LEMIEUX, T. "Regression Discontinuity Designs: A Guide to Practice". *Journal of Econometrics*, v. 142, n. 2, p. 615–635, 2008. Citado na página 31.

IMBENS, G. W.; RUBIN, D. B. *Causal inference in statistics, social, and biomedical sciences*. [S.l.]: Cambridge University Press, 2015. Citado na página 118.

LEE, D. S. Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, Elsevier, v. 142, n. 2, p. 675–697, 2008. Citado na página 100.

LEE, D. S.; LEMIEUX, T. "Regression Discontinuity Designs in Economics". *Journal of Economic Literature*, v. 48, n. 2, p. 281–355, 2010. Citado 15 vezes nas páginas 31, 32, 43, 46, 48, 50, 70, 71, 74, 75, 85, 87, 89, 91, and 105.

LEE, D. S.; MORETTI, E.; BUTLER, M. J. Do voters affect or elect policies? Evidence from the U.S. House. *The Quarterly Journal of Economics*, v. 119, n. 3, p. 807–859, 2004. Citado na página 100.

LOTT, J.; KENNY, L. "Did Women's Suffrage Change the Size and Scope of Government?". *Journal of Political Economy*, v. 107, n. 6, p. 1163–1198, 1999. Citado 2 vezes nas páginas 17 and 45.

MASON, A. D.; KING, E. M. *"Engendering Development through Gender Equality in Rights, Resources, and Voice"*. [S.l.], 2001. 1–388 p. Citado 2 vezes nas páginas 17 and 42.

McCrary, J. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test". *Journal of Econometrics*, v. 142, n. 2, p. 698–714, 2008. Citado 2 vezes nas páginas 34 and 35.

MCKENZIE, D. *Tools of the Trade: a joint test of orthogonality when testing for balance.* [S.l.], 2015. Citado na página 118.

MILLER, G. "Women's Suffrage, Political Responsiveness, and Child Survival in American History". *The Quarterly Journal of Economics*, v. 123, n. 3, p. 1287–1327, 2008. Citado na página 42.

NORRIS, P. The impact of electoral reform on women's representation. *Acta política*, Springer, v. 41, n. 2, p. 197–213, 2006. Citado na página 18.

OAKES, A.; ALMQUIST, E. Women in national legislatures. *Population Research and Policy Review*, Springer, v. 12, n. 1, p. 71–81, 1993. Citado na página 18.

OSBORNE, M. J.; SLIVINSKI, A. A model of political competition with citizen-candidates. *The Quarterly Journal of Economics*, v. 111, n. 1, p. 65–96, 1996. Citado na página 17.

PANDE, R. "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India". *The American Economic Review*, v. 93, n. 4, p. 1132–1151, 2003. Citado 2 vezes nas páginas 17 and 100.

PETTERSSON-LIDBOM, P. "Does the Size of the Legislature Affect the Size of Government? Evidence from Two Natural Experiments". *Journal of Public Economics*, v. 96, n. 3, p. 269–278, 2012. Citado na página 82.

ROMANO, J. P.; WOLF, M. Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, Taylor & Francis, v. 100, n. 469, p. 94–108, 2005. Citado 2 vezes nas páginas 36 and 112.

ROMANO, J. P.; WOLF, M. Stepwise multiple testing as formalized data snooping. *Econometrica*, Wiley Online Library, v. 73, n. 4, p. 1237–1282, 2005. Citado 2 vezes nas páginas 36 and 112.

ROMANO, J. P.; WOLF, M. Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, Elsevier, v. 113, p. 38–40, 2016. Citado 2 vezes nas páginas 36 and 112.

RUBIN, D. B. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies". *Journal of Educational Psychology*, v. 66, n. 5, p. 688–701, 1974. Citado na página 31.

SVALERYD, H. "Women's Representation and Public Spending". *European Journal of Political Economy*, v. 25, n. 2, p. 186–198, 2009. Citado 3 vezes nas páginas 17, 42, and 45.

THOMAS, S.; WELCH, S. The impact of gender on activities and priorities of state legislators. *Western Political Quarterly*, Sage Publications Sage CA: Thousand Oaks, CA, v. 44, n. 2, p. 445–456, 1991. Citado 2 vezes nas páginas 17 and 45.

TREMBLAY, M. Democracy, representation, and women: A comparative analysis. *Democratization*, Taylor & Francis, v. 14, n. 4, p. 533–553, 2007. Citado 2 vezes nas páginas 18 and 99.

VIGNA, S. D.; FERRARA, E. L. “Detecting Illegal Arms Trade". *American Economic Journal: Economic Policy*, v. 2, n. 4, p. 26–57, 2010. Citado na página 53.

YOUNG, A. Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The Quarterly Journal of Economics*, Oxford University Press, v. 134, n. 2, p. 557–598, 2019. Citado na página 112.