# Appendix (online only)

## Table of Contents

# A Mathematical Appendix

First, we rewrite the objective functions of the parties as follows

$$V_L(\mathbf{x}_L,\mathbf{x}_R) = \big(1-F(\mathbf{x}_L,\mathbf{x}_R)\big)\Big(w\big(\varphi(\mathbf{x}_R)-\varphi(\mathbf{x}_L)\big)+(1-w)\Big)-(1-c_L)\kappa+w\varphi(\mathbf{x}_R) \quad (1)$$

$$V_R(\mathbf{x}_L,\mathbf{x}_R) = F(\mathbf{x}_L,\mathbf{x}_R)\Big(w\big(\varphi(\mathbf{x}_R)-\varphi(\mathbf{x}_L)\big)+(1-w)\Big)-c_R\kappa+w\varphi(\mathbf{x}_L). \quad (2)$$

The probability that $R$ wins the election is given by

$$F(\mathbf{x}_L,\mathbf{x}_R) = \max\left\{\frac{\psi-(1-2q)\big(\varphi(\mathbf{x}_R)-\varphi(\mathbf{x}_L)\big)}{2\psi},0\right\}. \quad (3)$$

We first pin down party $L$'s strategy in equilibrium in the following Lemma.

**Lemma A.1.** *Party $L$ always announces its ideal point in equilibrium and never nominates a non-elite candidate.*

*Proof.* Fix $\mathbf{x}_R$. Assuming $c_L = 1$, any interior maximiser, $x_L^*$ of the problem in (1) solves the following equation

$$\frac{\partial(1-F(x_L,1,\mathbf{x}_R;q))}{\partial x_L}\Big(w\big(\varphi(\mathbf{x}_R)-\varphi(\mathbf{x}_L)\big)+(1-w)\Big)-w(1-F(x_L,1,\mathbf{x}_R;q)) = 0. \quad (4)$$

At any $x_L$ such that $F(x_L,1,\mathbf{x}_R;q) < 1$ both terms in (4) are negative because $(1-2q) > 0$, as $q < \frac{1}{2}$. If $F(x_L,1,\mathbf{x}_R;q) = 1$, then $L$ can move to the Left towards a more preferred policy. Thus, there is no interior solution to the maximization problem in (1), and $L$ best-response to any policy choice by $R$ is to announce its ideal point. Since it does not need to recruit a non-elite candidate to credibly commit to its ideal point, $L$ recruits an elite candidate because of the lower cost. □

Having now pinned $L$'s strategy, we turn to party $R$. First, note that the optimal policy announcement is different depending on the choice of candidate. If $R$ nominates an elite candidate, it simply chooses a policy position knowing well that it is constrained by its policy-motivation regarding its implementation. If, on the other hand, $R$ nominates a non-elite candidate, then $R$ chooses a policy position that it is fully committed to. We first characterize these policy positions in the following Lemma.

**Lemma A.2.** *If $c_R = 1$, then $R$'s optimal announcement is*

$$x_R^* = \max\left\{\min\left\{\frac{w\psi-(1-w)(1-2q)}{2w(1-2q)},1\right\},0\right\}. \quad (5)$$

*If $c_R = 0$, then $R$'s optimal announcement is*

$$\tilde{x}_R = \max\left\{\min\left\{\frac{w\psi-(1-2q)(w^2+(1-w)^2)}{2w(1-w)(1-2q)},1\right\},0\right\}. \quad (6)$$

*Proof.* Fix $c_R = 1$. Then, $R$'s problem reduces to

$$\begin{aligned} \max_{x_R} \quad & \left(\frac{\psi-(1-2q)x_R}{2\psi}\right)\big(wx_R+(1-w)\big)-\kappa \\ \text{s.t.} \quad & 0 \le x_R \le 1. \end{aligned} \tag{7}$$

The associated first order condition for an interior solution is

$$\frac{-(1-2q)}{2\psi}\big(wx_R+(1-w)\big)+w\left(\frac{\psi-(1-2q)x_R}{2\psi}\right) \quad = \quad 0,$$

which simplifies to

$$x_R \quad = \quad \frac{w\psi-(1-w)(1-2q)}{2w(1-2q)}.$$

When $w\psi > (1-w)(1-2q)$ and $w\psi < (1+w)(1-2q)$, the RHS above is stricly between 0 and 1 and characterizes $R$'s optimal announcement. If one of these inequalities does not hold, then one of $R$'s constraints binds. Thus, $R$'s optimal announcement is given by

$$x_R^* \quad = \quad \max\left\{\min\left\{\frac{w\psi-(1-w)(1-2q)}{2w(1-2q)},1\right\},0\right\}.$$

Now let $c_R = 0$. $R$'s problem reduces to

$$\begin{aligned} \max_{x_R} \quad & \left(\frac{\psi-(1-2q)(w+(1-w)x_R)}{2\psi}\right)\big(wx_R+(1-w)\big)-\kappa \\ \text{s.t.} \quad & 0 \le x_R \le 1. \end{aligned} \tag{8}$$

The associated first order condition for an interior solution is

$$\frac{-(1-2q)(1-w)}{2\psi}\big(wx_R+(1-w)\big)+w\left(\frac{\psi-(1-2q)(w+(1-w)x_R}{2\psi}\right) \quad = \quad 0,$$

which simplifies to

$$x_R \quad = \quad \frac{w\psi-(w^2+(1-w)^2)(1-2q)}{2w(1-w)(1-2q)}.$$

When $w\psi > (w^2+(1-w)^2)(1-2q)$ and $w\psi < (1-2q)$, the RHS above is stricly between 0 and 1 and characterizes $R$'s optimal announcement. If one of these inequalities does not hold, then one of $R$'s constraints binds. Thus, $R$'s optimal announcement is given by

$$\tilde{x}_R \quad = \quad \max\left\{\min\left\{\frac{w\psi-(1-2q)(w^2+(1-w)^2)}{2w(1-w)(1-2q)},1\right\},0\right\},$$

completing the proof. □

We now note that the objective problem in (2) is a two-part problem. Party $R$ can pay the cost to nominate a non-elite candidate and fully commit to $x_R^*$, or it can choose to go with an elite candidate forgoing full credibility and announce $\tilde{x}_R$. Thus, party $R$ nominates a non-elite candidate if and only if the following inequality is met:

$$F\big((0,1),(x_R^*,1)\big)\big(wx_R^*+(1-w)\big)-F\big((0,1),(\tilde{x}_R,0)\big)\big(w\tilde{x}_R+(1-w)\big) \quad \geq \quad \kappa. \tag{9}$$

The above inequality is key in proving the two Propositions in the main text.

## Proof of Proposition 1

*Proof.* Let $w=0$. Then, $x_R^*=\tilde{x}_R=0$ for all $q\in\big(0,\frac{1}{2}\big)$. Clearly then $F\big((0,1),(x_R^*,1)\big)=F\big((0,1),(\tilde{x}_R,0)\big)$, implying that the inequality in (9) cannot be met at any $q\in\big(0,\frac{1}{2}\big)$. Thus, when $w=0$, party $R$ never nominates a non-elite candidate. Since $x_R^*$ and $\tilde{x}_R$ are continuous in $w$, this observation holds for small $w$ as well. Indeed, consider $q'$ arbitrarily close to 0 and $w=1$. Then, $\varphi(\tilde{x}_R,0)=1$ and $x_R^*=\frac{\psi}{2(1-2q')}>\frac{\psi}{2}$, implying that $F\big((0,1),(x_R^*,1)\big)=1/4$ while $F\big((0,1),(\tilde{x}_R,0)\big)=0$. As a result, $R$'s payoff from $(\tilde{x}_R,0)$ is 0, while the payoff from $(x_R^*,1)$ is $\frac{\psi}{8(1-2q')}-\kappa>0$. Thus, the inequality in (9) is met for some $q\in\big(0,\frac{1}{2}\big)$ when $w=1$. Note that all components of the inequality in (9) are continuous in $w$. Also note that when $w=0$, there does not exist a $q\in\big(0,\frac{1}{2}\big)$ such that (9) holds, and that there exists a $q\in\big(0,\frac{1}{2}\big)$ such that (9) holds when $w=1$. All these facts imply that there exists a $\bar{w}\in(0,1)$ such that for all $w\leq\bar{w}$, equation (9) does not hold for all $q\in\big(0,\frac{1}{2}\big)$. □

## Proof of Proposition 2

The proof follows in three steps. First, we define $\bar{q}$ when $\tilde{x}_R=1$ and $x_R^*\in(0,1)$ such that (9) holds with equality. Then, we show that when $w=1$ this is the only solution to that equation. Finally, we use continuity of (9) to argue that there is a unique $\bar{q}$ where (9) holds with equality when $w$ is close enough to 1.

Let $\tilde{x}_R=1$ and $x_R^*\in(0,1)$. Set (9) to hold with equality and solve for $q$. There are two candidate solutions:

$$\bar{q} \quad = \quad \frac{(1+w)(3w+1)-2\psi w^2-4\psi w^2\kappa^2\pm w\psi\sqrt{(w-1)^2+16w^2\kappa(1+\kappa)}}{2(1+w)(3w+1)}.$$

Rewriting one of the solutions we get the following equation

$$w\psi\sqrt{(w-1)^2+16w^2\kappa(1+\kappa)}-4\psi w^2\kappa \quad = \quad 2\psi w^2-(1+w)(3w+1)(1-2q).$$

Since $x_R^* < 1$, we have that $\psi w < (1+w)(1-2\bar{q})$, and clearly since $w \in [0,1]$, we have that $2w < 3w+1$. As a result, the RHS is negative. Looking at the LHS,

$$\begin{aligned} w\psi\sqrt{(w-1)^2 + 16w^2\kappa(1+\kappa)} - 4\psi w^2\kappa &\geq w\psi\sqrt{16w^2\kappa(1+\kappa)} - 4\psi w^2 \\ &= 4\psi w^2\sqrt{\kappa(1+\kappa)} - 4\psi w^2\kappa \\ &> 0, \end{aligned}$$

which is a contradiction. Thus, the unique solution, if it exists, is

$$\bar{q}(w) = \frac{(1+w)(3w+1) - 2\psi w^2 - 4\psi w^2\kappa^2 - w\psi\sqrt{(w-1)^2 + 16w^2\kappa(1+\kappa)}}{2(1+w)(3w+1)}.$$

Now let $w = 1$. Then, $\varphi(x,0) = 1$ for all $x \in [0,1]$. That is, voters believe that $R$ implements its ideal point for sure if its candidate is elite. Furthermore, since $\kappa < \frac{\psi}{8}$, when $q = 0$ we have that $F\big((0,1),(x_R^*,1)\big)\big(x_R^*\big) - \kappa = \frac{\psi}{8} - \kappa > 0$, while $F\big((0,1),(\tilde{x}_R,1)\big) = 0$. And when $q = \frac{1}{2}$, we have that $F\big((0,1),(x_R^*,1)\big)\big(x_R^*\big) - \kappa < F\big((0,1),(\tilde{x}_R,0)\big)\big(\tilde{x}_R\big)$ because $x_R^* = \tilde{x}_R = 1$ when $q = \frac{1}{2}$. Since the LHS of (9) is continuous in $q$, there exists at least one $q$ where the equation (9) holds with equality. By the first step, this is given uniquely by $\bar{q}(1)$.

As the equation (9) is continuous in $w$, there exists a $\bar{w} < 1$ such that for all $w \geq \bar{w}$, the unique solution to the equality in (9) is given by $\bar{q}(w)$.

# B The Costs and Benefits of Less Educated Mayors

Our model simplifies candidate nomination to a binary decision and assumes that, all-else-equal, candidates that are descriptively similar to the poor are relatively more costly to nominate for both parties. In order to focus on the indirect, policy-driven relative benefits of descriptive representation, we also abstract away from modeling the direct benefits of descriptive representation. In this section, we discuss the assumption that less-educated mayors are more costly to nominate than higher educated ones, as well as voter preferences for descriptive representation.

## Political costs of less-educated mayors

We first discuss in detail the political costs of less-educated candidates. The determinants of the relative cost (or benefit) of nominating candidates of a specific background can be broadly divided into two categories: (i) supply-side and (ii) demand-side. The former category encompasses costs associated with the available supply of candidates, while the latter category comprises all (pre or post-election) costs associated with fielding a candidate of a particular type. These costs are always context specific. We show that the evidence from Brazil suggests that the nomination of less-educated candidates is more costly on balance in terms of both categories.

### *Supply-side factors*

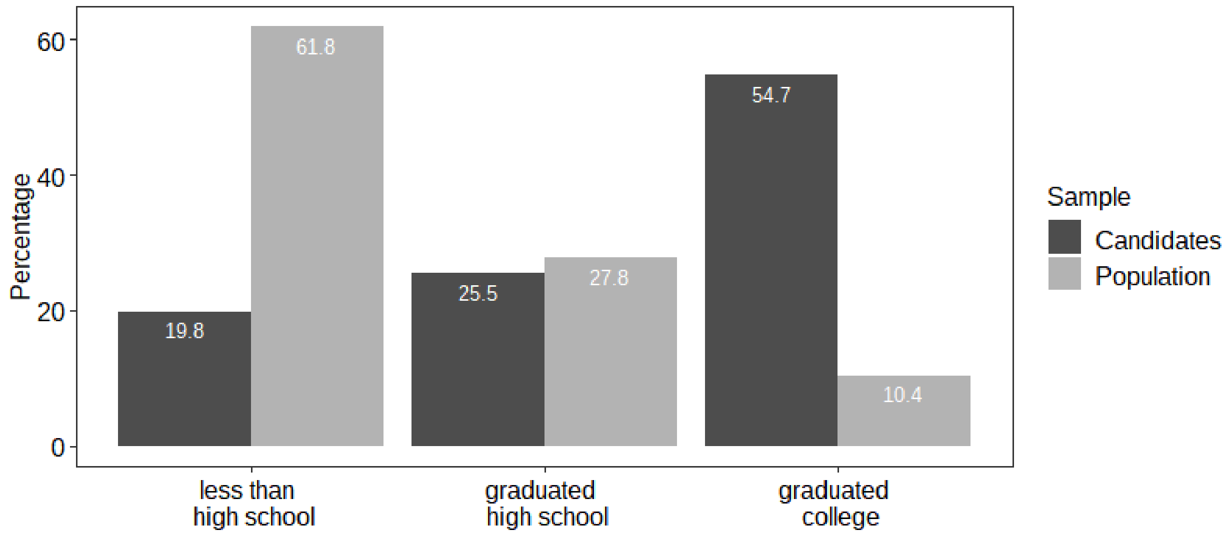
First, note that candidate selection does not happen in a vacuum and crucially depends on the candidate pool. Self-selection into politics is a well-documented phenomenon and shapes the candidate pools available to parties. For example, Dal Bó et al. (2017) examine the pattern of political selection for Swedish politicians. They find that politicians are on average better educated and wealthier than their constituents. They interpret their findings as positive selection into politics by individuals with higher human capital (that are evidently descriptively closer to the more affluent sections of society). We find similar patterns in Brazil, as is demonstrated in Figure B.1: while more than 60% of Brazilian voters did not graduate from high school, less than 20% of mayoral candidates did not; and while 10% of Brazilian voters graduated from university, almost 55% of candidates did.[34] This suggests that in Brazil, like in Sweden, there is a form of self-selection into politics by individuals that have on average higher human capital than voters. Thus, on the supply side, there is suggestive evidence that parties face higher costs to nominate less educated politicians in Brazil, simply because there are fewer such available candidates.

### *Demand-side factors*

Evidently, we cannot observe or list all the potential dimensions in which the education of politicians might hurt (or benefit) their parties. However, in this section we discuss some dimensions of performance

[34] This pattern is similar across both high and low poverty municipalities.

Figure B.1: Education of Politicians and Voters in Brazil




Data on politicians comes from the top 2 mayoral candidates in the 2004-2016 elections. Data for the general population considers only adults of 25 years or more, and comes from the 2010 census.

that are particularly important in the context of Brazil. We regress several policy and electoral outcomes on the education of mayors in 2004-2016 in a panel analysis.[35]

First, there may be a potential cost from nominating candidates that are less educated if they are of lower valence. The correlation between education and quality of politicians has been central to the burgeoning literature on political selection, where most work has directly associated the education of politicians with both their administrative abilities and subsequent performance in office – see a review in Dal Bó and Finan (2018). In the case of Brazil, Table B.1 shows that the presence of less educated mayors is associated with costs in at least two very relevant dimensions: administrative performance and ability to raise budget funds. Less educated mayors are significantly less likely to receive discretionary funds for special investment projects (*convênios*) from the federal government (columns 3-4). Event though the ultimate decision on the destination of these resources in made by the central government, mayors have the ability to negotiate and lobby for these transfers. Furthermore, school enrollment is lower under less educated mayors (columns 5-6), and the coverage of the public health sector is also lower (columns 7-8). Finally, there is no differential effect of performance of lower educated mayors by party ideology.

If on average uneducated candidates are of lower quality, all-else-equal, parties would limit nominating them despite potential benefits from descriptive representation. This is especially true in a political context where parties *are* punished for the weak administrative performance of their mayors (Feierherd, 2020; Ferraz and Finan, 2008; Klašnja and Titiunik, 2017). Note that we do not specifically model this electoral cost (a lower valence for lower educated candidates) but the results from including this va-

[35] We include fixed effects for both period and municipality, and also control for the education level of the runner-up. We also interact education with a dummy that indicates whether the mayor is from a Left-wing party

Table B.1: The education of mayors and some measures of performance

| Dep. Variable: | Party Votes | | Extra funds | | School | | Health | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Education | 0.451* | 0.395* | 0.015* | 0.014* | 0.001* | 0.001† | 0.016* | 0.014* |
| | (0.108) | (0.121) | (0.004) | (0.004) | (0.001) | (0.001) | (0.004) | (0.005) |
| Left-wing | | -1.971 | | 0.057 | | -0.007 | | 0.027 |
| | | (1.526) | | (0.053) | | (0.009) | | (0.058) |
| Educ. x L-wing | | 0.235 | | 0.003 | | 0.000 | | 0.008 |
| | | (0.222) | | (0.008) | | (0.001) | | (0.008) |
| Observations | 16995 | 16995 | 16995 | 16995 | 16995 | 16995 | 16995 | 16995 |

†p<0.1, **p<0.05. Standard errors are clustered by municipality (parenthesis). All regressions include fixed effects for municipality and election. The variables are defined as follows: Party Votes: Votes for the mayor's party in the midterm congressional elections. Extra funds: Discretionary resources for special investment projects negotiated with the federal government. School: School enrollment as a share of the population (measure in the last available school census of the mayoral tenure). Health: Number of public health teams per 4000 inhabitants, which is the target level of health coverage in the public sector (at the end of the mayoral tenure).

lence difference between low and high educated candidates would be qualitatively similar to the results presented in text.

Second, mayors in Brazil are important vote brokers for their parties in subsequent federal elections (Brollo and Nannicini, 2012; Novaes, 2018). We examine whether the mayor's education level influences her ability to mobilize votes for her party in congressional elections. We highlight that public funding for parties in the Brazilian electoral system is split according to the number of congressional seats. Thus, this analysis provide a clear and direct measure of the mayor's overall contribution to her party's national strength. Columns (1-2) clearly show that parties that elect less educated mayors lose valuable votes for their house candidates. Interestingly, this is the case for both Left and Right-wing parties, which is in line with the assumption in our theory. We highlight that, as suggested by (Feierherd, 2020), this brokerage deficit by uneducated candidates might be in itself a consequence of their poor administrative performance discussed above. Furthermore, recent work also indicates that less educated mayors in Brazil are more likely to switch parties after being elected (Hott and Sakurai, 2020), suggesting that higher educated mayors are also safer candidates from a party organizational perspective. Therefore, the evidence broadly suggests that even on the demand side the cost of nominating lower educated candidates in Brazil is relatively higher.

*Alternative cost structures*

If the cost of nominating less educated candidates is context specific, what are implications of alternative cost structures? Evidently, different sets of assumptions result in differing comparative static predictions. Here we discuss informally what our model would predict with the imposition of two alternative cost structures: (i) the same cost for nominating elite and non-elite candidates and (ii) a higher cost for nominating more educated candidates.

First, consider a cost framework where there is no cost differential in nominating candidates of a particular socio-economic class. This can be modeled simply by setting $\kappa = 0$. The Left is now completely indifferent between the two types of candidates. On the other hand, if programmatic brands are salient for both parties and voters (i.e. when $w$ is high), the Right *always* nominates lower educated candidates, even in relatively wealthy municipalities. With this cost framework, the model predicts that we should be equally likely to see an education gap in high poverty and low poverty municipalities. This prediction is clearly at odds with the empirical pattern we uncover in the data from Brazil.

Second, consider the case when it is relatively more costly to nominate a higher educated candidate, i.e. set $\kappa < 0$. As before, the Left has an overwhelming advantage with respect to policy competition, and so its policy announcement is always in line with its ideals. However, since lower educated candidates are less costly to nominate, the Left always nominates a non-elite candidate. The Right benefits from nominating lower educated candidates, since they lend legitimacy to its deviations from its ideal point. Since they are less costly than higher educated candidates to nominate, it is strictly beneficial for the Right to nominate them. Thus, the Right also always prefers to nominate lower educated non-elite candidates. This particular cost framework would predict that there is no education gap between candidates from the Left and the Right at *all* levels of poverty. However, the lack of an education gap is driven by both parties nominating lower educated candidates. Again, these predictions are not borne out in data from Brazil.

## Direct benefits of descriptive representation

Finally, voters in our model do not have any direct preferences for descriptively representative candidates, even though the literature finds that they are very likely to possess them. We choose to abstract away from these considerations in order to isolate the indirect, policy-related effects of descriptive representation on party strategies, and to have a cleaner exposition of our results. In this section we discuss the consequences of including these direct preferences in our framework. In lieu of presenting a fully solved version of the model, we discuss the logic in an informal manner.

Consider the most interesting case that generates the main prediction in our model when parties are fully policy-motivated (i.e. $w = 1$). Now suppose that poor voters receive a positive bump $d > 0$ to their utility when parties nominate a non-elite candidate, thereby directly improving the electoral prospects of both parties the *same* way. Party $R$'s incentives to nominate a descriptively representative candidate *always remain greater* than those of $L$. In addition to the direct electoral benefits from convincing more poor voters, the choice of a non-elite candidate *also* provides credibility to *R*'s deviations from its ideal point. Clearly, the effectiveness of descriptive representation as an electoral tool is dependent on the proportion of poor in the region. When there are more poor voters, the potential swing in vote share due to the nomination of a descriptively representative candidate is higher for both parties.

Thus, if $d$ is small enough relative to $\kappa$,[36] then the above facts imply the following for high and low poverty areas. When poverty is very high, *only* $R$ prefers to nominate non-elite candidates. While the direct benefits of nominating a descriptively representative candidate do not increase party payoffs enough relative to the cost in very poor areas, $R$ still indirectly benefits by credibly reducing programmatic differentiation with $L$. In less poor areas, the effectiveness of direct descriptive representation as an electoral tool is too low for both parties, and they both nominate elite candidates. Note that these results are qualitatively identical to those in Proposition 2 in the paper.

If $d$ is relatively high, both parties would be more likely to nominate non-elite candidates in all areas. In the context of this framework, our empirical results then suggest that the direct benefits of descriptive representation in Brazil – although likely present – are not exceedingly high. Our empirical results highlight that the Left is very likely to select highly educated candidates in both high and low poverty areas, and the Right also does so in low poverty areas.

[36] That is, fixing the cost of a non-elite candidate, poor voter preferences for a descriptively representative candidate are relatively small in magnitude. We stress here that it is *not* necessary that $d < \kappa$. It suffices that for each $\kappa > 0$, we have $d < \alpha(\kappa)$, where $\alpha(\kappa)$ is increasing in $\kappa$ and could potentially be greater than $\kappa$.

# C Candidate's Campaign Proposals

Table C.1: Candidate's proposals: RDD

| Dependent Variable: | Pro-poor proposals I (gap) | | | Pro-poor proposals II (gap) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| High Poverty | 0.305 | 0.120 | 0.018 | 0.304 | 0.103 | -0.001 |
| | (0.291) | (0.262) | (0.246) | (0.298) | (0.270) | (0.252) |
| Low Poverty | -0.516* | -0.430* | -0.366$^{\dagger}$ | -0.505* | -0.449* | -0.386* |
| | (0.227) | (0.213) | (0.199) | (0.211) | (0.198) | (0.185) |
| Bandwidth | 3.52 | 4.70 | 5.87 | 3.61 | 4.81 | 6.01 |
| Observations | 545 | 707 | 856 | 557 | 717 | 874 |

$^{\dagger}$p<0.1, *p<0.05. Standard errors are clustered by municipality (in parenthesis). The estimates represent the difference in outcomes between municipalities with Right and Left-wing mayors for each subsample, at the discontinuity. The coefficients come from the estimation of equation 8. The first outcome variable includes the count of all "pro-poor" words. The second has this count subtracted by the number of "law-and-order" words.

Here we show that our main result is robust to an alternative measure of the candidates' pro-poor policies. The Brazilian Electoral Court (TSE) discloses the written policy proposals of all mayoral candidates in the country since 2012. We collected all the available readable files in the website, and converted each proposal to text.[37] We then created a pro-poor score for each document, which is the number of "pro-poor" words as a percentage of the total word count – "pro-poor" words are the ones related to education, health, sanitation and housing policies.[38]

These scores can be calculated for both candidates, which allows us to estimate the RDD for the gap in pro-poor proposals between the winner and the loser in the election. The estimates resemble our main policy result: in very poor municipalities, the policy proposals of the Left and the Right are indistinguishable. In low poverty areas, the Right is much less likely to campaign on pro-poor issues than the Left.

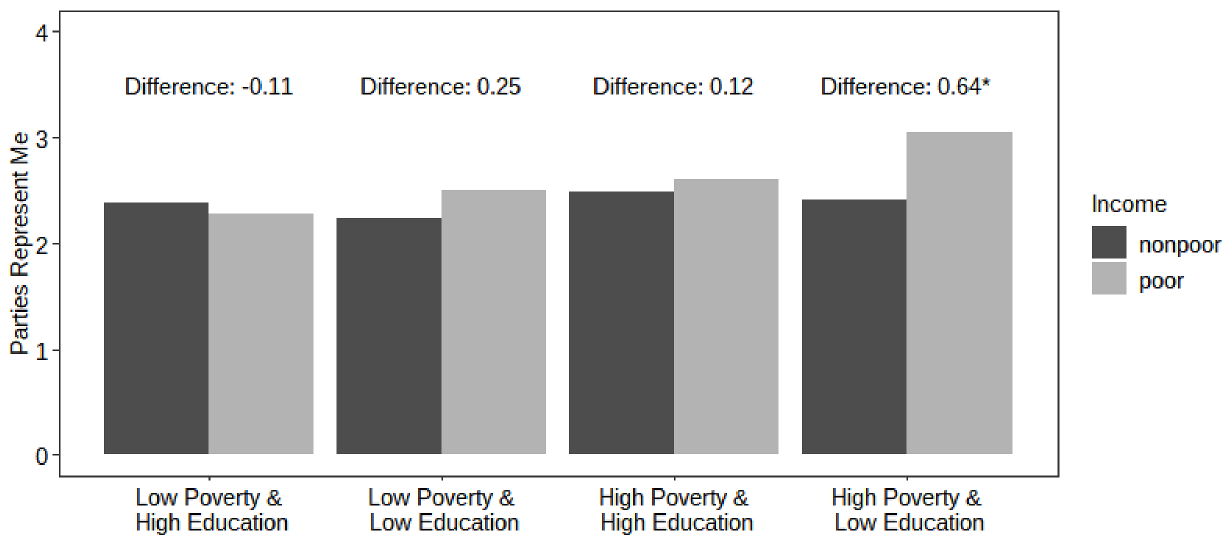
[37]PDF files available at https://divulgacandcontas.tse.jus.br/. Some files were missing, some were not in pdf format, and some were uploaded as low quality scans of paper documents that could not be converted to text. Thus, we recovered data for both the Left and Right-wing candidates for 83% of our sample in 2012-16.

[38]The full list is: education, health, school, pharmacy, teacher, medical doctor, water, sewage, sanitation, house (and housing), poverty, inclusion, vulnerability, misery and hunger. As a slightly different specification we also counted "law-and-order" words, which are typically related to Right-wing policies, and subtracted them from the pro-poor count. These are: police, violence, crime, guard and security.

# D Self-Reported Representation in LAPOP

We provide support for the main findings using survey results from the America's Barometer conducted by the LAPOP.[39] The surveys of 2008 and 2012 contain a question where voters express the extent to which they believe political parties represent them.[40] Proximity is measured in a 1-7 scale, 7 being the highest level, and it is available from 2030 voters in 197 municipalities. We use these questions to show that the correlation between poverty, the politician's education, and the voter's self-reported 'proximity' to political parties is consistent with both the theory and the results presented in this article. Note that this proximity could be both in terms of policy preferences or descriptive representation. Our theory predicts that it is only in high-poverty municipalities, and under a less educated politician, that we should expect poor voters to always feel more represented by mayors from either ideological group, compared to nonpoor voters.[41]

Figure D.1: Self-reported political representation




†p<0.1, *p<0.05. The data include 2030 voters in 197 municipalities. Low education refers to municipalities where the mayor has graduated high school, at most.

We classify municipalities into four groups, by poverty level (low and high), and by the mayor's education (low and high).[42] We also split voters in each municipality into poor and non-poor.[43] Figure D.1

[39] www.lapopsurveys.org. We thank the Latin American Public Opinion Project (LAPOP) and its major supporters (the United States Agency for International Development, the Inter-American Development Bank, and Vanderbilt University) for making the data available.

[40] In the 2008 wave, the question was: *To what extent political parties are close to people like me?* In 2012: *How much do political parties listen to people like you?*

[41] Poor voters might feel less represented by a Rightist mayor that is more educated, or by a Rightist mayor that implements less pro-poor spending (as it is the case in low-poverty municipalities).

[42] Low education is defined as having, at most, a high school degree.

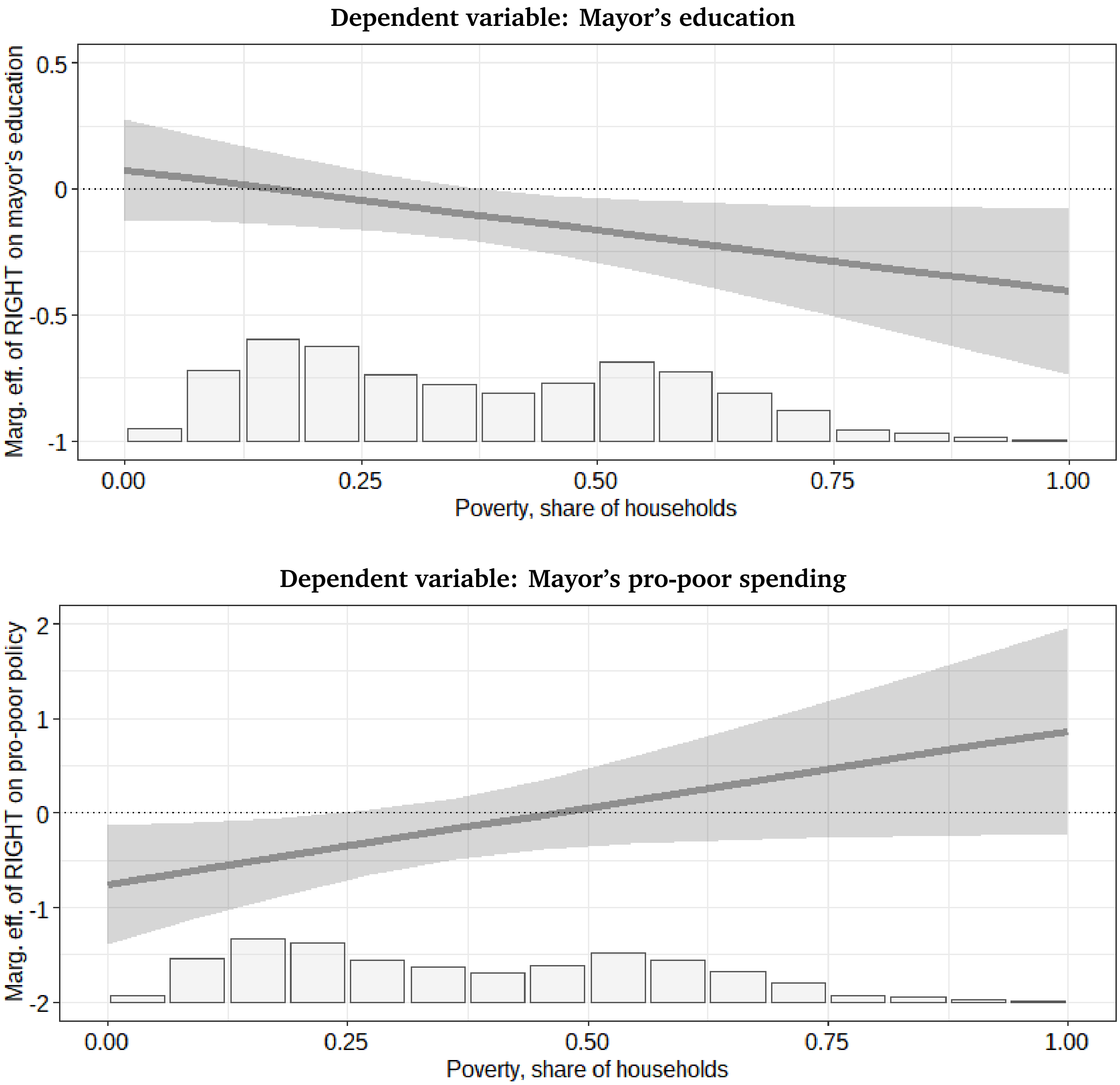
[43] In both waves, income is reported in categories. Poor households are the ones with monthly income below R$380 in 2008,

shows that, for three out of the four groups, poor and nonpoor voters feel equally represented by political parties. It is only in high-poverty areas under a less educated mayor that the poor feels significantly closer to politicians than the nonpoor.

which is the first recorded category, and it is slightly below the national minimum wage level. Adjusted by inflation, this corresponds to R$440 in 2012, so poor voters in 2012 are the ones with income below the closest recorded level (R$410).

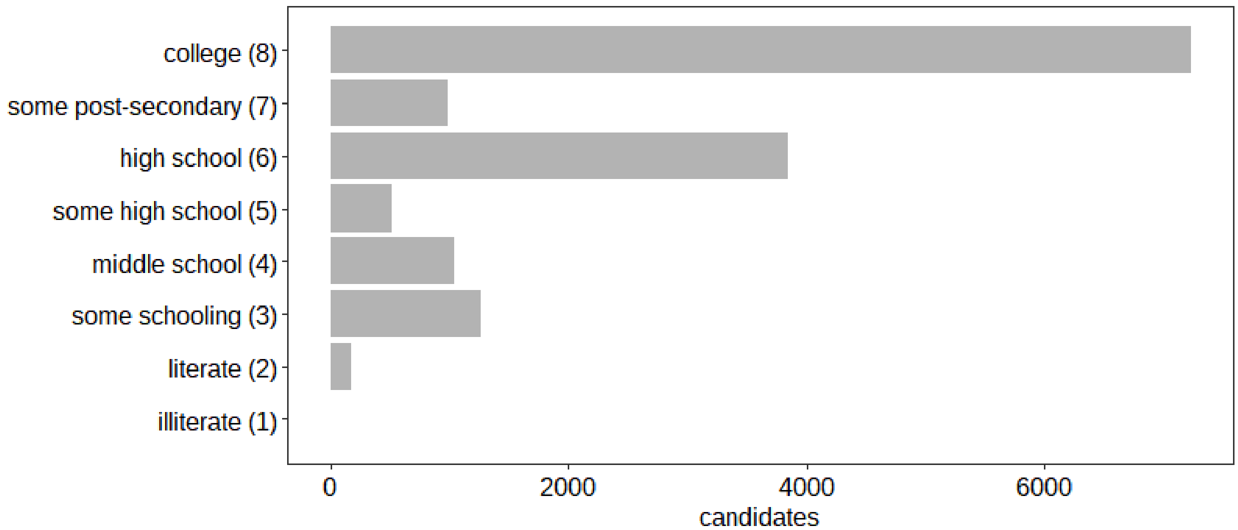
# E Tables and Figures

Figure E.1: Marginal effect of electing a Right-wing mayor, by poverty level



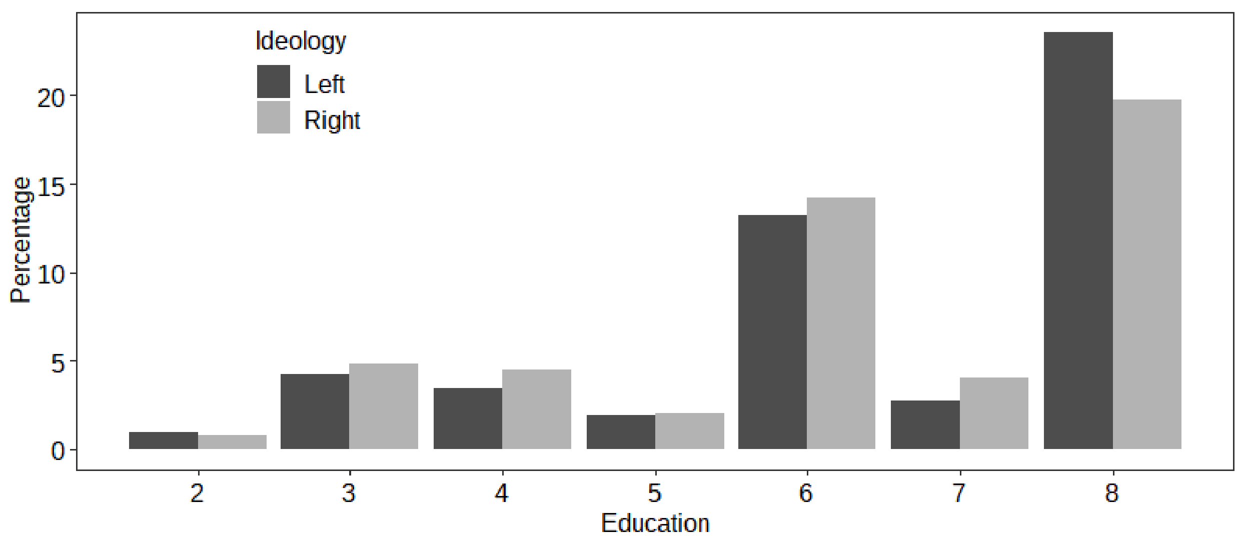

Data from 2004-2016. The lines show how the marginal effect of electing a Right-wing mayor (instead of a left-wing contender) changes as poverty increases in municipalities. The plot is based on a regression that includes fixed-effects for election and municipality, and pre-treatment, time-variant covariates. The coefficients that generate these plots are shown in Table E.9 of this appendix, where the regression is also described. The lines show a linear fit with a 95% confidence interval; and the columns represent the share of the sample at each poverty level. In the case of the policy variable, the dependent variable is coded as the change in the share of budget spent on pro-poor categories between the current and the last mayoral tenures.

Figure E.2: Education scale for mayoral candidates (2004-2016)



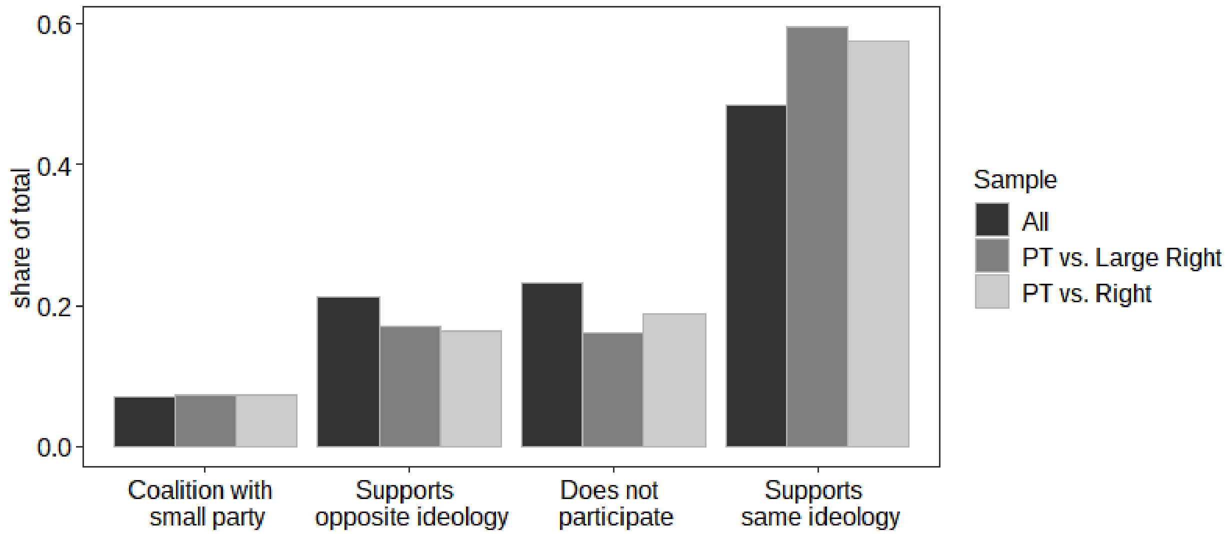

The bars show the number of candidates on each category.

Figure E.3: Education gap in the high poverty sample



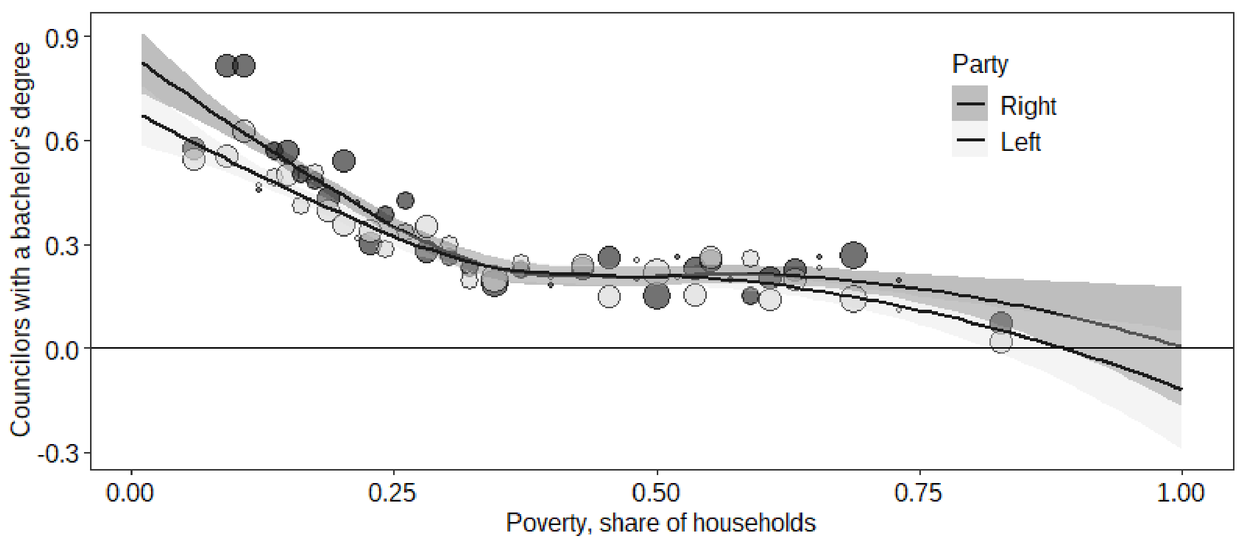

The bars show the percentage of candidates on each category. The description of these categories can be found in Figure E.2 above.

Figure E.4: Coalition choices by the main Brazilian parties



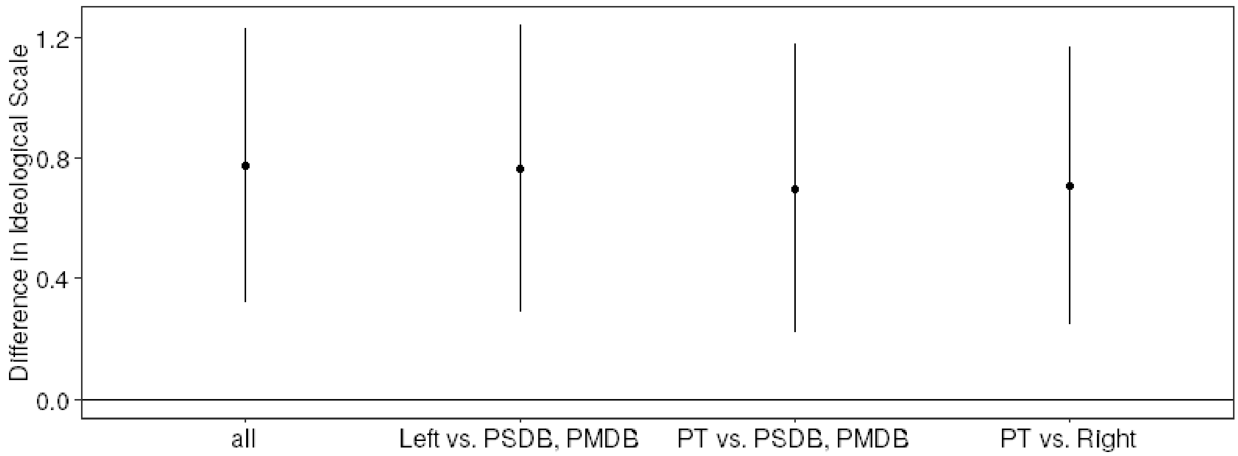

The bars show the share of the total cases. Large Right includes PMDB, PSDB, PFL, PP and PSD.

Figure E.5: Highly educated councilors, by party ideology



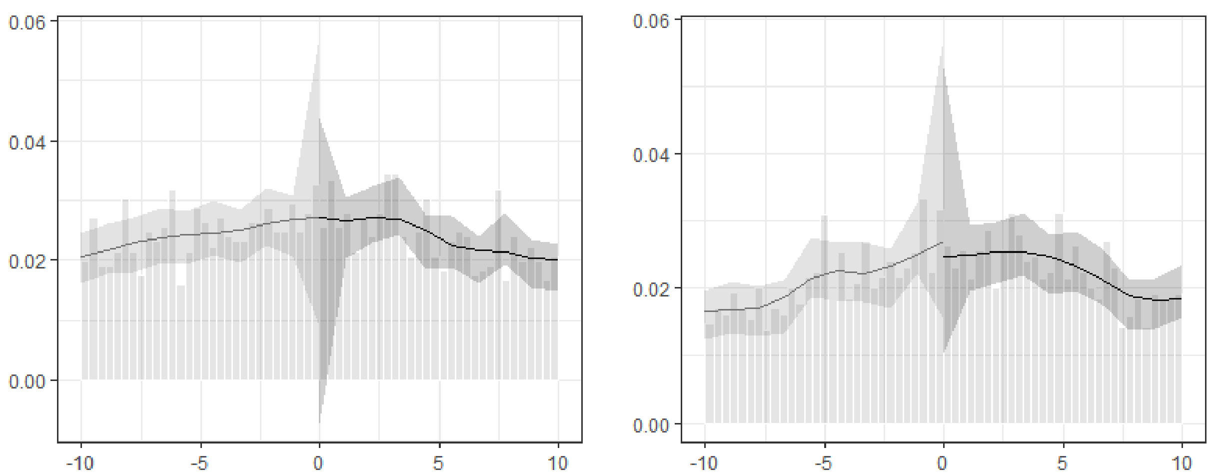

The plot shows the number of councilors with a bachelor's degree, immediately preceding each election in 2004-2016, for the Right and Left-wing parties. For each municipality, the count includes councilors from the top 2 parties in the mayoral election – only considering this paper's sample of races between Right and Left parties. The size of the dots represent the number of observations at each poverty level.

Figure E.6: Correlation between self-identified ideology and party preference in Brazil




The point estimates represent the average difference in ideology of voters that identify with Right vs. Left parties. The self-identified ideological score is on a L-R scale, from 0 to 10. The sample comes from the LAPOP 2010 survey, with 560 respondents when all parties in Figure 2 are included. 95% confidence intervals are shown.

Figure E.7: Density of the running variable




The plot in the left shows the high poverty sample, the plot in the right shows the low poverty sample.

Table E.1: Balance of covariates

| Dependent Variable | Sample: High Poverty | | | Sample: Low Poverty | | |
|---|---|---|---|---|---|---|
| Budget | 0.158 | 0.127 | 0.080 | -0.180 | -0.167 | -0.069 |
| (past tenure) | (0.130) | (0.114) | (0.102) | (0.146) | (0.128) | (0.118) |
| | [3.46] | [4.61] | [5.77] | [3.46] | [4.61] | [5.77] |
| | 0.55 | 0.56 | 0.77 | 0.12 | 0.36 | 0.63 |
| Left for President | 2.949 | 2.776 | 2.502 | 1.754 | 1.901 | 1.489 |
| (last election) | (2.591) | (2.218) | (1.974) | (1.996) | (1.720) | (1.541) |
| | [3.15] | [4.19] | [5.24] | [3.15] | [4.19] | [5.24] |
| | 0.70 | 0.17 | 0.09 | 0.36 | 0.79 | 0.62 |
| Longitude | 0.156 | 0.231 | 0.525 | -1.047$^{\dagger}$ | -0.880$^{\dagger}$ | -0.566 |
| | (1.104) | (0.969) | (0.866) | (0.576) | (0.515) | (0.468) |
| | [3.61] | [4.81] | [6.02] | [3.61] | [4.81] | [6.02] |
| | 0.67 | 0.70 | 0.97 | 1.00 | 0.98 | 0.96 |
| Latitude | 0.789 | 0.744 | 0.457 | -0.203 | 0.101 | 0.207 |
| | (0.996) | (0.853) | (0.762) | (0.685) | (0.597) | (0.540) |
| | [4.29] | [5.72] | [7.15] | [4.29] | [5.72] | [7.15] |
| | 0.54 | 0.27 | 0.92 | 0.47 | 0.96 | 0.94 |
| Semi-arid | 0.090 | 0.058 | 0.049 | -0.008 | -0.010 | -0.009 |
| | (0.078) | (0.068) | (0.061) | (0.007) | (0.007) | (0.006) |
| | [3.68] | [4.91] | [6.14] | [3.68] | [4.91] | [6.14] |
| | 1.00 | 0.97 | 1.00 | 1.00 | 1.00 | 1.00 |
| Share of voters | 0.032 | 0.019 | 0.008 | -0.027 | -0.022 | -0.017 |
| | (0.022) | (0.020) | (0.018) | (0.018) | (0.016) | (0.014) |
| | [3.63] | [4.84] | [6.05] | [3.63] | [4.84] | [6.05] |
| | 0.67 | 0.92 | 0.47 | 0.94 | 0.98 | 0.98 |
| Households (log) | 0.122 | 0.156 | 0.194$^{\dagger}$ | 0.054 | 0.023 | 0.024 |
| (census 2000) | (0.128) | (0.115) | (0.104) | (0.181) | (0.159) | (0.144) |
| | [3.66] | [4.88] | [6.10] | [3.66] | [4.88] | [6.10] |
| | 0.47 | 0.69 | 0.55 | 0.90 | 0.29 | 0.52 |
| Bandwidth rules | 0.75 x op. | optimal | 1.25 x op. | 0.75 x op. | optimal | 1.25 x op. |

$^{\dagger}$p<0.1, *p<0.05. For each variable: standard errors in parenthesis, optimal bandwidth in brackets, and the p-value of a KS-test in the last row. The Table continues in the next page – see all notes at the end.

Table E.1: Balance of covariates (continued)

| Dependent Variable: | Sample: High Poverty | | | Sample: Low Poverty | | |
|---|---|---|---|---|---|---|
| Inequality | 0.007 | 0.006 | 0.005 | -0.001 | 0.001 | 0.002 |
| (census 2000) | (0.010) | (0.009) | (0.008) | (0.011) | (0.010) | (0.009) |
| | [3.34] | [4.46] | [5.57] | [3.34] | [4.46] | [5.57] |
| | 0.98 | 0.98 | 0.99 | 0.99 | 0.95 | 0.83 |
| GDP pc | 0.069 | 0.040 | -0.020 | -0.678 | -0.588 | -0.368 |
| (IBGE 2000) | (0.219) | (0.187) | (0.162) | (0.772) | (0.787) | (0.737) |
| | [3.71] | [4.94] | [6.18] | [3.71] | [4.94] | [6.18] |
| | 0.59 | 0.67 | 0.81 | 0.10 | 0.14 | 0.42 |
| Gender gap | 0.071 | 0.064 | 0.056 | 0.079 | 0.081 | 0.076 |
| | (0.074) | (0.065) | (0.059) | (0.068) | (0.059) | (0.054) |
| | [3.98] | [5.31] | [6.63] | [3.98] | [5.31] | [6.63] |
| | 1.00 | 1.00 | 0.96 | 1.00 | 1.00 | 0.97 |
| Career gap | 0.061 | 0.022 | -0.006 | -0.087 | -0.099 | -0.101 |
| (Health / Education) | (0.073) | (0.062) | (0.055) | (0.080) | (0.069) | (0.063) |
| | [4.53] | [6.05] | [7.56] | [4.53] | [6.05] | [7.56] |
| | 0.14 | 0.36 | 0.49 | 0.95 | 0.82 | 0.78 |
| Career gap | -0.049 | -0.015 | -0.009 | 0.025 | 0.038 | 0.046 |
| (Business) | (0.083) | (0.074) | (0.066) | (0.095) | (0.083) | (0.075) |
| | [3.56] | [4.74] | [5.93] | [3.56] | [4.74] | [5.93] |
| | 0.98 | 0.90 | 0.96 | 0.69 | 0.46 | 0.26 |
| Career gap | -0.029 | -0.005 | 0.004 | -0.002 | -0.014 | -0.019 |
| (Public sector) | (0.055) | (0.050) | (0.046) | (0.061) | (0.052) | (0.046) |
| | [3.68] | [4.90] | [6.13] | [3.68] | [4.90] | [6.13] |
| | 1.00 | 1.00 | 0.94 | 1.00 | 1.00 | 1.00 |
| Career gap | 0.009$^{\dagger}$ | 0.001 | -0.004 | -0.014 | -0.007 | -0.006 |
| (Religious) | (0.006) | (0.006) | (0.006) | (0.015) | (0.012) | (0.010) |
| | [3.80] | [5.07] | [6.33] | [3.80] | [5.07] | [6.33] |
| | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Bandwidth rules | 0.75 x op. | optimal | 1.25 x op. | 0.75 x op. | optimal | 1.25 x op. |

$^{\dagger}$p<0.1, *p<0.05. For each variable: standard errors in parenthesis, optimal bandwidth in brackets, and the p-value of a KS-test in the last row. The Table continues in the next page – see all notes at the end.

Table E.1: Balance of covariates (continued)

| Dependent Variable: | Sample: High Poverty | | | Sample: Low Poverty | | |
|---|---|---|---|---|---|---|
| Competitive coalitions | 0.007 | -0.018 | -0.005 | 0.042 | 0.111 | 0.137 |
| | (0.292) | (0.256) | (0.230) | (0.306) | (0.265) | (0.237) |
| | [4.09] | [5.45] | [6.81] | [4.09] | [5.45] | [6.81] |
| | 0.91 | 0.98 | 0.86 | 0.55 | 1.00 | 1.00 |
| Past mayor's party is | 0.005 | 0.032 | 0.057 | -0.008 | -0.015 | -0.033 |
| Top 2 (current) | (0.076) | (0.066) | (0.059) | (0.074) | (0.065) | (0.059) |
| | [4.07] | [5.43] | [6.78] | [4.07] | [5.43] | [6.78] |
| | 0.27 | 0.22 | 0.94 | 0.66 | 0.90 | 1.00 |
| Past mayor's party is | 0.046 | 0.031 | 0.032 | 0.013 | -0.008 | -0.027 |
| Left-wing | (0.063) | (0.057) | (0.052) | (0.069) | (0.060) | (0.055) |
| | [4.17] | [5.55] | [6.94] | [4.17] | [5.55] | [6.94] |
| | 0.99 | 1.00 | 1.00 | 0.78 | 0.97 | 1.00 |
| Past mayor's party is | 0.013 | 0.002 | -0.003 | -0.018 | -0.023 | -0.023 |
| PT | (0.039) | (0.035) | (0.033) | (0.053) | (0.046) | (0.042) |
| | [4.13] | [5.51] | [6.89] | [4.13] | [5.51] | [6.89] |
| | 1.00 | 1.00 | 1.00 | 0.97 | 0.73 | 0.80 |
| Past mayor's party is | -0.047 | -0.035 | -0.023 | 0.083 | 0.076 | 0.067 |
| Big Right | (0.073) | (0.064) | (0.058) | (0.076) | (0.066) | (0.060) |
| | [3.98] | [5.30] | [6.63] | [3.98] | [5.30] | [6.63] |
| | 1.00 | 1.00 | 1.00 | 0.92 | 0.56 | 0.59 |
| Bandwidth rules | 0.75 x op. | optimal | 1.25 x op. | 0.75 x op. | optimal | 1.25 x op. |

†p<0.1, *p<0.05. For each variable: standard errors in parenthesis, optimal bandwidth in brackets, and the p-value of a KS-test in the last row. Standard errors are clustered by municipality. The estimates represent the difference in outcomes between municipalities with Right and Left-wing mayors for each subsample, and include election fixed-effects. Big right is PMDB/PSDB/PP/DEM/PSD.

**Variable description:** (1) Budget: Municipal budget per household in the previous mayoral tenure (R$ mn); (2) PT for President: Percentage of local votes for PT in the previous presidential election; (3) Turnout: as share of registered voters; (4-5) Longitude and Latitude: in degrees; (6) Municipal Area: km2; (7) Assumes value of one when municipality is part of the semi-arid region; (8) Share of voters: Able voters in each election as a share of the 2000 population; (9) Population: in municipality, census 2000; (10) Inequality: GINI, census 2000; (11) GDP pc: Per capita GDP, IBGE 2000; (12) Gender gap: Difference in gender (female=1, male=0) between winner and runner-up; (13-16) Career gap: Difference in the career dummy (past career in the field=1, otherwise=0) between winner and runner-up in the election; (17) Competitive coalitions: Number of large parties (as defined in the text) that are part of either the winner or runner-up's coalition in the municipality; (18) Past mayor's party is top 2: Dummy that assumes value of 1 when the party of the past mayor is either the winner or the runner-up in the current election; (19) Past mayor's party is left-wing: Dummy that assumes value of 1 when the party of the past mayor is left-wing.

Table E.2: Balance of past policy outcomes (placebo)

| Dependent Variable: | Sample: High Poverty | | | Sample: Low Poverty | | |
|---|---|---|---|---|---|---|
| Past outcome 1 | 0.956 | 0.822 | 0.452 | -0.965 | -0.837 | -0.581 |
| | (1.134) | (0.982) | (0.879) | (0.901) | (0.782) | (0.714) |
| | [4.82] | [6.42] | [8.03] | [4.82] | [6.42] | [8.03] |
| | 0.34 | 0.35 | 0.97 | 0.63 | 0.58 | 0.55 |
| Past outcome 2 | 0.921 | 0.821 | 0.502 | -0.422 | -0.369 | -0.153 |
| | (1.120) | (0.964) | (0.861) | (0.917) | (0.798) | (0.730) |
| | [4.74] | [6.32] | [7.90] | [4.74] | [6.32] | [7.90] |
| | 0.23 | 0.52 | 0.96 | 0.60 | 0.63 | 0.76 |
| Past outcome 3 | 0.685 | 0.610 | 0.321 | -0.496 | -0.484 | -0.282 |
| | (1.111) | (0.960) | (0.859) | (0.915) | (0.796) | (0.726) |
| | [4.71] | [6.28] | [7.85] | [4.71] | [6.28] | [7.85] |
| | 0.40 | 0.35 | 0.95 | 0.42 | 0.59 | 0.66 |
| Bandwidth rules | 0.75 x op. | optimal | 1.25 x op. | 0.75 x op. | optimal | 1.25 x op. |

†p<0.1, *p<0.05. For each variable: standard errors clustered by municipality in parenthesis, optimal bandwidth in brackets, and the p-value of a KS-test in the last row. The estimates represent the difference in outcomes between municipalities with Right and Left-wing mayors for each subsample, and include election fixed-effects. Past outcomes are defined according to columns (1) through (3) of Table E.7 in this appendix.

Table E.3: Placebo test with alternative party categorization

| Dependent Variable: | Pro-poor spending as % of budget | | | Education Gap (winner minus loser) | | |
|---|---|---|---|---|---|---|
| | (A) | (B) | (C) | (A) | (B) | (C) |
| High Poverty | -0.721 | -0.080 | 0.816 | 0.323 | -0.060 | -0.036 |
| | (0.850) | (0.651) | (0.701) | (0.326) | (0.289) | (0.289) |
| Low Poverty | 0.005 | 0.711 | 0.702 | 0.167 | -0.155 | -0.353 |
| | (0.994) | (0.680) | (0.722) | (0.297) | (0.235) | (0.275) |
| Bandwidth (optimal) | 4.60 | 6.17 | 5.62 | 5.23 | 5.31 | 4.79 |
| Observations | 1576 | 2866 | 2661 | 1789 | 2487 | 2280 |

†pp<0.1, *p<0.05. Standard errors are clustered by municipality (parenthesis). The estimates represent the difference in outcomes between municipalities with mayors from placebo party groups 1 and 2, at the discontinuity, from equation 8. Placebos are defined as follows: (A) Treatment group is PPS/PL/PDT/PTB; (B) Treatment group is PT/PSB/PFL/PP; (C) Treatment group is PT/PSB/PFL/PSDB. The control group is always the remaining parties.

Table E.4: Robustness of RDD results

| Dependent Variable: | Pro-poor spending as % of budget | | | Education Gap (winner minus loser) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Excludes Covariates (bandwidth is optimal)** | | | | | | |
| High Poverty | 1.236 | 1.413 | 0.886 | -0.682* | -0.728* | -0.803* |
| | (0.959) | (1.009) | (1.033) | (0.307) | (0.320) | (0.327) |
| Low Poverty | -1.778* | -1.939* | -2.027* | 0.083 | 0.050 | 0.048 |
| | (0.867) | (0.917) | (0.939) | (0.274) | (0.290) | (0.294) |
| Bandwidth (optimal) | 5.29 | 10.37 | 17.81 | 5.40 | 10.59 | 17.88 |
| Observations | 2026 | 3563 | 5217 | 2061 | 3618 | 5224 |
| Polynomial | Linear | Quadratic | Cubic | Linear | Quadratic | Cubic |
| **Quadratic Polynomial (includes covariates)** | | | | | | |
| High Poverty | 1.103 | 0.582 | -0.076 | -0.838* | -0.791* | -0.706* |
| | (0.871) | (0.785) | (0.725) | (0.350) | (0.303) | (0.271) |
| Low Poverty | -2.110* | -2.230* | -2.046* | 0.445 | 0.225 | 0.120 |
| | (0.911) | (0.813) | (0.753) | (0.295) | (0.261) | (0.238) |
| Bandwidth | 7.78 | 10.37 | 12.96 | 7.94 | 10.59 | 13.24 |
| Observations | 2824 | 3563 | 4263 | 2864 | 3618 | 4316 |
| Bandwidth rules | 0.75 x op. | optimal | 1.25 x op. | 0.75 x op. | optimal | 1.25 x op. |
| **Cubic Polynomial (includes covariates)** | | | | | | |
| High Poverty | 1.099 | 0.227 | -0.163 | -0.873* | -0.830* | -0.757* |
| | (0.883) | (0.802) | (0.747) | (0.355) | (0.309) | (0.281) |
| Low Poverty | -2.221* | -2.273* | -2.009* | 0.391 | 0.219 | 0.093 |
| | (0.928) | (0.835) | (0.771) | (0.301) | (0.265) | (0.244) |
| Bandwidth | 13.36 | 17.81 | 22.26 | 13.41 | 17.88 | 22.35 |
| Observations | 4334 | 5217 | 5789 | 4347 | 5224 | 5806 |
| Bandwidth rules | 0.75 x op. | optimal | 1.25 x op. | 0.75 x op. | optimal | 1.25 x op. |

$^{\dagger}p<0.1$, $^{*}p<0.05$. Standard errors are clustered by municipality (parenthesis). The estimates represent the difference in outcomes between municipalities with Right and Left-wing mayors for each subsample, at the discontinuity. Covariates are the ones shown in Table E.1 of this appendix.

Table E.5: Main results for different specifications of Left-Right groups

| Dependent Variable: | Pro-poor spending as % of budget | | | Education Gap (winner minus loser) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| High Poverty | -0.251 | -0.686 | -0.331 | -1.041* | -1.281* | -1.266* |
| | (1.050) | (1.362) | (1.398) | (0.417) | (0.479) | (0.480) |
| Low Poverty | -2.304* | -2.366* | -4.530* | -0.071 | -0.356 | 0.376 |
| | (1.028) | (1.165) | (1.633) | (0.354) | (0.380) | (0.606) |
| Bandwidth | 5.67 | 5.15 | 4.84 | 5.17 | 5.22 | 5.18 |
| Observations | 1093 | 741 | 505 | 1003 | 748 | 532 |
| Right-Wing includes | Largest 5 | All | Far Right | Largest 5 | All | Far Right |
| Left-Wing includes | Largest 2 | PT | Far Left | Largest 2 | PT | Far Left |

†pp<0.1, *p<0.05. Standard errors are clustered by municipality (parenthesis). The estimates represent the difference in outcomes between municipalities with Right and Left-wing mayors for each subsample, at the discontinuity, from equation 8. Far Right parties are PP/PFL/PL/PSD, Far Left are PT/PSB. Party size is based on the number of mayors elected in the period 2004-2016.

Table E.6: Main results for different poverty cutoffs

| Dependent Variable: | Pro-poor spending as % of budget | | | Education Gap (winner minus loser) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| High Poverty | 0.948 | 0.731 | 0.231 | -0.619* | -0.753* | -0.629* |
| | (0.678) | (0.739) | (0.871) | (0.271) | (0.290) | (0.314) |
| Low Poverty | -2.797* | -2.026* | -1.323* | 0.243 | 0.250 | -0.007 |
| | (0.841) | (0.762) | (0.670) | (0.260) | (0.249) | (0.243) |
| Bandwidth (optimal) | 5.29 | 5.29 | 5.29 | 5.40 | 5.40 | 5.40 |
| Observations | 2026 | 2026 | 2026 | 2061 | 2061 | 2061 |
| Cutoff quantile | 40th | 50th | 60th | 40th | 50th | 60th |

†pp<0.1, *p<0.05. Standard errors are clustered by municipality and presented in parenthesis. The estimates represent the difference in outcomes between municipalities with Right and Left-wing mayors for each subsample, at the discontinuity, from equation 8.

Table E.7: Main results for alternative variable categorizations

| Dependent Variable: | Pro-poor spending as % of budget | | | Education Gap (winner minus loser) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| High Poverty | 0.731 | 0.710 | 0.620 | -0.148* | -0.520* | -0.112† |
| | (0.739) | (0.711) | (0.719) | (0.072) | (0.234) | (0.064) |
| Low Poverty | -2.026* | -1.689* | -1.800* | 0.015 | -0.002 | 0.006 |
| | (0.762) | (0.756) | (0.760) | (0.065) | (0.191) | (0.064) |
| Bandwidth (optimal) | 5.29 | 5.68 | 5.63 | 5.69 | 5.55 | 5.28 |
| Observations | 2026 | 2157 | 2141 | 2160 | 2114 | 2020 |

†pp<0.1, *p<0.05. Standard errors are clustered by municipality and presented in parenthesis. The estimates come from equation 8 and represent the difference in outcomes between municipalities with Right and Left-wing mayors for each subsample, at the discontinuity. The columns are defined as follows: Column (1) shows the main specification; column (2) adds social assistance spending; and column (3) subtracts spending with security. Column (4) codes the education of each candidate as a dummy that assumes value of 1 if the candidate has a bachelor's degree, and uses the difference between winner and loser as before; column (5) uses the education of the mayor only (winner); and column (6) also uses the education of the mayor only, coded as the *college* dummy.

Table E.8: Campaign Expenses

| Dependent Variable: | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| High Poverty | -0.039 | 0.011 | 0.045 | -0.059 | -0.004 | 0.033 |
| | (0.168) | (0.145) | (0.129) | (0.169) | (0.146) | (0.130) |
| Low Poverty | 0.166 | 0.091 | 0.043 | 0.176 | 0.099 | 0.053 |
| | (0.175) | (0.154) | (0.142) | (0.173) | (0.153) | (0.141) |
| Bandwidth (optimal) | 4.61 | 6.15 | 7.69 | 4.61 | 6.15 | 7.69 |
| Observations | 1614 | 2123 | 2552 | 1614 | 2123 | 2552 |
| Bandwidth rules | 0.75 x op. | optimal | 1.25 x op. | 0.75 x op. | optimal | 1.25 x op. |
| Covariates | Yes | Yes | Yes | No | No | No |

†pp<0.1, *p<0.05. Standard errors are clustered by municipality (parenthesis). The estimates represent the difference in outcomes between municipalities with Right and Left-wing mayors for each subsample, at the discontinuity, from equation 8.

Table E.9: Main Results: OLS and Panel

| Dep. Variable: | Education gap | | Mayor's education | | Pro-poor spending | |
|---|---|---|---|---|---|---|
| Right-wing | 0.044 | 0.054 | 0.062 | 0.076 | -0.628* | -0.757* |
| | (0.110) | (0.160) | (0.084) | (0.102) | (0.216) | (0.322) |
| Poverty | 0.286 | 0.168 | -0.656* | 0.575 | -0.488 | -0.913 |
| | (0.186) | (0.782) | (0.152) | (0.573) | (0.392) | (1.862) |
| Right-wing x Poverty | -0.486† | -0.659† | -0.460* | -0.482* | 1.151* | 1.625* |
| | (0.256) | (0.376) | (0.201) | (0.246) | (0.542) | (0.805) |
| Observations | 7511 | 7511 | 7511 | 7511 | 7511 | 7511 |
| Fixed effects | No | Yes | No | Yes | No | Yes |

†p<0.1, *p<0.05. Standard errors are clustered by municipality (parenthesis). The dependent variable is regressed on a dummy that indicates whether the mayor is Right-wing; on the continuous level of poverty; and their interaction. For pro-poor spending, the dependent variable is the difference between the current and the last period's share of budget. Fixed effects are by period and municipality. The regression also controls for the pre-treatment value of the following variables defined in page E-19: budget; gender gap; career gaps (health/edu and business); share of voters; on dummmies the indicate whether the past mayor's party is PT, or from one of the big Right-wing parties, or is now one of the top 2 contenders; and also on the past municipal margin of victory.

Table E.10: Heterogeneity in the education gap in high-poverty areas

| Sample split: | PT vs. PSDB races only | | Leftist mayor in previous 4 years | | Leftist candidate has bachelor's degree | |
|---|---|---|---|---|---|---|
| | (No) | (Yes) | (No) | (Yes) | (No) | (Yes) |
| RDD estimate | -0.697* | -2.157* | -0.641 | -1.353* | 0.847† | -2.704* |
| | (0.328) | (0.961) | (0.404) | (0.586) | (0.439) | (0.337) |
| Bandwidth | 4.72 | 4.72 | 4.43 | 4.43 | 4.72 | 4.72 |
| Observations | 884 | 63 | 568 | 271 | 505 | 442 |

†p<0.1, *p<0.05. Standard errors are clustered by municipality (parenthesis). The dependent variable is the education gap, and the sample is the one with high poverty municipalities. The estimates represent the difference in outcomes between municipalities with Right and Left-wing mayors, at the discontinuity. The coefficients come from a pooled regression that estimates RDD effects for the two subsamples, according to the sample split described in the header.

Table E.11: Education of partisan council members: panel

| Councilors with: | secondary education | | bachelor's degree | | | |
|---|---|---|---|---|---|---|
| | (A) | (B) | (A) | (B) | (Ax) | (Bx) |
| Right-wing | 0.231 | -0.014 | 0.012 | -0.021 | -0.333 | -0.091 |
| | (0.172) | (0.072) | (0.097) | (0.041) | (0.204) | (0.083) |
| Wealth | -0.656 | -0.424$^{\dagger}$ | -0.251 | -0.020 | -0.635 | -0.042 |
| | (0.546) | (0.231) | (0.317) | (0.136) | (0.637) | (0.258) |
| Right-wing x Wealth | -0.189 | 0.173 | -0.028 | 0.060 | 0.019 | 0.081 |
| | (0.273) | (0.115) | (0.163) | (0.067) | (0.338) | (0.133) |
| Observations | 7503 | 7503 | 7503 | 7503 | 2858 | 2858 |

$^{\dagger}$p<0.1, *p<0.05. The dependent variable is always the gap between the outcomes for the winner and loser. Includes time and municipality fixed effects, and standard errors are clustered by municipality (parenthesis). The dependent variable is regressed on a dummy that indicates whether the mayor is Right-wing; on the continuous level of wealth; and their interaction. Included pre-treatment covariates are listed in the footnote of Table E.9. Columns (A) consider all councilors. Columns denoted by (B) only consider the 2 most voted councilors in each municipality. The (x) denotes the subsample that compares Right-wing parties to PT only.

Table E.12: Results for non-binary measures of poverty

| Dependent Variable: | Pro-poor spending as % of budget | | | Education Gap (winner minus loser) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Effect with MAX poverty | 1.975 | 3.377$^{\dagger}$ | 5.166$^{\dagger}$ | -1.082* | -1.025$^{\dagger}$ | -1.954$^{\dagger}$ |
| | (1.246) | (1.728) | (2.683) | (0.468) | (0.621) | (1.038) |
| Marg. effect of wealth | -0.826* | -6.181* | -11.072* | 0.270* | 1.199 | 3.230$^{\dagger}$ |
| | (0.373) | (2.625) | (5.093) | (0.131) | (0.897) | (1.865) |
| Bandwidth | 5.29 | 5.29 | 5.29 | 5.40 | 5.40 | 5.40 |
| Observations | 2026 | 2026 | 2026 | 2061 | 2061 | 2061 |

$^{\dagger}$p<0.1, *p<0.05. Standard errors are clustered by municipality and presented in parenthesis. This specification reflects the estimates of equation 8, where the variable that measures municipal WEALTH is NOT binary. In this case, rather than showing the treatment effects for two different subsamples, we present the estimate for $\beta_1$ and $\beta_5$, which can be interpreted as the RDD effect when poverty is maximum, and the marginal change in this effect for every unit increase in wealth; respectively. Columns (1) and (4) code wealth as a categorical variable with 5 different levels, defined according to the quantiles of poverty. Columns (2) and (5) code wealth using an alternative (and continuous) measure of municipal development, the Human Development Index, calculated by IBGE in 2000. Columns (3) and (6) code wealth using an the continuous poverty measure defined in this paper.

## REFERENCES

Brollo, Fernanda and Tommaso Nannicini. 2012. "Tying Your Enemy's Hands in Close Races: The Politics of Federal Transfers in Brazil." *American Political Science Review* 106:742–761.

Dal Bó, Ernesto and Frederico Finan. 2018. "Progress and Perspectives in the Study of Political Selection." *Annual Review of Economics* 10 (1):541–575.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne. 2017. "Who Becomes A Politician?" *The Quarterly Journal of Economics* 132 (4):1877–1914.

Feierherd, Germán. 2020. "How Mayors Hurt Their Presidential Ticket: Party Brands and Incumbency Spillovers in Brazil." *The Journal of Politics* 82 (1):195–210.

Ferraz, Claudio and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123 (2):703–745.

Hott, Henrique Augusto Campos Fernandez and Sergio Naruhiko Sakurai. 2020. "Party switching and political outcomes: evidence from Brazilian municipalities." *Public Choice* :1–36.

Klašnja, Marko and Rocío Titiunik. 2017. "The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability." *American Political Science Review* 111 (1):129–148.

Novaes, Lucas M. 2018. "Disloyal brokers and weak parties." *American Journal of Political Science* 62 (1):84–98.