Can Descriptive Representation Help The Right Win Votes From The Poor? Evidence From Brazil

Zuheir Desai & Anderson Frey[†]

December, 2020

Abstract

The electoral success of the Right in poor nations is typically attributed to non-policy appeals such as clientelism. Candidate profiles are usually ignored, because if voters value class-based descriptive representation, it should be the Left that uses it. In this article we develop and test a novel theory of policy choice and candidate selection that defies this conventional wisdom: it is the Right that capitalizes on descriptive representation in high poverty areas. The Right is only competitive in poor regions when it matches the Left's pro-poor policies. To credibly shift its position, it nominates candidates that are descriptively closer to the poor. Using a regression discontinuity design in Brazilian municipal elections, we show that Right-wing mayors spend less on the poor than Left-wing mayors only in low-poverty municipalities. In high-poverty municipalities, not only does the Right match the Left's policies, it also does so while nominating less-educated candidates.

Word count: 9963

Department of Politics, Princeton University. Fisher Hall, 315. Princeton University, NJ, 08540. email: zdesai@princeton.edu.

[†]Department of Political Science, University of Rochester. Harkness Hall, 320B. Rochester, NY, 14627. email: anderson.frey@rochester.edu.

Right-wing parties often win elections in developing nations where voters are overwhelmingly poor. Prevailing explanations for this puzzle typically focus on how they build a portfolio of electoral appeals such as clientelism (Murillo and Calvo, 2019), ethnic mobilization (Huber, 2017), or private provision of social services (Thachil, 2014). The case of Brazil is similar: clientelism and personalistic politician-voter ties have been the primary explanation for why "conservative parties fare best electorally among relatively poor, less educated" voters (Mainwaring et al., 2000), despite the fact that the Left is more likely to support redistributive policies.

Not surprisingly, these explanations seldom focus on the descriptive profile of the candidates nominated by the Right. The literature on political behavior suggests that voters value descriptive representation (Carnes and Lupu, 2016; Dal Bo et al., 2019), and are more likely to trust and feel included by politicians descriptively closer to them (Gay, 2002; Hayes and Hibbing, 2017; Lawless, 2004). In turn, when politicians stress that "I am one of you", their common identity helps them to better understand the needs of voters (Carnes and Lupu, 2015), and provides incentives for the betterment of the status of their shared social group (Shayo, 2009). Thus, if there are electoral returns to class-based descriptive representation, it is natural to expect that Left -wing parties are the ones that capitalize on it in poor areas. Former Brazilian President Lula (2003-2010) is a clear example. He often used his lack of education to emphasize his ability to succeed as a politician, and to implement redistributive policies, mentioning for example that "a steelworker without a bachelor's degree created more universities than the PhDs that previously governed the country."[1]

However, in this article we uncover an empirical pattern in Brazilian municipalities that at first defies this conventional wisdom: it is the Right that capitalizes on descriptive representation in the poorest areas. We interpret this finding within the literature on party strategies in developing nations, with a novel theory of policy choice and candidate selection. The idea is simple: Right-wing parties are only competitive in very poor areas if they implement pro-poor policies that voters most often identify with the Left, and might not be credible for the Right. However, if voters are also more likely to trust candidates that 'look like them', the Right can credibly shift local policy positions leftwards by nominating candidates that are less educated than the average politician, and therefore descriptively closer to the poor.[2]

Our argument is best illustrated by the 2016 mayoral race in Camaçari (BA). Right-wing DEM (*Democratas*) and Left-wing PT (*Partido dos Trabalhadores*) had mayoral candidates with opposite profiles: DEM nominated Elinaldo Araújo, a former manual laborer with only a secondary education. PT's candidate was Luiz Caetano, a congressman and a biochemist. Camaçari is a small, but strategic municipality, as it houses the largest petrochemical complex in Brazil. Nevertheless, it is highly unequal and poor. Not surprisingly, Elinaldo's campaign emphasized that he was a 'true' representative of the people, in spite of being nominated by the traditional elite party in the state. A party leader described him as "a humble person, who

[1] Tania Monteiro, "Lula diz querer eleger alguém para fazer mais do que fez", *Politica*, June 2009, https://bit.ly/2X0U3rq.
[2] Education is highly correlated with economic class within countries (Krueger and Lindahl, 2001).

does not have many possessions, an individual that identifies with the poorer people in Camaçari."[3] After winning, Elinaldo himself framed his low education as a virtue, saying that his opponent "cannot accept the fact that he lost the election to a humble person, without a college degree, but that understands the people."[4]

We develop this theory in a formal model of electoral competition between two ideologically opposed parties, building on Desai (2020). The model first provides the following hypothesis for both the implemented policy and the profile of Left and Right-wing candidates: (i) in high poverty areas, both parties offer similar policies. Lower programmatic differentiation at the local level boosts the chances of the Right winning the election. However, because pro-policies are only in line with the Left's ideals, the Right nominates less educated candidates; and (ii) in low poverty areas the prediction is reversed: policies are more divergent, following the national pattern of party ideals, and both candidates come from the educated elite.

The empirical evidence comes from four Brazilian municipal elections (2004-2016), which offer a suitable environment to test this theory. First, Brazil is a large and unequal democracy where we can observe candidate profiles and policy choices by the same parties in municipalities of high and low poverty. Second, Brazil's multiparty system exhibits a highly consensual, broad Left-Right divide between the main parties, as shown by surveys with voters, experts and politicians (Power and Jr., 2009; Samuels and Zucco Jr., 2018). Also, politicians in these groups display significantly different preferences for redistribution in the period under analysis (Power and Zucco Jr., 2012).

Our measure of pro-poor policy implemented by mayors is the share of the municipal budget spent on health, sanitation, education and housing. We identify the causal effect of party ideology on policies with a regression discontinuity design (RDD) in close races between Left and Right-wing parties. We also use the education level of mayoral candidates as a measure of their ability to descriptively identify with the poor. Because this variable is determined before the election, the estimates based on this outcome cannot be interpreted as a causal effect, but rather as a correlation between party ideology and candidate education.

The main estimates are in line with our first theoretical hypothesis: in low poverty areas, Right-wing mayors spend significantly less on pro-poor categories, in line with the preferences revealed by Left and Right-wing politicians in national surveys. In these locations, both parties field highly educated mayoral candidates. In high poverty areas, policy differences disappear, as both Right and Left-wing mayors increase their pro-poor spending to similar levels. However, Right-wing candidates are less educated than their Leftist competitors.

Our interpretation of these results is only valid in a context where party brands drive policy choices at the municipal level, and their brands are recognized by voters. This at first poses at threat to our applica-

[3] Aparecido Silva, "Recepção de Elinaldo em Camaçari é demonstração de que o povo confia, diz Neto", *BNews*, December 2015, http://bit.ly/32qpAmJ.

[4] Alexandre Galvão and Gabriel Nascimento, "Elinaldo nega dedo de Neto, de Azi em reforma administrativa", *Metro1*, April 2018, http://bit.ly/2IYqGys.

tion, given the conventional wisdom that Brazilian municipal races are often driven by non-policy issues, and that the ideological politician-voter linkages are frail. Thus, we take additional steps to show that our framework is indeed consistent with the application. First, although it is evident that programmatic competition is *not* the primary driver behind municipal elections, our results show that it does play a role in these races. Otherwise, we would not see a consistent difference between Right and Left-wing spending in low poverty municipalities.[5] When we consider our finding in the context of the extensive evidence in the literature that Brazilian voters both *observe* and *recognize* the policies and performance of mayors, and also *punish/reward* parties accordingly (Boas, Hidalgo, and Toral, 2020; Feierherd, 2020; Ferraz and Finan, 2008; Klašnja and Titiunik, 2017), then party brands do matter locally, at least to a certain extent.

Second, we show that the average profile of local party coalitions also reflects the national Left-Right divide. In the Brazilian multiparty system, nearly every mayor is supported by a large coalition averaging more than 6 parties. In fact, on average, parties are more likely to formally support a coalition mayoral candidate than to run their own. In this context, although "inconsistent" alliances often exist (e.g. a Left party supporting a Rightist mayoral candidate), parties are still much more likely to support candidates within their ideological corner than the alternative.

Third, our theory also accounts for the fact that party brands might be irrelevant in many local elections. Formally, we allow parties to have mixed-motivation, caring both about holding office as well as policy. We show that the less policy-motivated parties are, the lower the relevance of programmatic competition in the mayoral race, and the less likely we are to observe the empirical patterns outlined before. We test this second hypothesis using the ideological alignment within local party coalitions as a proxy for the relevance of party brands in each local race. The results are consistent: the patterns of policy differentiation and candidate selection are stronger in high alignment races and all but disappear in others.

The fact the Left consistently nominates highly educated politicians in poor areas also has strong implications for how we interpret the overall pattern in the data. Our results suggest that although descriptive representation has direct value for voters as shown by previous scholarship, it *also* has an indirect effect by making policy deviations credible. If this was not the case, both parties would be equally likely to nominate lower educated candidates.

We also asses three alternative explanations for this surprising nomination pattern in high poverty areas. First, if uneducated Right-wing politicians are systematically better at clientelism, the pattern could be interpreted as consequence of the practice, which is ubiquitous in Brazil (Nichter, 2018). However, we provide evidence suggesting that this is unlikely the case: (i) the nomination pattern remains robust for a subsample with the most programmatic parties only; (ii) Right and Left-wing parties spend a similar

[5]Note that we are not the first to use a RDD to evaluate whether national party brands matter for local policy. Pettersson-Lidbom (2008) also show that party ideology matters for local economic outcomes in Swedish municipalities. Our article goes further to show that ideology matters locally *only* when it is optimal for parties.

amount of funds in mayoral campaigns, a proxy for clientelistic capacity; and (iii) the 2010 LAPOP[6] survey shows that the education of the Right-wing candidate is uncorrelated with voters' perceptions on vote buying. Second, we show that this pattern is not driven by systematic differences in the pool of potential candidates of Left and Right-wing parties. There is no difference in the number of highly educated local councilors across the Left and the Right.[7] Third, the pattern could also arise if parties face differential costs of electing uneducated mayors. That said, we use various measures of administrative performance and mobilization capacity to show that the cost indeed exists, but it is uniform across ideological groups.

Finally, this article complements the existing literature on party portfolio strategies, particularly on programmatic shifts. Recent work in Latin America has already shown that the Right becomes more attractive to poor voters by shifting policy towards redistribution across time and constituencies (Murillo and Calvo, 2019). However, less attention has been paid to the credibility problem around these policy changes, which is the basis of the theory proposed here. In interpreting the empirical findings within this framework, this article also relates to the literature on mechanisms used by parties to commit to targeted redistribution (Huber, 2017). Our findings also have significant implications for the burgeoning literature on political selection and its concern with the profile and quality of citizens that enter politics (Carnes and Lupu, 2015; Dal Bó et al., 2017), especially to the extent that it reveals a context where parties optimally nominate candidates with lower human capital.

A MODEL OF CANDIDATE NOMINATION AND POLICY CHOICE

We develop a simple theory that incorporates findings of the political behavior literature on descriptive representation with mixed-motivation parties that strategically offer programmatic policies. The model is intended to be applied to developing democracies, where the majority of voters are poor, and in particular to Brazil.

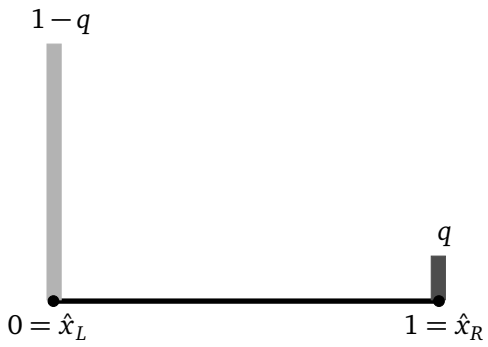
MODEL SETUP

Our framework is based on Desai (2020), with two ideologically opposed parties L and R. The programmatic Left-Right dimension is defined as the $[0, 1]$ interval, and positions closer to 0 represent Leftist, pro-poor policies. The ideal point of party i is given by $\hat{x}_i$. Accordingly, L's ideal point is situated at 0, while that of R is situated at 1. There are two classes of voters, poor (P) and affluent (A), which share the ideal points of L and R; respectively. The distribution of voters is indexed by the proportion of A voters, denoted by q. Since we focus on a developing context, we assume that the affluent are always in the minority ($q < \frac{1}{2}$). This setup is reflected in Figure 1.

[6] 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.

[7] Given that the council is a stepping stone for mayoral candidacies, and they come from the very same pool as mayors (party membership rolls), this is a very natural measure of party candidate pools.

Figure 1: The electoral environment



The light and dark bars represent the distribution of voter types P and A. Poor voters share their ideal point with party L and affluent voters share their ideal point with party R.

Parties: Before the election, parties announce policies and choose candidates. The candidate pool for each party contains elite candidates, which are descriptively closer to affluent voters, and non-elite ones, which are descriptively closer to the poor. An observable feature of elitism is, for example, education.

While policies are chosen from the $[0,1]$ interval, candidate selection is binary. We say that if $c_i = 1$, then i's candidate provides descriptive representation to the group that does not share its ideal point, and $c_i = 0$ otherwise. Let $\mathbf{x}_i = (x_i, c_i) \in [0,1] \times \{0,1\}$ be the policy announcement and candidate choice of party i. Both parties have mixed motivation, i.e., they benefit both from holding office and policy outcomes. The weight attached to policy benefits is given by $w \in [0,1]$, and the office benefit is normalized to 1. Denote

$$\varphi(\mathbf{x}_i) \quad = \quad \begin{cases} w\hat{x}_i + (1-w)x_i & \text{if } c_i = 0 \\ x_i & \text{if } c_i = 1 \end{cases} \tag{1}$$

as the final policy implemented by i on winning the election with candidate c_i. The policy function (1) indicates that the more policy-motivated parties are, the more likely it is that their implemented policy deviates from their announced policy position in absence of a descriptively representative candidate. All-else equal, both parties face a cost of κ when choosing a non-elite candidate. We interpret this cost framework as context specific. In Appendix B, we provide an extensive discussion on the motivation behind this assumption. We use both the literature and new empirical evidence to show that this particular cost framework is not only apt to the Brazilian context, but that it also fits the empirical results better than alternative assumptions.[8]

[8] In summary, on the supply-side the candidate pool in Brazil is biased in favor of higher educated candidates due to self-selection into politics. On the demand-side, we show that less educated candidates perform worse on many measures of administrative performance indicative of lower valence, and are also worse at brokering votes for their parties in subsequent elections. We also discuss alternative cost structures to the one presented in the text and show that the resulting predictions are at odds from the empirical patterns observed in Brazilian data.

The objective functions of the two parties are given by

$$\begin{aligned} V_L(\mathbf{x}_L,\mathbf{x}_R) &= w\big((1-F(\mathbf{x}_L,\mathbf{x}_R))\cdot u_L(\varphi(\mathbf{x}_L))+F(\mathbf{x}_L,\mathbf{x}_R)\cdot u_L(\varphi(\mathbf{x}_R))\big) \\ &\quad +(1-w)(1-F(\mathbf{x}_L,\mathbf{x}_R))-(1-c_L)\kappa \qquad (2) \\ V_R(\mathbf{x}_L,\mathbf{x}_R) &= w\big((1-F(\mathbf{x}_L,\mathbf{x}_R))\cdot u_R(\varphi(\mathbf{x}_L))+F(\mathbf{x}_L,\mathbf{x}_R)\cdot u_R(\varphi(\mathbf{x}_R))\big) \\ &\quad +(1-w)F(\mathbf{x}_L,\mathbf{x}_R)-c_R\kappa, \qquad (3) \end{aligned}$$

where $u_i(x)=-|\hat{x}_i-x|$ and $F(\mathbf{x}_L,\mathbf{x}_R)$ is the probability that party R wins the election.

Voter behavior: A voter of class j receives the following utility from party i

$$u_j(\mathbf{x}_i) \quad = \quad -|\hat{x}_j-\varphi(\mathbf{x}_i)|. \qquad (4)$$

Let

$$\Delta u_j(\mathbf{x}_L,\mathbf{x}_R) \quad := \quad u_j(\mathbf{x}_L)-u_j(\mathbf{x}_R) \qquad (5)$$

be the utility differential to voter of class j from the candidate-policy pairs of both parties. Each voter j has two idiosyncratic components to her utility, individual and aggregate. The voter has an individual preference η_j for party R, which is drawn identically and independently from a distribution G. This represents how voter j evaluates party R's characteristics on any other criteria other than economic policies (e.g. clientelism). In addition to this individual-level idiosyncratic component, all voters receive an aggregate shock ϵ, which is distributed according to the distribution H. This shock represents the aggregate popularity of party L over party R. It affects each voter identically, thereby resulting in parties facing aggregate uncertainty about the election outcome. A negative realization of ϵ means that the electorate is biased towards party R.

Voter j votes for party R if and only if the condition below holds

$$\begin{aligned} u_j(\mathbf{x}_R)+\eta_j &\geq u_j(\mathbf{x}_L)+\epsilon \\ \iff \eta_j &\geq \Delta u_j(\mathbf{x}_L,\mathbf{x}_R)+\epsilon. \end{aligned}$$

Thus, the proportion of voters voting R is $1-G(\Delta u_j(\mathbf{x}_L,\mathbf{x}_R)+\epsilon)$. The total vote share for party R is given by the following random variable

$$VS_R(\mathbf{x}_L,\mathbf{x}_R;\epsilon) \quad = \quad \underbrace{(1-q)\left(1-G\big(\Delta u_P(\mathbf{x}_L,\mathbf{x}_R)+\epsilon\big)\right)}_{\text{Vote share from poor}} \quad + \quad \underbrace{q\left(1-G\big(\Delta u_A(\mathbf{x}_L,\mathbf{x}_R)+\epsilon\big)\right)}_{\text{Vote share from affluent}}, \qquad (6)$$

and the vote share of party L is analogously $1-VS_R(\mathbf{x}_L,\mathbf{x}_R;\epsilon)$. Note that the model implies that the smaller Δu_j, the *less* voters vote on the basis of their economic preferences. The probability that R wins

the election is the probability that its vote share is greater than that of party L, and is given by

$$F(\mathbf{x}_L, \mathbf{x}_R) := \int \mathbb{I}\left\{VS_R(\mathbf{x}_L, \mathbf{x}_R; \epsilon) \geq \frac{1}{2}\right\} h(\epsilon)\, \mathrm{d}\epsilon. \tag{7}$$

The probability that L wins the election is simply $1 - F(\mathbf{x}_L, \mathbf{x}_R)$. We assume that G is uniform on $[-2, 2]$ and H is uniform on $[-\psi, \psi]$, where $\psi < 1$, for tractability.

The game proceeds as follows:

1. Parties choose their policy announcement x_i and candidate c_i.
2. Individual and aggregate shocks η_j and ϵ are realized.
3. Voters sincerely vote for their preferred party.
4. The winning party implements its policy according to $\varphi(\mathbf{x}_i)$.

In what follows we make the following assumption on the nomination cost.

Assumption 1. *The cost to nominate a non-elite candidate is such that* $\kappa < \frac{\psi}{8}$.

Note that while policy choice is continuous, we adopt a discrete candidate selection framework for the sake of a cleaner exposition. In this class of models, candidate selection follows a cost-benefit analysis. Evidently, descriptive representation as a costly strategy is only viable if the cost is outweighed by the benefit, which is a weighted average of office and policy related benefits. Assumption 1 precludes the existence of a trivial equilibrium where a purely policy-motivated Right rationally chooses to lose the election *for sure* in very poor districts.

Before proceeding, we make a remark on the structure of the model. Our model captures multiple competitive frameworks. The parameter w measures how "programmatic" competition is: it simultaneously measures how motivated parties are on policy, as well as how final policy reflects party ideal points. If $w = 1$, we are in a purely policy motivated setting where parties care *exclusively* about policy, and no policy position other than the party's ideal point is ex-ante credible. On the contrary, if $w = 0$, then parties are purely office motivated, and since party ideal points have no meaning, there is no disconnect between implemented policy and party positions.

Model results

We present our results in two propositions, each focusing on a particular kind of competitive framework. Party incentives depend on the nature of competition as well as poverty. These two considerations drive parties to reduce or increase programmatic differentiation, which in turn shapes candidate selection patterns.

First, we look at a relatively office motivated competitive framework.

Proposition 1 (Office-motivated framework)**.** *There exists a* $\underline{w} \in (0,1)$ *such that for all* $w \leq \underline{w}$ *and for all* $q \in \left(0, \frac{1}{2}\right)$

1. *Party L never nominates a non-elite candidate and implements* $x_L^* = 0$*;*

2. *Party R never nominates a non-elite candidate and implements* $\varphi(\tilde{x}_R, 0)$ *upon winning the election where*

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

It is best to first focus on the case when both parties are purely office-motivated in order to understand this result. In this case, L and R both choose policies to maximize their probability of winning. Since their policy preferences are irrelevant, all promises are credible. In such a scenario, L and R *always* converge to the policy of the median voter. Since descriptive representation as a tool to establish credibility is unnecessary, both parties nominate elite candidates. When w is positive but small, this logic continues to hold. While parties have some policy-motivation, it is not as high as the pressure to converge to the (poor) median voter's position to maximize the chance of winning. At the same time, R can credibly converge to a large degree because voters know that parties are very office-motivated and that their policy preferences do not matter for competition. As a result, in equilibrium, the expected policies of both parties are relatively close to each other and to the median voter,[9] and neither party resorts to the use of descriptive representation.

Next, we turn to a competitive framework where parties place a high weight on their policy preferences and because voters are aware of their policy-motivation, they are skeptical of deviations from party ideals. The following proposition outlines equilibrium strategies by the Left and the Right.

Proposition 2 (Policy-motivated framework)**.** *There exists a* $\bar{w} \in (0,1)$ *such that for all* $w \geq \bar{w}$*, there exists a* $\bar{q} \in \left(0, \frac{1}{2}\right)$ *such that*

1. *Party L never nominates a non-elite candidate and implements* $x_L^* = 0$*;*

2. *If* $q < \bar{q}$*, party R nominates a non-elite candidate and implements* x_R^* *upon winning the election where*

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

3. *If* $q > \bar{q}$*, party R nominates an elite candidate and implements its ideal point upon winning the election.*

Since $q < \frac{1}{2}$, party L maximizes policy divergence with R in order to capitalize on its brand advantage. L's ideal point maximizes both its probability of winning as well as its policy benefit, thereby making it unnecessary to nominate a non-elite candidate. Because w is high, party R faces a trade-off between

[9] Note that $\varphi(\tilde{x}_R, 0)$ is weakly increasing in q but is close to 0 when w is small and q is not near $\frac{1}{2}$.

proposing a policy position very close to its ideal point with a standard elite candidate, or using a tailored (pro-poor) policy position with a non-elite candidate. This latter candidate choice is costly but can significantly increase R's probability of winning. When policy promises are very similar, individual and aggregate shocks matter more to the election result, which may swing the race in R's favor.

As q increases, this trade-off becomes more binding because the optimal policy by R is increasing in q. As the share of poor voters decreases, the gains from class-based descriptive representation start to decrease in relation to the nomination cost. Thus, there exists a cutoff $\bar{q}$ that divides the parameter space into low and high poverty regions. In the high poverty region, R pays the cost to nominate non-elite candidates and credibly reduce programmatic differentiation with L. In the low poverty region, R nominates elite candidates, and diverges more in terms of policy from L.

Importantly, we do not consider direct voter preferences regarding descriptive representation in the model. As discussed in the introduction, there is certainly a wealth of evidence suggesting that voters value descriptive representation independent of its effect on policy commitments. However, in order to focus on this indirect policy effect, we choose to abstract away from direct preferences over representative candidates. We revisit this in Appendix B where we discuss the robustness of our theoretical results when poor voters also directly value descriptive representation, in addition to its indirect effect on policy credibility. To summarize that discussion, our results continue to hold in a qualitatively similar fashion as long as poor voters' preferences for descriptive representation are independent of the candidate's party affiliation, and low enough relative to the cost of nominating a non-elite candidate. That is, the direct benefits of nominating an uneducated candidate are the same for both parties, and do not mitigate electoral uncertainty enough relative to the cost.

Empirical implications

We derive two main theoretical insights that we can take to the data, summarized in the following hypotheses.

Hypothesis 1. *When parties are policy-motivated **and** their brands are valued by voters:*

(a) In high poverty regions, policy differentiation is low, but the Right-wing candidates are less educated.

(b) In low poverty regions, candidate profiles are similar, but policy differentiation is high.

Hypothesis 2. *When programmatic labels have little meaning for voters and parties:*

(a) Policies and candidate profiles of Right and Left are indistinguishable, in both low and high poverty regions.

First, when parties are policy-motivated and their brands are recognized by voters, our model suggests that the Right chooses to reduce policy differentiation with the Left when the electorate is extremely poor. Since voters recognize national brands the Right cannot easily commit to pro-poor promises. In order to

demonstrate such commitment, they nominate candidates that are less educated like poor voters, and are therefore more likely to be trusted by them. As the electorate becomes wealthier, the policies offered by the parties diverge, in line with their ideological positions, and candidate profiles converge.

Second, we show that it is necessary that *both* parties and voters care about programmatic brands for these results to hold. When programmatic labels have little meaning for both parties and voters, both policies and candidate profiles converge.

Party ideology and local politics in Brazil

Before moving on to the empirical exercise, we provide some background regarding the ideological partisan divide and local politics in Brazil. Brazil has a fragmented party system with institutions that favor candidate-centered elections, and where a sizable share of the electorate does not identify with any party (Ames and Smith, 2010).[10] Moreover, clientelism is pervasive (Nichter, 2018), and parties often form ideologically inconsistent coalitions. In recent work, Samuels and Zucco Jr. (2018) argue that the main partisan cleavage in the period under analysis is between partisans and 'anti-partisans'. This dichotomy has, on one side, voters that favor the labor party PT (the 'partisans'), and on the other side, voters that reject it but do not necessarily identify with any specific Right-wing party.

However, even if this context undermines the scholars' ability to build a fine-grained ideological scale for all parties, there is a widely accepted consensus by experts, voters, and politicians alike on what constitutes the broader Left-Right (L-R) divide in Brazil. Recent surveys with legislators place the main parties in a fairly cohesive scale (Power and Jr., 2009; Power and Zucco Jr., 2012), displaying a clear-cut divide between Left and Right (Figure 2).[11] In fact, they emphasize how the ideological distances between parties are less meaningful within each group, but significantly large across the divide. This categorization is confirmed by the DALP expert survey (2008),[12] and supported by the 2010 LAPOP survey,[13] which shows that voters that identify with a Left party are also significantly more likely to label themselves as Leftist (Figure E.6).

The way in which voters and politicians understand Left-Right ideology might also vary across countries and time. For the period of our analysis (2004-2016), this divide is highly aligned with politicians' views on redistribution: Leftist legislators are much more likely to express pro-poor preferences (Power and Zucco Jr., 2012). The same is shown in contemporary voter surveys: Lupu (2016) for example finds that, "when asked which party most protects them, poor respondents in 2006 were far more likely to choose the PT than any other party."

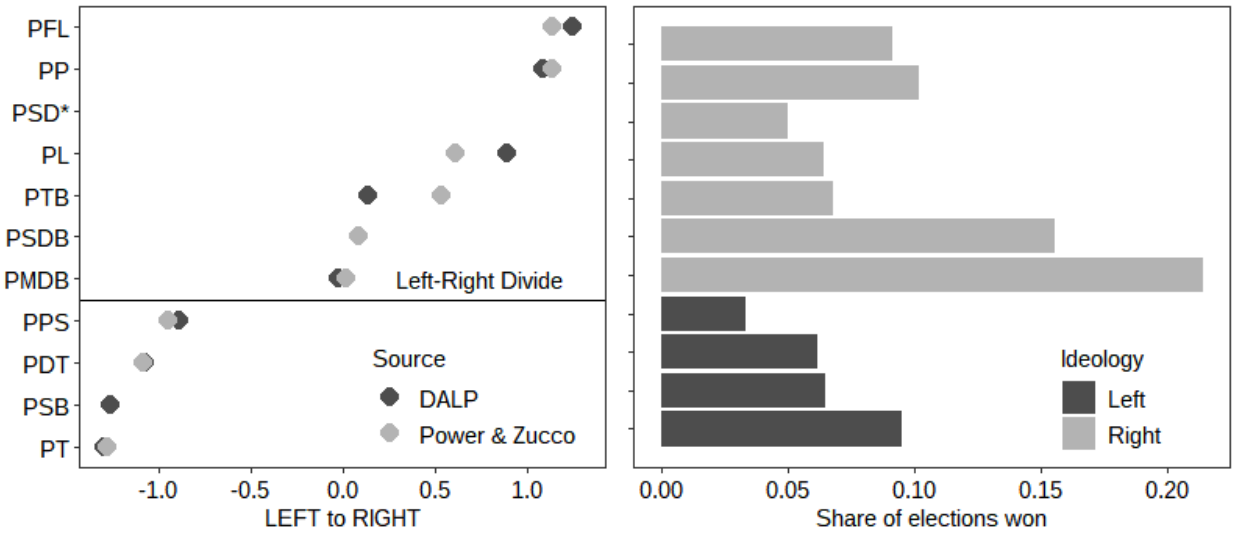
[10] The authors show that the voters less likely to identify themselves with party ideologies also "tend to be latent Rightists."

[11] The 11 largest parties by the number of mayors in the period were PMDB, PSDB, PT, PFL, PP, PTB, PL, PDT, PSB, PPS and PSD. PSD was only created in 2011 – that is why we do not have PSD's ideological scores for this period in the plot. Given that the party was formed mainly by dissidents from the right-wing parties PFL and PP, we consider PSD as rightist. Many of these changed their name since. PMDB became MDB, PFL became DEM, and PPS became CIDA.

[12] Democratic Accountability Linkages Survey. https://sites.duke.edu/democracylinkage/.

[13] www.lapopsurveys.org.

Figure 2: Left-Right Categorization of Brazilian Parties



All parties below (above) the line are categorized as Left-wing (Right-wing) for the purpose of this empirical application. DALP data is from 2008; the "Power and Zucco" data shown here is from 2005. PSD became the third largest party in Brazil after its creation, in 2011, by dissidents of other right-wing parties (namely PP and PFL).

Samuels and Zucco Jr. (2014) also show that partisanship in Brazil is, in fact, meaningful to voters, and that party identification shapes voter attitudes even in this political environment. They focus on PT and PSDB, known for being more 'programmatic' than their counterparts. They argue that "even though Brazil's two main parties have converged on the political center, agree on many of the issues, and have allied with a confusing array of parties, party labels for the PT and PSDB have the same effects scholars find for parties in older democracies."

Accordingly, in our empirical application, we conservatively classify parties in these two L-R groups. That said, this categorization warrants two additional comments. First, parties within a group are not necessarily cohesive in other dimensions, especially the Right (Power and Rodrigues-Silveira, 2018). For example, there is heterogeneity in *modus operandi* (some are highly clientelistic), origins, or "rent-seeking" behavior (some are more prone to alliances with the Left). We acknowledge that any idiosyncratic source of party heterogeneity could also influence both candidate selection and local spending. Unfortunately, it is beyond the scope of this project to precisely examine the effects of alternative party categorizations. Nevertheless, the appendix provides evidence that our results are robust to slight variations in these L-R groups (Table E.5), and nonexistent for alternative party categorizations that are uncorrelated with the L-R scale like small vs. large parties (Table E.3).

Second, these surveys reflect national ideological positions that, although consistent, might not trickle down to local races. In fact, the conventional wisdom in Brazil discounts the role of left-right ideology in municipal elections in favor of other issues such as dynastic rivalries (Boas, Hidalgo, and Melo, 2019).

In response, we note that our theoretical framework takes into account the possibility that ideological brands are not very meaningful for parties and voters. Our primary hypothesis should *only* hold when they do. Thus, the relevance or irrelevance of national brands for local politics in Brazil is a dispute that our empirical results can help adjudicate. However, before delving in the results, we look at the pattern of local mayoral coalitions to provide at least some preliminary evidence that national brands drive local politics to a certain extent.

Given the high number of parties in Brazil, nearly every mayor is supported by a large, pre-electoral coalition averaging more than 6 parties. In fact, even the 11 parties above are more likely to formally support a candidate from a competing party than to run their own. In this context, the ideological consistency of these alliances can reflect the degree of influence of party brands in mayoral races. The data from 2004-2016 show that, although "inconsistent" alliances often exist (e.g. a Left party supporting a Rightist mayoral candidate), parties were 3x times more likely to support a candidate from the same ideological group, or to refrain from any support, than to rally behind a candidate with opposing ideology as shown in Figure E.4. This statistic is even stronger if we limit the data to the largest parties only, which suggests that party brands are highly informative of local leadership's behavior.

Candidate selection and policymaking in municipalities

The Brazilian party system in fairly decentralized, which gives the municipal party branches ample control over the nomination of mayoral candidates.[14] The nomination can be a highly competitive process that mobilizes a relatively high share of the population: even though many voters do not identify with a party in surveys, the party membership rate in Brazil is among the highest across democracies (10+% of voters). Recruitment is also highly concentrated in the year before local elections, often as a display of electoral strength by mayoral hopefuls (Frey, 2020). These pre-candidates not only compete with other partisans for the candidacy, but also with politicians from parties that are potential coalition members. This pre-electoral competition is emphasized by the following press reports from Pontal do Araguaia (MT), Batayporã (MS), and Sorocaba (SP); where local parties/coalitions even rely on polls to find the optimal candidate.[15]

Within this decentralized system, why would local party branches implement policies aligned with national party preferences? The answer lies in the quid-pro-quo between mayors and other partisan politicians that characterizes Brazilian politics. Mayors play an important role in supporting the vertical strength of their parties, as they have ample control over the distribution of public goods: in Brazil, municipal administrations implement the bulk of spending in the areas of health, education, and infrastructure. Not surprisingly, mayors can effectively help (or hurt) their parties in higher elections (Feierherd, 2020;

[14] State/national leadership could actively interfere in the nomination process, but only in larger or more strategic municipalities – see for example Camaçari (BA) in the introduction.

[15] See the links in: https://bit.ly/2HuZ86e, https://bit.ly/3lafqzv, and https://bit.ly/3meDhis; respectively.

Novaes, 2018). On the other hand, mayors also depend on party leaderships to obtain budget resources.[16] Although the bulk of these funds comes in the form of non-discretionary federal transfers, congress members have (limited) access to budget amendments that can be targeted to municipalities in a way that is electorally efficient for parties. Many local electoral campaigns are also dependent on the party coffers, and from corporate donations obtained with the influence of the leadership,[17] and many mayors sustain careers within the party after leaving office.

Data and empirical design

We use data from the municipal elections of 2004, 2008, 2012, and 2016; and consider only races between the largest parties in the country – Figure 2 shows both the classification and the number of mayorships won by each party in the period.[18]

Our measure of pro-poor policy implemented by mayors is the share of the four-year municipal budget invested in the following four categories: health, sanitation, education, and housing.[19] In Appendix C, we show results for an alternative measure of pro-poor policy, based on the campaign proposals of each mayoral candidate released in advance of the 2012 and 2016 elections.[20] As mentioned before, our measure of descriptive representation is based on the education level of mayoral candidates. Unlike the main policy variable, here we observe the education for both candidates. The variable is thus defined as the difference between the education of winner and loser in the election. Education is measured on a scale of 1 to 8, the lowest level meaning that the candidate is illiterate, and the highest that she has at least a 4-year bachelor's degree.[21] In Table E.7, we show that the results are robust to alternative specifications where we use only the education of the winner, or code education as a dummy that indicates whether or not candidates have a bachelor's degree.

This measure is an attractive proxy for descriptive representation of poor voters for three reasons. First, it is easily obtainable, verifiable, and not open to interpretation, as for example is race in Brazil (Bueno and Dunning, 2017). Second, education has been shown to be highly correlated with socioeconomic status within countries (Krueger and Lindahl, 2001). Third, our theory does not require a perfect correlation, i.e., we do not expect every less educated politician to be poor, or vice versa. In fact, the

[16] Local taxes play only a very minor role in financing such investments.

[17] Public campaign funds are allocated to parties according to their congressional seats. Corporate donations were allowed in Brazil until 2012.

[18] Data on both the profile of candidates and election results was obtained from the Superior Electoral Court (TSE).

[19] For 2016, the measure only includes 3-years of spending, given that the 2020 data was not yet released. In Brazil, although public health and education services are formally 'universal', they are effectively used only by the poor – most middle and upper-class citizens use private alternatives. In Table E.7, we show that the results are robust to alternative specifications where we either subtract public security spending (a salient Right-wing policy) or add expenditures on social assistance. In Table E.2, we also show that the past values of this variable are balanced at the discontinuity. The breakdown of budget expenses for Brazilian municipalities was obtained from the National Treasury.

[20] This appendix describes the construction of this alternative variable, and shows the estimation results in Table C.1 – the empirical findings are consistent with the main results.

[21] This is how TSE categorizes the education of candidates. See the full scale in Figure E.2.

idea here is to measure the candidate's ability to descriptively *identify* herself with the lower classes of the population, as opposed to actually *be* poor. Many of these candidates are self-made entrepreneurs that grew up in poverty, and thus identify with the poor on their "humble origins" and lack of education, even if later they became successful.[22] Thus, all that our theory requires is that poor voters, when in the presence of a less educated candidate, perceive him to be "one of them."

Regression discontinuity design

The empirical evidence that supports our hypotheses comes from comparing both the policies and education of mayors elected by Right and Left-wing parties in Brazil. Our main explanatory variable is a dummy that indicates whether the elected mayor belongs to a Right-wing party, following Figure 2. A simple comparison of our outcomes between Right and Left-wing mayors is likely to be biased by unobserved municipal characteristics that either influence policies or are correlated with the education of the candidates that run and/or win elections. We address this problem with a regression discontinuity design (RDD) that compares only municipalities where a Right-wing party won (or lost) to a Left-wing one by a close margin.

For the policy variable, the RDD estimates represent the local treatment effect of electing a Right-wing mayor, precisely identified for a municipality were the margin of victory in the election was zero. However, our estimates for the education outcome cannot be interpreted as an *effect* of electing a Rightist politician, given that the nominations happen before elections. Instead, they should be interpreted as the correlation between education and the winner's party ideology.

Nevertheless, there are benefits from also using the RDD to estimate this correlation. First, by using an empirical approach consistent with the one used to identify the policy treatment effect, and a comparable sample, the RDD allows to precisely connect both results, as required by our theory. Second, the RDD is a very transparent way to show that this empirical pattern is not driven by a potential correlation between ideology and other observed variables, including other characteristics of candidates. Accordingly, Table E.1 shows the balance around the discontinuity of pre-determined or fixed covariates. In addition, we also show that the observed relationship between partisanship, education and poverty in Brazil is robust to alternative empirical approaches, such as OLS (cross-section) estimation and panel analysis, and not driven by the RDD assumptions. Figure E.1 and Table E.9 in the appendix show the results of these empirical strategies.

We provide estimates for two sub-samples with municipalities with poverty rate above and below the median.[23] Municipal poverty is measured by the share of poor families, estimated by the Ministry of

[22] Former president Lula, for example, was by no means poor at the time of his presidential runs. Nevertheless, he often used his lack of formal education to vouch for his ability to be in touch with the populace.

[23] The appendix (Table E.6) shows that the results are robust to the choice of poverty cutoff. In Table E.12, we show that they are also robust to non-binary measures of poverty, and a different definition of the poverty variable that uses the municipal Human Development Index.

Social Development (MDS).[24] The main estimating equation is:

$$y_{mt} = \beta_0 + \beta_1 R_{mt} + \beta_2 W_{mt} + \beta_3 R_{mt} W_{mt} + \Big(\beta_4 + \beta_5 R_{mt} + \beta_6 W_{mt} + \beta_7 R_{mt} W_{mt}\Big) M_{mt} + \delta_t + \theta_{mt} + \xi_{mt} \quad (8)$$

where outcome y_{mt} for municipality m in period t is regressed on the Right-wing dummy R_{mt}, and on the dummy that indicates whether the municipality is in the low poverty group (W_{mt}). The margin of victory is the difference in the vote share between the winner and runner-up (M_{mt}), δ_t are election fixed-effects, and θ_{mt} is a vector of pre-treatment covariates and other candidate characteristics.[25] Accordingly, β_1 is the effect of having a Right-wing mayor in a high-poverty municipality, while $\beta_1 + \beta_3$ is the effect in a low-poverty one.

Results

Table 1 shows the RDD estimates from equation 8 for different bandwidths.[26] Robustness to the exclusion of covariates and polynomial choice are shown in Table E.4.

Table 1: Mayor's partisanship, education, and pro-poor spending

Dependent Variable:	Pro-poor spending as % of budget			Education Gap (winner minus loser)		
	(1)	(2)	(3)	(4)	(5)	(6)
High Poverty	1.061	0.731	0.348	-0.802*	-0.753*	-0.681*
	(0.828)	(0.739)	(0.679)	(0.335)	(0.290)	(0.260)
Pre-treatment baseline	59.659	59.630	59.579	-0.033	-0.021	-0.003
Low Poverty	-1.904*	-2.026*	-1.923*	0.519†	0.250	0.104
	(0.858)	(0.762)	(0.695)	(0.281)	(0.249)	(0.226)
Pre-treatment baseline	50.307	50.352	50.421	0.164	0.136	0.117
Bandwidth	3.97	5.29	6.61	4.05	5.40	6.76
Observations	1544	2026	2464	1566	2061	2504
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. 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. The coefficients come from the estimation of equation 8.

The pattern that emerges from the estimation is in line with **Hypothesis 1**: in high-poverty locations, Right and Left implement indistinguishable policies, i.e., both parties allocate nearly 60% of the budget

[24] This is the base for several federal government benefits including *Bolsa Família*.

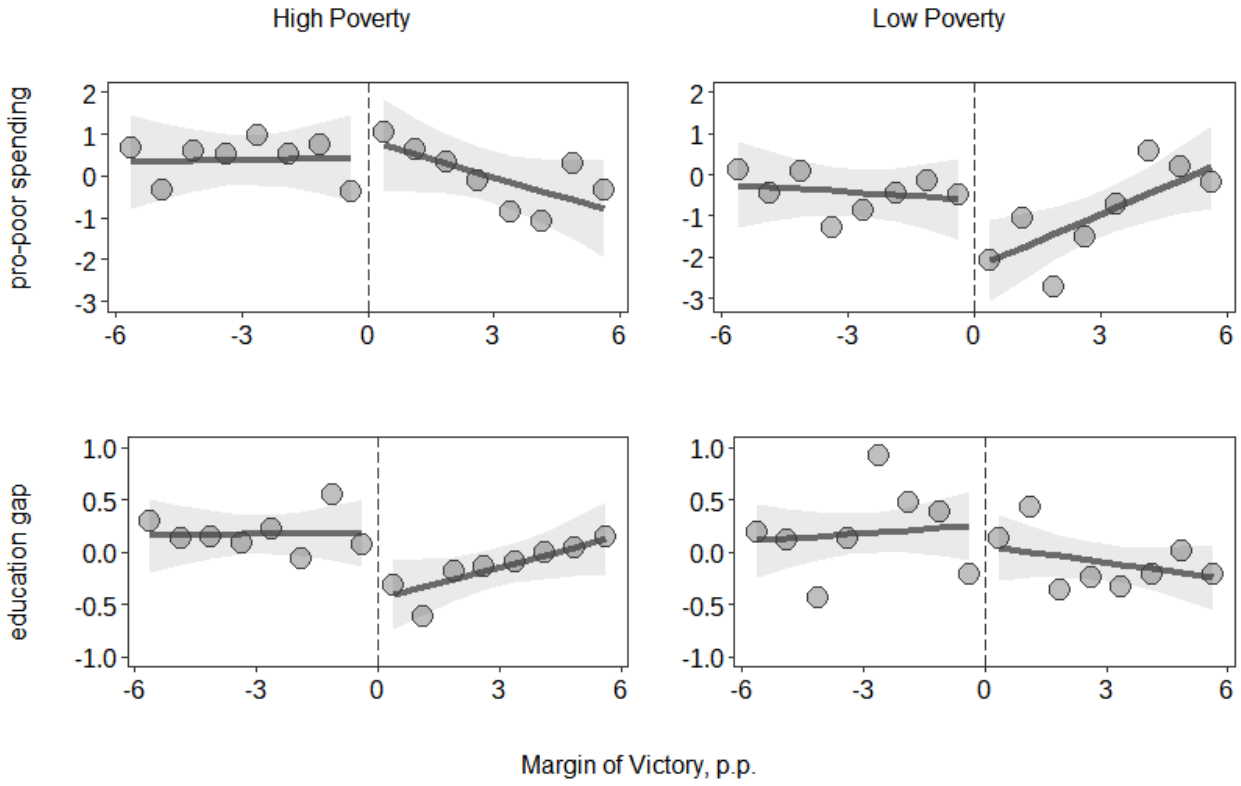
[25] As it is usual in RD designs, covariates are often included to reduce the variance of the estimated coefficients. We include all covariates listed in Table E.1, which also shows that these variables are balanced at the discontinuity. Table E.4 shows robustness to the exclusion of covariates.

[26] Bandwidths are estimated using the algorithm in Calonico, Cattaneo, and Titiunik (2014).

to pro-poor spending. However, for the same sample, Right-wing winners are on average less educated than their Leftist opponents by nearly one point on the 1-8 scale. As municipalities become less poor, the policy effect becomes significant: Right-wing incumbents spend 2.0pp less than Left-wing ones on the poor, from a baseline of 50% of the budget. However when municipalities are less poor, both parties nominate politicians with virtually indistinguishable education.

The overall pattern in the results is illustrated in Figure 3, and is robust to slightly different definitions of Left-Right groups (Table E.5). In fact, our baseline categorization provides estimates that are more conservative than these alternative specifications. For example, when we restrict our comparison to PT vs. Right-wing parties only, both the policy and education coefficients are higher in magnitude, by 17% and 70%; respectively.

Figure 3: RDD effects by variable and poverty level



For every plot, the Right (Left) side shows the municipalities where a Right-wing (Left-wing) party has won the mayoral election. The lines are a linear fit, and the points represent the average outcome for the corresponding level of margin of victory in each bin. For presentation, the outcomes are normalized by the average of the subsample.

In addition to being highly consistent with the theory, the policy result in low poverty areas provides an additional insight to the specific Brazilian case: parties consistently implement policies in line with their ideological brands, i.e., national ideologies are informative of local policies.[27] This result is also

[27]We also stress that, if the mechanism we propose is true, we would never observe policy differentiation in high poverty

present when we use the alternative policy variable based on the campaign proposals (Appendix C). This evidence provides strong support to our theoretical assumption that party brands play a significant role in many local races, especially in a political context where voters have been shown to *recognize* the policies and performance of mayors, and to *punish/reward* parties accordingly (Boas, Hidalgo, and Toral, 2020; Feierherd, 2020; Ferraz and Finan, 2008; Klašnja and Titiunik, 2017).

As for the education gap in high poverty areas, we show in the appendix that it is primarily driven by cases where the Left fields college-educated candidates – Table E.10. In other words, the descriptive representation gains accrued by the Right come primarily from the comparison between candidates with and without a bachelor's degree,[28] in line with the representation gap depicted in Figure B.1. Table E.7 also shows that the education result remains robust when the variable is built with dummies that indicate whether the candidates have a bachelor's degree. Finally, the OLS and panel results in Figure E.1 and Table E.9 are also consistent with our findings.

Heterogeneity by the degree of local ideological competition

Our theory accommodates the possibility that party brands might be less informative of local policies, which could often be the case in Brazil. In such locations, **Hypothesis 2** predicts that the pattern depicted in Table 1 should be weaker, or even nonexistent. In order to check whether our results exhibit such heterogeneity, we split the overall sample into high and low ideological alignment groups, based on the profile of the party coalitions of Left and Right-wing candidates in each municipal election. The intuition is that ideological competition is more salient for parties and voters in races where these coalitions are highly aligned with the national L-R divide.[29] Table 2 shows the results of this exercise. Columns (1) and (4) run our main specifications on the full sample. In columns (2) and (5), where we run the analysis on the *low alignment* sample, all relevant coefficients are statistically insignificant as expected. In columns (3) and (6), we show that the results in the *high alignment* sample are both significant and higher in magnitude.

Table E.5, columns (3,6), provide additional evidence in favor of Hypothesis 2. We show results for a subsample with races between parties at both extremes of the L-R scale (i.e. PT/PSB vs. PFL/PP/PSD/PL). Since these parties have more extreme ideological positions, we also expect the effects to be stronger for this subsample, which is indeed the case.

Finally, in the appendix we show two additional tests of the overall framework. In Table E.10 columns (3,4) show that the education gap results are stronger in municipality-years where the previous mayor was Leftist. This is in support of the mechanism that the Right nominates less educated candidates to

areas, even if party brands are relevant and informative.

[28] As opposed to with and without high-school, for example. Figure E.3 shows that, while the Left fields more candidates with a bachelor's degree, the Right has more candidates in every other education category.

[29] More specifically, for every race, we compute the share of the 11 main parties that support a mayoral candidate from the same L-R group. This is our municipality-election measure of alignment, which has maximum value of 1, when all parties support aligned candidates, and a median value of 0.45, used to divide the sample.

Table 2: Effect Heterogeneity, by the local level of ideological alignment

Dependent Variable:	Pro-poor spending as % of budget			Education Gap (winner minus loser)		
	Full Sample	Low Align.	High Align.	Full Sample	Low Align.	High Align.
High Poverty	0.731	0.704	0.986	-0.753*	-0.472	-0.978*
	(0.739)	(0.984)	(1.047)	(0.290)	(0.379)	(0.424)
Low Poverty	-2.026*	-1.126	-2.652*	0.250	0.180	0.263
	(0.762)	(0.998)	(1.139)	(0.249)	(0.397)	(0.344)
Optimal Bandwidth	5.29	5.29	5.29	5.40	5.40	5.40
Observations	2026	969	1057	2061	984	1077

†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. The coefficients come from the estimation of equation 8. The sample split is described in the text.

reinforce their commitment in matching the Left's pro-poor policies. In this context, the Right's strategy should be more relevant where voters have not recently been exposed to the policies of a Right-wing mayor, and there is more uncertainty around the Right's commitment to policy shifts. In Appendix D (page D-11) we use the LAPOP survey to show that poor voters feel more represented by parties than non-poor voters only in high poverty municipalities that are governed by a less educated mayor.

ALTERNATIVE EXPLANATIONS

In this section we assess three alternative mechanisms that could be driving the observed candidate nomination pattern. First, there is no doubt that clientelism plays a significant role in Brazilian politics. Rather than ruling out the relevance of this practice, we show several pieces of evidence suggesting that clientelism is unlikely to drive the candidate selection pattern shown before. On the contrary, the evidence suggests that descriptive representation and clientelism are likely *substitute* electoral strategies in this context.

The nomination pattern we uncover in Brazil could be an artefact of clientelism as opposed to our explanation if (i) Right-wing parties are consistently better at it than the Left (e.g. either because they possess better broker networks or more resources); and (ii) less educated candidates have a competitive advantage in the practise. Under these assumptions, the Right selects less educated candidates to boost clientelism in poor areas, which would not be efficient for the Left.[30]

We assess this explanation with three empirical exercises. First, we examine the correlation between ideology and education gap only in races between the two most "programmatic" large parties on each

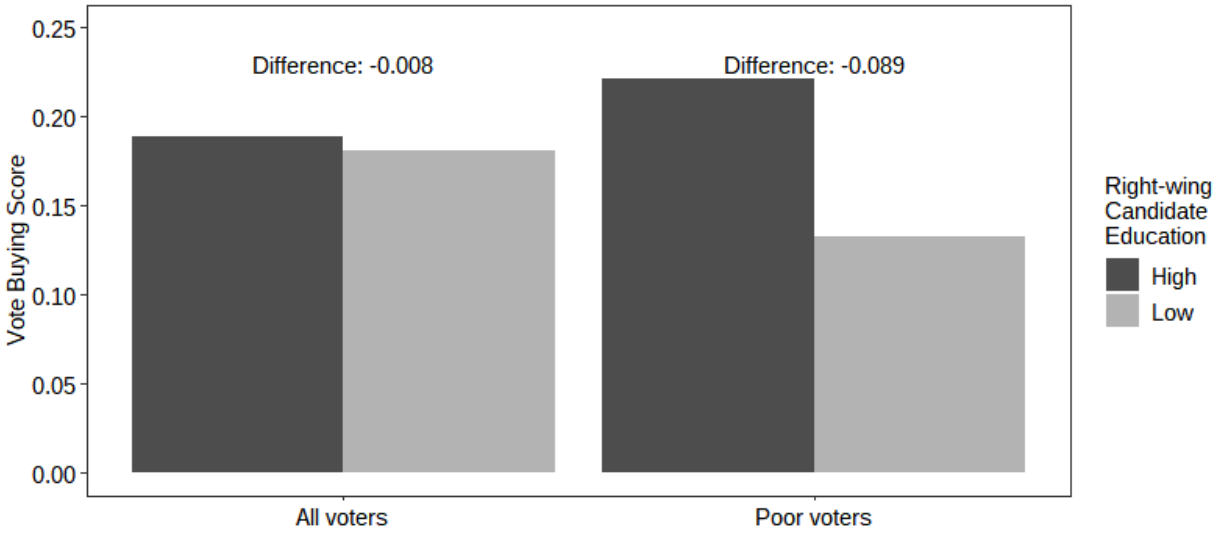
[30]We emphasize that this narrative does not threaten our results on programmatic differentiation in low poverty municipalities.

group, PT (Left) and PSDB (Right). Table E.10 shows not only that the political selection mechanism still holds for this less clientelistic subset, but that the magnitude of the correlation is more than double the one of the full sample. This could suggest that rather than being an alternative explanation to our results clientelism *attenuates* the Right's need to appeal to descriptive representation to become competitive among poor voters.

Second, we show that there is no evidence that Right-wing parties spend more than the Left in mayoral campaigns, both in high and low poverty areas. Although this is not a direct measure of clientelistic capacity, it is a good proxy for the parties' ability to buy votes during the election period. In Table E.8 we show that our RDD framework produces statistically insignificant estimates when the outcome variable is the gap between the spending of the Right and Left candidates.

Third, we use the 2010 LAPOP survey to elicit voters' perceptions on vote buying, and how they vary according to the education level of the Right-wing mayoral candidate in the municipality.[31] Vote buying is coded as a dummy that indicates whether the voter was offered to sell their vote. We focus on the municipalities with higher than median poverty level, where on average 19% of voters were offered to sell their vote. Figure 4 shows that the presence of a more educated Right-wing candidate in the 2008 race did not trigger higher reporting of vote buying by voters (both differences are statistically insignificant, and negative).

Figure 4: Self-reported vote buying and the education of Right-wing candidates



†p<0.1, *p<0.05. Data includes 983 voters in 25 municipalities. High education refers to the group of candidates with more than a secondary degree.

A second alternative explanation for the education gap result is that Right and Left-wing parties face

[31] The 2010 wave of this survey had the following question: In recent years and during electoral campaigns, did any candidate or any member of a political party offered you something like a favor, food, or any other benefit or good in exchange for your vote or support?

a different pool of candidates in high-poverty municipalities. For example, the same nomination pattern would arise if the Left simply has better recruitment networks than the Right among the highly educated in these areas. We assess this narrative using the education of incumbent councilors for each party at the moment of the election, given that councilors come precisely from the same pool as mayors (party membership rolls), and the council can be a stepping stone to a mayoral candidacy.[32] Here the outcome variable is the difference in the number of highly educated councilors[33] between the mayor's party and the runner-up's (similar to the construction of the education gap described in page 13). Table 3 shows the RDD estimates for the high-poverty sample. Columns denoted by (A) consider all councilors, and columns denoted by (B) only consider the two most voted councilors in each municipality.

Table 3: Education of partisan council members

Dep. Variable:	Councilors with secondary education		Councilors with a bachelor's degree	
	(A)	(B)	(A)	(B)
RDD estimate	0.134	0.061	0.019	0.041
	(0.185)	(0.075)	(0.103)	(0.051)
Bandwidth	5.52	5.43	5.05	3.98
Observations	1082	1070	998	799

†p<0.1, *p<0.05. The dependent variable is always the gap between the outcomes for winner and loser. Standard errors are clustered by municipality (parenthesis). The RDD estimates are the difference in outcomes between municipalities with Right and Left-wing mayors, at the discontinuity, for the high-poverty sample only. Columns denoted by (A) and (B) consider all councilors, or only the 2 most voted members in the council; respectively.

As it is evident, the coefficients are statistically indistinguishable at the discontinuity. If anything, a victory of a Right-wing mayor is (weakly) correlated with the Right-wing party also having a larger pool of educated councilors. Table E.11 and Figure E.5 show similar findings in both a panel analysis and in the raw data. Together, these results suggest that it is unlikely that the nomination pattern in poor municipalities is being driven by the candidate pools available to the parties.

Finally, the selection pattern could arise if the recruitment of less educated candidates is particularly costly to the Left. We provide an extensive discussion in Appendix B about the potential costs of recruiting less educated candidates. In short, our multiple measures of administrative performance and mobilization capacity strongly suggest that uneducated mayors are indeed costly to parties, but also that the cost is uniform across ideological groups.

[32] We must note here that Right-wing parties, on average, have much larger party memberships than Leftist ones in most Brazilian municipalities, already making this alternative argument less likely.

[33] We use two different specifications for "high education". The first includes at least secondary education (levels 6, 7 and 8 of our scale), and other that considers only a college degree (level 8).

CONCLUSION

This paper uses a regression discontinuity design in Brazilian municipal elections to uncover a puzzling empirical regularity: in high poverty municipalities, Right and Left-wing mayors devote a similar share of budget to pro-poor policies, but Right-wing candidates are relatively less educated. In low-poverty areas the pattern is reversed: candidate profiles are similar, but Right-wing mayors spend less on the poor. We interpret these findings within a theory where the Left always chooses pro-poor policies in line with party ideals. The Right, on the other hand, can only credibly promise pro-poor policies when it nominates candidates that are descriptively closer to the poor. This strategy allows the Right to become competitive in poorer regions, despite its less popular brand. In less poor areas, where more voters are aligned with the Right's preferences, this strategy is unnecessary.

These findings have significant implications for the future study of party strategies in developing democracies, where most research attributes the success of elite-driven parties to the erosion of programmatic brands, and the prevalence of non-policy politician-voter linkages. While these findings in no way imply that non-policy strategies such as clientelism are irrelevant, they suggest that their success may also depend on how they interact with programmatic differentiation.

Given the relative paucity of descriptive representation of the poor in the literature (Carnes and Lupu, 2015), our results suggest that, at least in Brazil, it is surprisingly Right-wing parties who more often capitalize on this shortage. More broadly, the results also imply that there is more to descriptive representation than its direct effects on substantive representation, which has been the focus of the bulk of this literature. Our article suggests that parties might also use strategic descriptive representation as a tool to convey their commitment to policy shifts.

Finally, the mechanism proposed here is context driven, where fielding less educated candidates is costly for parties, and there is a trade-off between this cost and the electoral benefits (direct or policy-driven) of descriptive representation. Otherwise, the nomination strategy uncovered here would be optimal for all parties.

REFERENCES

Ames, Barry and Amy Erica Smith. 2010. "Knowing Left from Right: Ideological Identification in Brazil, 2002-2006." *Journal of Politics in Latin America* 2 (3):3–38.

Boas, Taylor C., F. Daniel Hidalgo, and Marcus André Melo. 2019. "Norms versus Action: Why Voters Fail to Sanction Malfeasance in Brazil." *American Journal of Political Science* 63 (2):385–400.

Boas, Taylor C., F. Daniel Hidalgo, and Guillermo Toral. 2020. "Competence versus Priorities: Negative Electoral Responses to Education Quality in Brazil." *working paper* .

Bueno, Natália S. and Thad Dunning. 2017. "Race, Resources, and Representation: Evidence from Brazilian Politicians." *World Politics* 69 (2):327365.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82 (6):2295–2326.

Carnes, Nicholas and Noam Lupu. 2015. "Rethinking the comparative perspective on class and representation: Evidence from Latin America." *American Journal of Political Science* 59 (1):1–18.

———. 2016. "Do voters dislike working-class candidates? Voter biases and the descriptive underrepresentation of the working class." *American Political Science Review* 110 (4):832–844.

Dal Bo, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne. 2019. "Economic Losers and Political Winners: Sweden's Radical Right." *Mimeo* URL https://eml.berkeley.edu/~ffinan/Finan_SwedenDemocrats.pdf.

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.

Desai, Zuheir. 2020. "A Theory of Electoral Competition in Developing Democracies."

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.

Frey, Anderson. 2020. "Strategic Allocation of Irrevocable and Durable Benefits." *Mimeo* URL http://goo.gl/ucMR8E.

Gay, Claudine. 2002. "Spirals of trust? The effect of descriptive representation on the relationship between citizens and their government." *American Journal of Political Science* :717–732.

Hayes, Matthew and Matthew V Hibbing. 2017. "The symbolic benefits of descriptive and substantive representation." *Political Behavior* 39 (1):31–50.

Huber, John D. 2017. *Exclusion by Elections: Inequality, Ethnic Identity, and Democracy*. Cambridge Studies in Comparative Politics. Cambridge University Press.

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

Krueger, Alan B. and Mikael Lindahl. 2001. "Education for Growth: Why and for Whom?" *Journal of Economic Literature* 39 (4):1101–1136.

Lawless, Jennifer L. 2004. "Politics of Presence? Congresswomen and Symbolic Representation." *Political Research Quarterly* 57 (1):81–99.

Lupu, Noam. 2016. "Building Party Brands in Argentina and Brazil." In *Challenges of Party-Building in Latin America*, edited by Steven Levitsky, James Loxton, Brandon Van Dyck, and Jorge I. Dominguez, chap. 3. Oxford: Cambridge University Press, 76–99.

Mainwaring, Scott, Meneguello, Rachel Meneguello, and Timothy Power. 2000. "Conservative Parties, Democracy, and Economic Reform in Contemporary Brazil." In *Conservative Parties, the Right, and Democracy in Latin America*, edited by Kevin J. Middlebrook, chap. 6. The Johns Hopkins University Press, 164177.

Murillo, Maria Victoria and Ernesto Calvo. 2019. *Non-policy Politics: Richer Voters, Poorer Voters, and the Diversification of Electoral Strategies*. Cambridge University Press.

Nichter, Simeon. 2018. *Votes for Survival: Relational Clientelism in Latin America*. Cambridge Studies in Comparative Politics. Cambridge University Press.

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

Pettersson-Lidbom, Per. 2008. "Do Parties Matter for Economic Outcomes? A Regression-Discontinuity Approach." *Journal of the European Economic Association* 6 (5):1037–1056.

Power, Timothy J. and Cesar Zucco Jr. 2009. "Estimating Ideology of Brazilian Legislative Parties, 1990-2005: A Research Communication." *Latin American Research Review* 44 (1):218–246.

Power, Timothy J. and Rodrigo Rodrigues-Silveira. 2018. "The Political Right and Party Politics." In *Routledge Handbook of Brazilian Politics*, edited by Barry Ames. Taylor Francis Group, 251268.

Power, Timothy J. and Cesar Zucco Jr. 2012. "Elite Preferences in a Consolidating Democracy: The Brazilian Legislative Surveys, 19902009." *Latin American Politics and Society* 54 (4):1–27.

Samuels, David and Cesar Zucco Jr. 2014. "The Power of Partisanship in Brazil: Evidence from Survey Experiments." *American Journal of Political Science* 58 (1):212–225.

Samuels, David J. and Cesar Zucco Jr. 2018. *Partisans, Antipartisans, and Nonpartisans: Voting Behavior in Brazil*.

Shayo, Moses. 2009. "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution." *American Political Science Review* 103 (2):147174.

Thachil, Tariq. 2014. *Elite parties, poor voters: How Social services win votes in India*. Cambridge University Press.

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

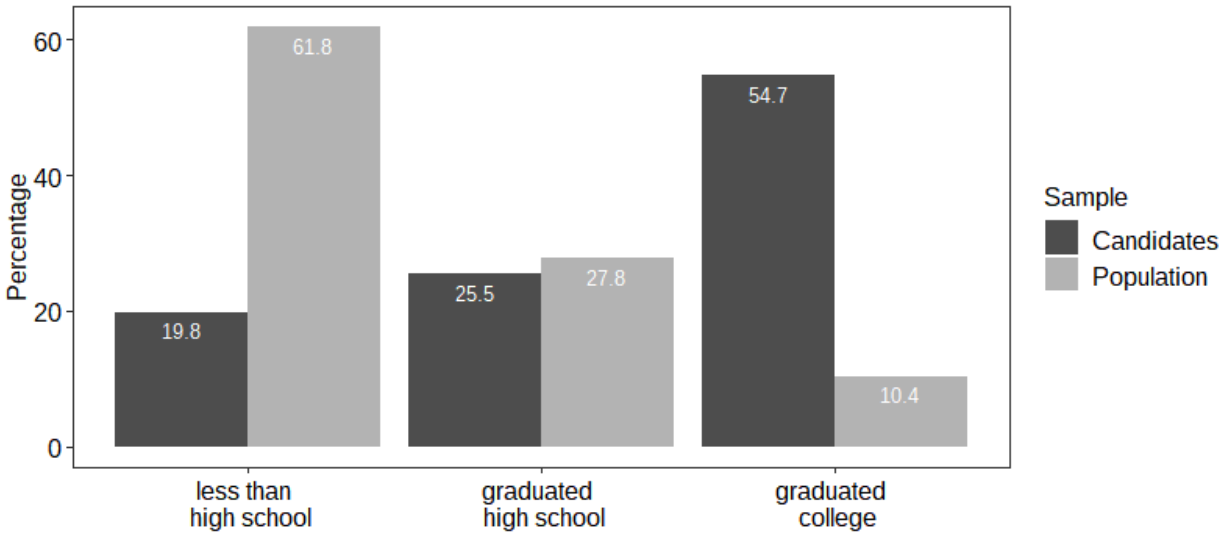
First, note that candidate selection does not happen in a vacuum and crucially depends on the candidate pool. Self-selection into politics is a well-documented phenomenon and shapes the candidate pools available to parties. For example, Dal Bó et al. (2017) examine the pattern of political selection 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 κ,[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 κ and could potentially be greater than κ.

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†	-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

†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.

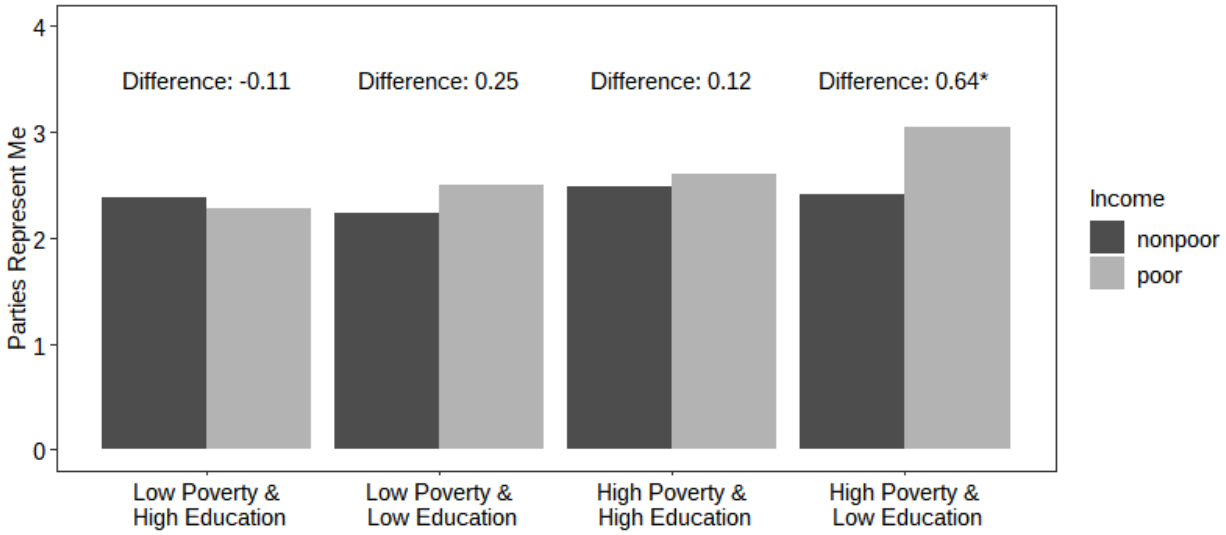
[37] PDF files available at https://divulgacandcontas.tse.jus.br/. Some files were missing, some were not in pdf format, and some were uploaded as low quality scans of paper documents that could not be converted to text. Thus, we recovered data for both the Left and Right-wing candidates for 83% of our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

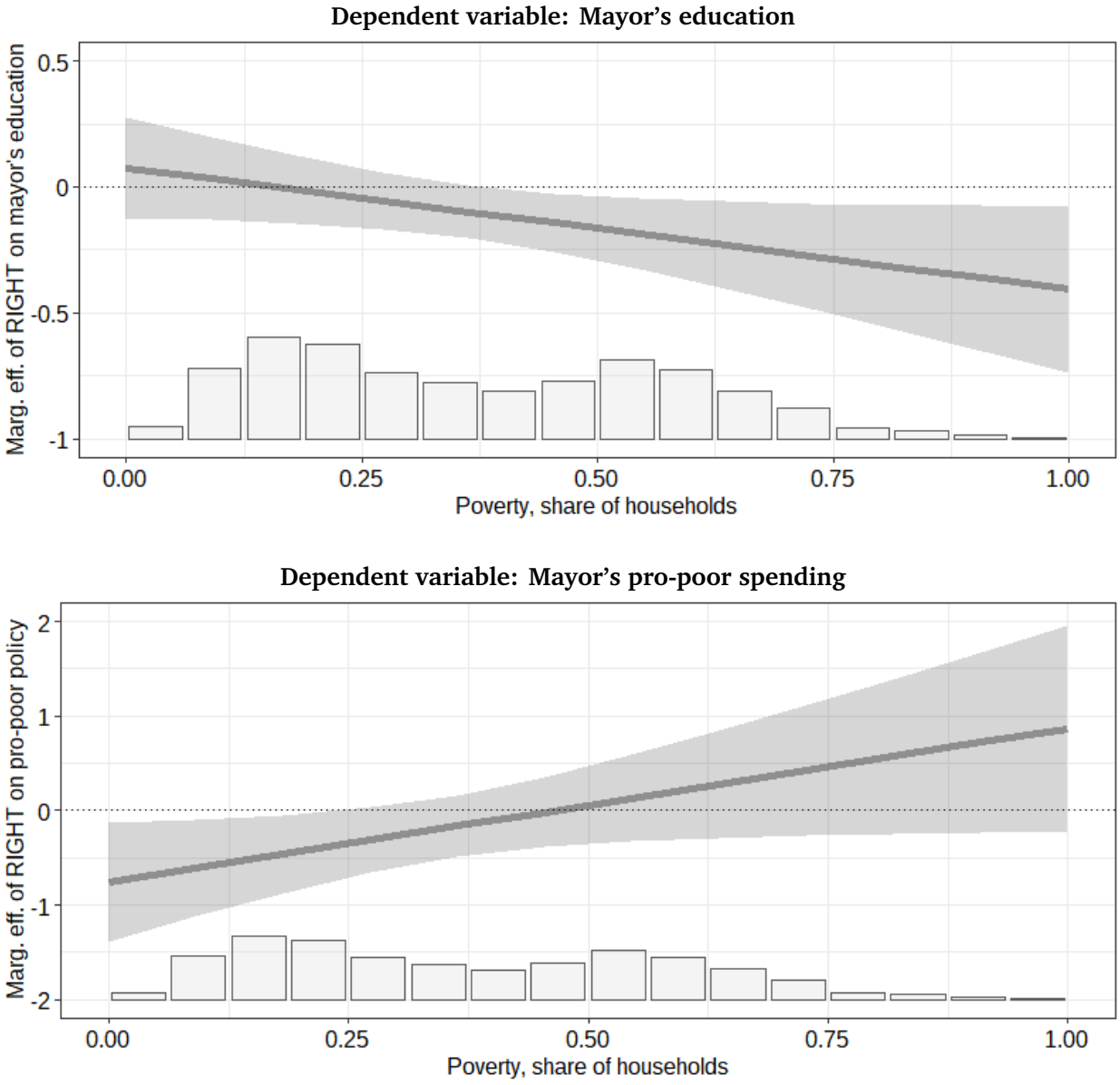
[43] In both waves, income is reported in categories. Poor households are the ones with monthly income below R$380 in 2008,

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).

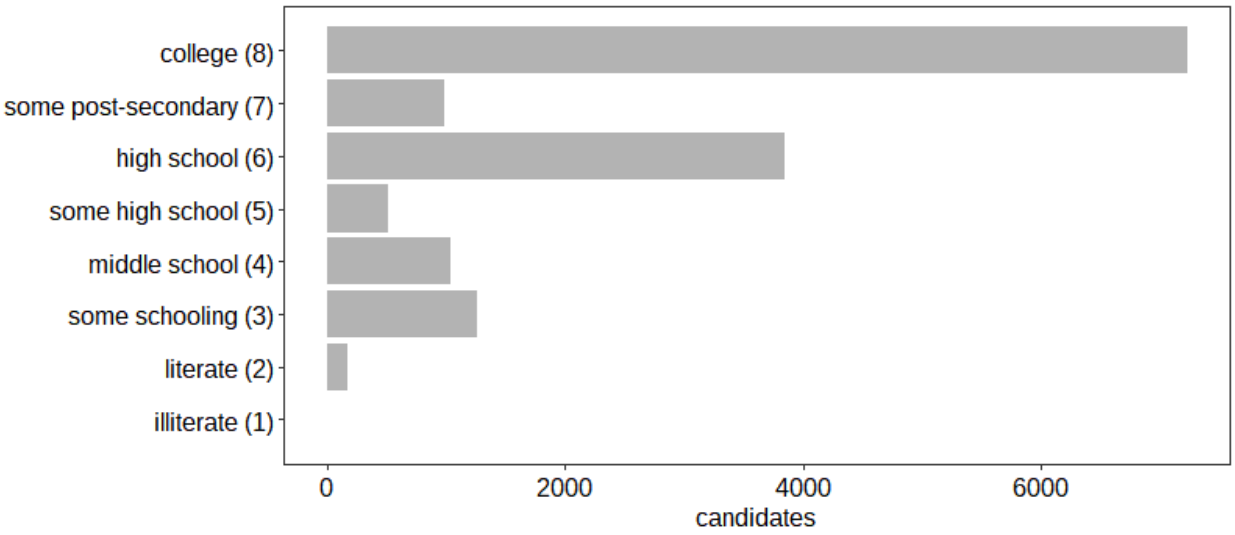
E TABLES AND FIGURES

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



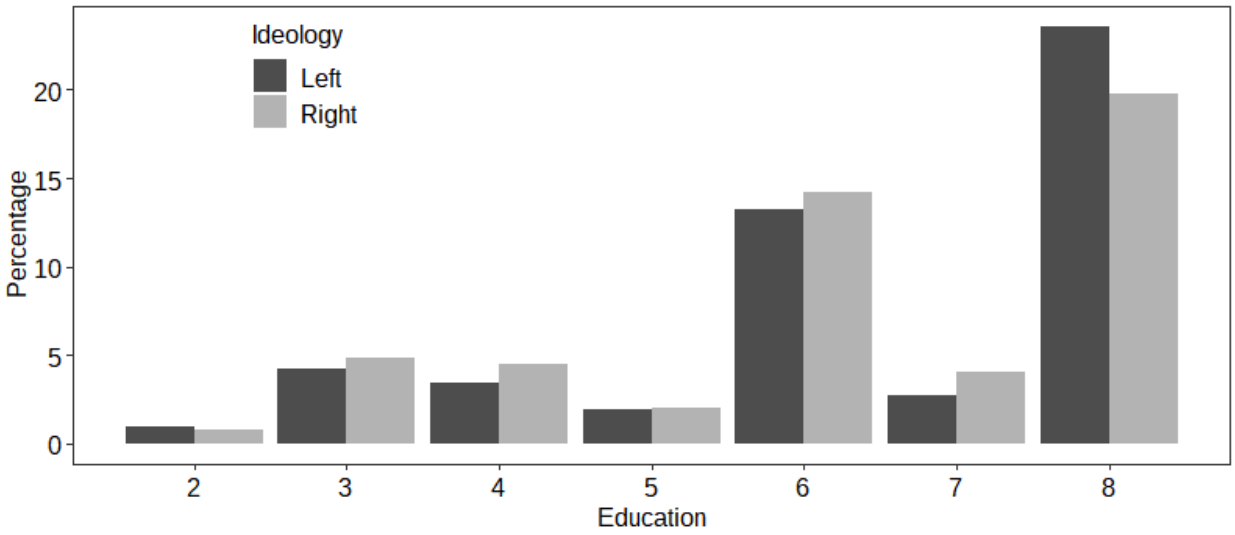
Data from 2004-2016. The lines show how the marginal effect of electing a Right-wing mayor (instead of a left-wing contender) changes as poverty increases in municipalities. The plot is based on a regression that includes fixed-effects for election and municipality, and pre-treatment, time-variant covariates. 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)



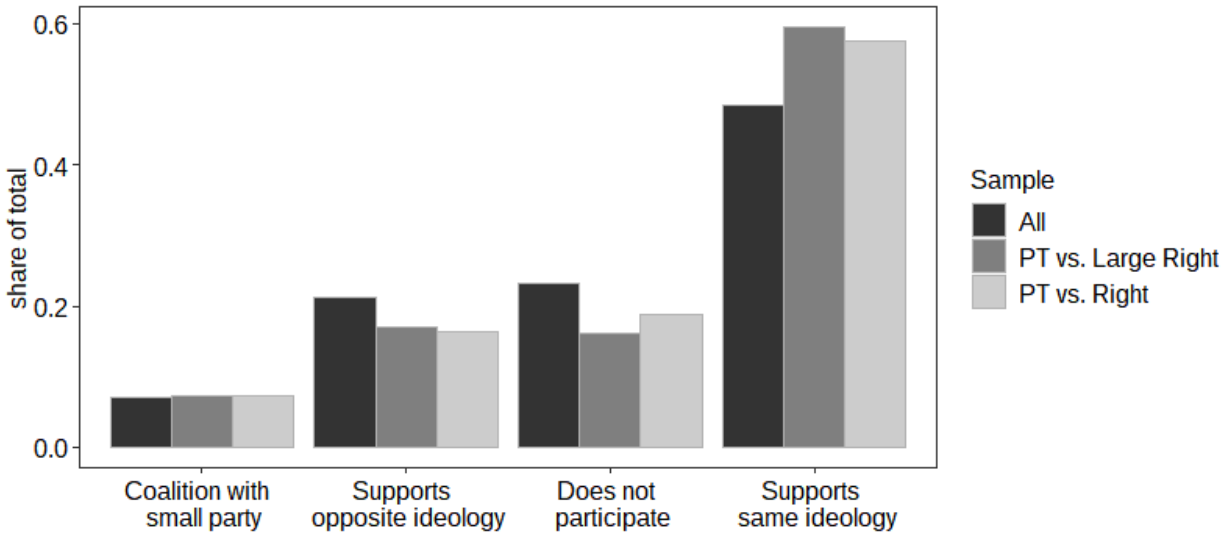
The bars show the number of candidates on each category.

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



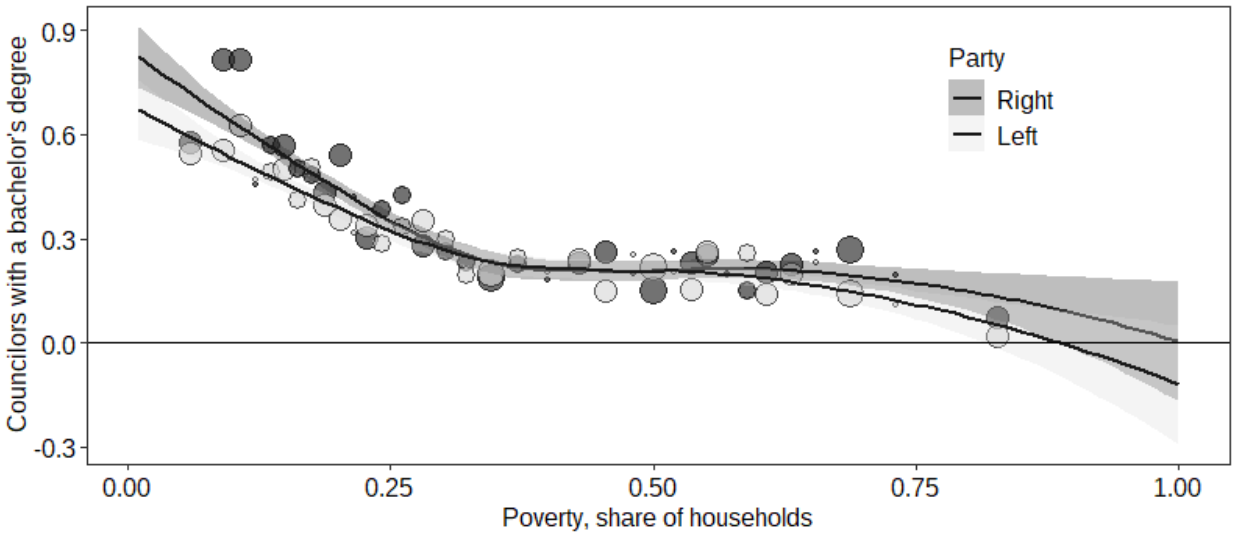
The bars show the percentage of candidates on each category. The description of these categories can be found in Figure E.2 above.

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



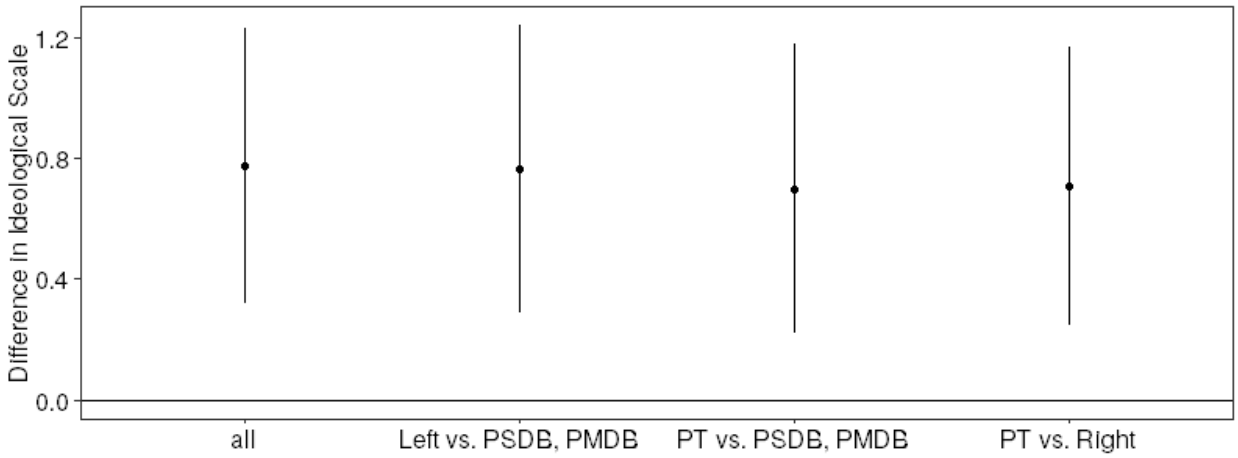
The bars show the share of the total cases. Large Right includes PMDB, PSDB, PFL, PP and PSD.

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



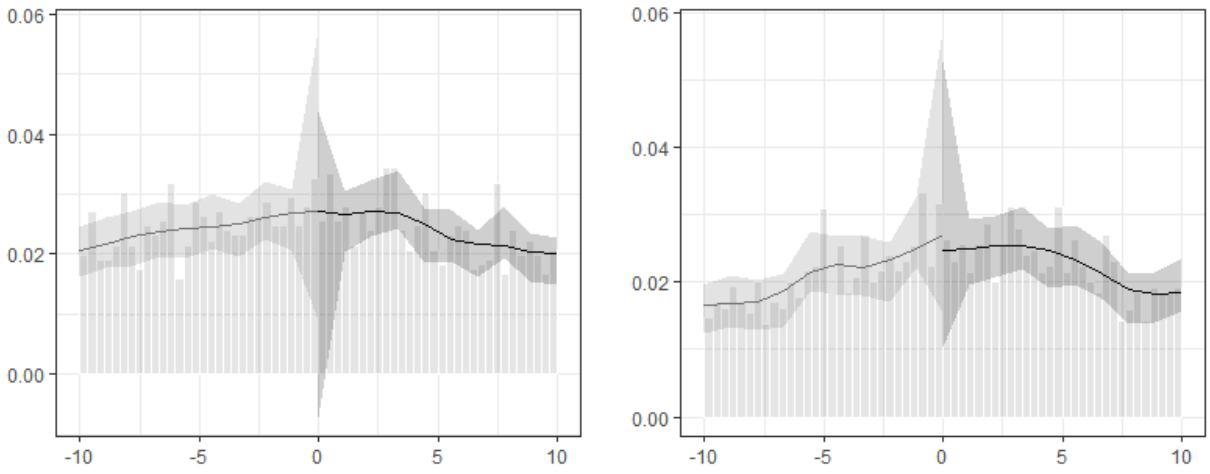
The plot shows the number of councilors with a bachelor's degree, immediately preceding each election in 2004-2016, for the Right and Left-wing parties. For each municipality, the count includes councilors from the top 2 parties in the mayoral election – only considering this paper's sample of races 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†	-0.880†	-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†	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.

†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†	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.

†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.

†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†	-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

†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†	5.166†	-1.082*	-1.025†	-1.954†
	(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†
	(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

†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 β_1 and β_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.