When politicians appeal broadly: The economic consequences of electoral rules in Brazil

Moya Chin *

September 30, 2019

Please click here for the latest version

Abstract

Electoral rules determine how voters’ preferences are aggregated and translated into political representation, and their design can lead to the election of representatives who represent broader or narrower constituencies. This project examines how single- and two-round elections in Brazil affect municipal mayoral races using a regression discontinuity design. Two-round elections use two rounds of voting to elect a winner, ensuring the winner obtains at least 50% of the vote. Theoretically, this provides incentives for candidates to secure a broader base of support. I provide evidence that, in two-round systems, candidates represent a more geographically diverse group of voters, more resources are allocated to public schools, and there is less variance in resources allocated to public schools across the municipality. I find evidence suggesting that strategic behavioral responses by candidates, rather than differential entry of candidates into races, explain a larger part of these effects. These findings suggest that two-round systems lead candidates to secure broader voter bases and subsequently exhibit less political favoritism when implementing policy.

*Harvard University, Department of Economics, mchin@g.harvard.edu. Sincere thanks to my advisors Melissa Dell, Nathan Nunn, Vincent Pons, and Gautam Rao for their extensive advice and guidance. I also thank Juan Pablo Chauvin, Siddharth George, Nir Hak, Tzachi Raz, Jonathan Roth, Juan Francisco Santini, Henrik Sigstad, Niharika Singh, and Clemence Tricaud. Roberto Acioly, Renata Dias, and Julia Orrico provided excellent research assistance. This project also benefited from numerous participants at the Development Lunch, Political Economy and Culture Lunch, and Labor and Public Finance Lunch at Harvard.
1. Introduction

There is growing consensus that institutions that are more inclusive lead to better outcomes\(^1\). This leads to the question of how to translate this into a political system. Even within democracies, there are many ways the political system can be structured. This paper studies how a particular political institution, the electoral rule, may lead to the election of representatives who represent broader or narrower groups of voters and affect the manner in which policies are implemented.

I take advantage of a unique rule in Brazil that assigns a municipality’s electoral rule based on a threshold of 200,000 registered voters. Municipalities below this threshold elect their mayor in a single-round election and municipalities above this threshold elect their mayor in a two-round election. This setting allows me to isolate the causal impact of the electoral rule by utilizing a regression discontinuity framework that compares municipalities close to the threshold. This paper provides empirical evidence that politicians elected under two-round systems secure geographically broader bases of support, provide more resources to public schools, and allocate these resources more equitably.

In a single-round system, voters vote once and the candidate with the most votes wins. In a two-round system, voters first vote and, if no candidate receives a majority, vote a second time between the top two candidates\(^2\). This distinction serves two main purposes. First, two-round systems require winners to attain a vote share above 50%. Second, the existence of a second round effectively limits the number of candidates (Lizzeri and Persico, 2005). If any one candidate is unlikely to attain a majority, the second round by construction limits the number of candidates. If one candidate emerges as a clear front-runner, she effectively only needs to be concerned with the runner-up, who threatens either a victory in the first round or will be the likely opponent in the second round. Because of these distinctions, two-round systems have been argued to incentivize candidates to secure a broader base of support and legitimize the winner’s position once elected (Bouton, 2013; Bouton and Gratton, 2015)\(^3\).

The intuition behind this is that the rules imposed in a two-round system make it more difficult for politicians to win with policies that appeal to a narrow group of voters.

---

\(^1\)This is most strongly argued in the literature on elite capture, which emphasizes how groups holding disproportionate power undermine reforms that benefit broader constituencies. See Acemoglu and Robinson (2008), Acemoglu et al. (2014), and Bardhan and Mookherjee (2006).

\(^2\)This is the case in Brazil. In other countries, variants of the two-round system are implemented.

\(^3\)While not a focus of this paper, there is also a large literature arguing that two-round systems allow voters to vote more sincerely in the first round (see Bouton et al., 2019 for a review) and to better communicate their policy preferences to candidates (Piketty, 2000).
If politicians must commit at some level to their campaign promises, the policies that politicians offer in order to win the election may translate into economic consequences. This paper aims to test the hypothesis that electoral rules can affect the public goods that are provided and the manner in which these are allocated across the electorate. When electoral rules make it more difficult for politicians to win with a narrow constituency, this reduces their incentives to provide public goods supported by a narrower constituency once in office.

To guide the empirical analysis, I propose a simple model of electoral competition to explain why two-round systems may incentivize politicians to engage broader groups of voters and allocate resources more equally across different groups. I adapt a standard probabilistic voting model but allow candidates to offer policy proposals that target government resources to specific observable and distinct localities within a municipality, as in Genicot et al. (2018). These policy proposals are announced prior to the election and, in a two-round system, are binding between rounds. In these elections, the top two candidates must contend with a small, third candidate, who is non-strategic and allocates all resources to a single locality.

In the model, because the locality is dominated by a third candidate, the top two candidates in a single-round system face low incentives to allocate resources to that locality as the marginal return to doing so is low. Consequently, candidates’ optimal strategy is to allocate few resources to that locality. In contrast, the existence of a second round raises the marginal return to allocating resources to that locality. Even in the first round, the top two candidates must look ahead to the possibility of a second round, where the third candidate will not be present, and appeal to these voters. As a result, in a two-round system, the top candidates allocate more resources to that locality even if they are unlikely to attract those votes in the first round.

Two-round systems are the most widely used system in democratic presidential elections: 64.0% of elections used the two-round system between 2000 and 2010 (Bormann and Golder, 2013 and Figure 1). Together, single- and two-round systems account for 86.7% of elections in this period. Despite the prevalence of single- and two-round systems, causally identifying the impacts on economic and political outcomes is difficult. While countries employ a variety of electoral rules, these rules are often chosen with economic or political interests in mind or implemented in conjunction with other reforms (Cusack et al., 2007). By utilizing a regression discontinuity framework that compares municipalities close to the threshold, this study aims to provide causal, empirical evidence that these rules have electoral consequences and impact the policies that politicians implement once in office.

I obtain two main empirical results. First, candidates running in two-round elections
receive broader geographical support. Using vote counts that each candidate receives at each polling station, I calculate (i) indices of voter concentration, to quantify the overall level at which voters for specific candidates are geographically concentrated or dispersed, and (ii) the standard deviation of candidates’ vote shares, to quantify a candidate level of geographic dispersion. In two-round municipalities, voters are less geographically concentrated, corresponding to a reduction that is $26.7 - 44.4\%$ of the average level of concentration in single-round municipalities. The reduction in concentration is only among the top two candidates, indicating that the main effects of the electoral rule are on the candidates with a chance of winning. These results are not driven by the increased number of candidates in two-round elections. In turn, I find that voters are more engaged in the political process. While turnout is unaffected, as turnout is mandatory in Brazil, I find that the number of blank and invalid ballots is significantly lower in two-round elections.

Second, once in office, politicians elected under two-round systems provide and allocate public goods differently. I find that two-round elections impact both the level and distribution of municipal resources by measuring the level of equipment and infrastructure resources present in public schools in the municipality. More specifically, I find that (i) public schools in two-round municipalities are $0.081$ and $0.057$ percentiles higher in the national distribution of equipment and infrastructure resources, respectively, and (ii) the standard deviation of these resources across schools is lower. It is the schools with the fewest resources in the municipality who benefit the most from these additional resources. When politicians are securing broader bases of support, they provide more public goods and distribute these resources more evenly across the municipality.

I find evidence that municipalities in two-round systems experience improved economic outcomes. While I am unable to estimate the effects precisely, I find improvements along several indicators. First, literacy rates are higher among cohorts who are of elementary school age during the mayoral term. Second, income per capita is higher and the fraction of low income households is lower.

My proposed model predicts that two-round systems lead to different outcomes because candidates adopt different strategies during the campaign. I find suggestive evidence for different strategic behavioral responses of candidates and find less evidence for other explanations. More specifically, I do not find that different types of candidates enter the races nor do different types of candidates win the elections. Candidates in two-round elections are not observably different in terms of demographic characteristics, place of birth, educational attainment, and previous occupation. Winners also do not differ along these observable
characteristics. There are more candidates entering the race in two-round elections – many of these additional candidates are represented by smaller parties and have ran in previous elections. However, these candidates are not more likely to win the election.

Instead, I find evidence that is more consistent with candidates changing their behavior. First, geographic concentration decreases between the first and second round of the election, suggesting that candidates adopt strategies between rounds to consolidate their voter bases. Second, candidates in two-round elections finance their campaigns differently: they rely less on donations from corporations. While I am unable to observe candidates’ actual behavior during the campaign, this suggests that candidates employ different campaign strategies, such as exerting more effort on door-to-door campaigns or offering policies that appeal to individuals rather than corporations. Finally, I do not find that different outcomes in two-round systems arise from different re-election incentives, suggesting that different outcomes come instead from a) mayors offering and implementing different policy platforms or b) mayors rewarding their supporters differently, as a result of greater geographic dispersion in their supporters.

This paper adds to a growing empirical literature providing causal evidence on the impacts of electoral rules by leveraging population thresholds that determine local rules. These studies, which typically compare proportional to single-round systems or single- to two-round systems, have examined how electoral rules affect turnout (Eggers, 2013); strategic voting and the number of candidates (Fujiwara, 2011); the election of clientelistic parties (Pellicer and Wegner, 2013); and fiscal expenditures (Chamon et al., 2018). Of particular interest are Bordignon et al. (2016), who compare single- and two-round systems in Italian municipalities and find that municipalities under two-round systems exhibit more policy moderation, as measured by the volatility across elections of a municipal tax rate. In contrast to Bordignon et al. (2016), I study not only an aggregate policy outcome, the overall level of public goods provision, but also the allocation of this policy across the electorate. This paper’s contribution is to provide evidence that electoral rules have economic consequences, both on public goods provision and how these public goods are distributed.

4There is empirical evidence that candidates qualifying for the second round rally votes from supporters of the candidates eliminated after the first round. Pons and Tricaud (2018) find in France that the qualification of a third candidate in the second round reduces the votes for the top two candidates, indicating that when the third candidate is not present in the second round (as is always the case in Brazil), the top two candidates capture votes from supporters for the third placed candidate.

5The effect of the electoral rule on the number of candidates is known as Duverger’s Law, which states that single-round plurality elections will lead to a two-party system, while a two-round majority system will lead to a multi-party system. This prediction has been formalized, and sometimes challenged, in recent literature (Bouton, 2013; Bouton and Gratton, 2015; Callander, 2005; Cox, 1997; Osborne and Slivinski, 1996).
This study builds on a large, mostly theoretical literature studying the role of electoral rules on political incentives. These studies document how different rules impact electoral accountability (Persson et al., 2003) and personal vote-seeking (Carey and Shugart, 1993). Importantly, these studies highlight how electoral rules may incentivize politicians to appeal to broader groups of voters (Myerson, 1993), provide public goods with broader benefits (Lizzeri and Persico, 2001, 2005; Persson and Tabellini, 1999), and target public spending to specific groups of voters (Genicot et al., 2018; Milesi-Ferretti et al., 2002). With the exception of Lizzeri and Persico (2005), these studies do not compare single- to two-round systems. The goal of this paper is to provide a simple theoretical framework to examine how public goods provision may differ between single- and two-round systems, and to subsequently test these predictions empirically.

More broadly, this paper connects to a literature examining inequalities in the allocation of state resources. A large literature documents the role of political factors in creating these inequalities, in particular how politicians politically favor certain subgroups such as those of the same ethnicity or partisanship. A key insight that emerges from this literature is that the extent to which politicians practice political favoritism is reduced when political institutions are stronger, elections are more competitive, and citizens are more broadly engaged in the electoral process (Burgess et al., 2015; Fujiwara and Wantchekon, 2013; Hodler and Raschky, 2014). Notably, Golden and Min (2013) emphasize the importance of policy responsiveness to voter preferences. Electoral rules serve as a key channel through which voter preferences are translated into policy outcomes. The final contribution of this paper is to document how electoral rules can influence the way state resources are allocated.

The remainder of this paper is organized as follows. Section 2 presents a simple theoretical framework for single- and two-round electoral systems. Section 3 describes the context and section 4 describes the empirical strategy that will be used. Section 5 presents the results on electoral outcomes and public goods provision. Section 6 discusses the results and mechanisms. Section 7 concludes.

---

2. Theoretical framework

In this section, I present a simple, stylized model to illustrate how two-round elections may incentive politicians to secure a broader base of support. My model of electoral competition adapts a standard probabilistic voting model and follows the setup in Genicot et al. (2018) by allowing for targeting of government interventions to specific localities within a municipality. I extend this model by (i) introducing a third non-strategic candidate who appeals to a single locality and (ii) adapting it to the context of single- and two-round elections.

The basic intuition is that two round elections perform two functions. First, two round elections require the winner to attain a vote share that is above 50%. This makes it difficult for politicians who appeal to a minority of the electorate to gain enough votes to win the election. Second, the existence of a second round effectively limits the number of candidates. Candidates who expect to gather sufficiently many first round votes to qualify for the second round but less than the 50% required to win in the second round must act as if a second round will occur where only one other candidate will stand. As a result, the existence of a second round incentivizes candidates to offer policy proposals that appeal to all localities, even to localities that are likely to vote for another candidate in the first round. While this effect runs counter to a large literature documenting the positive impacts of political competition, in a single-round election, higher electoral competition incentivizes candidates to appeal to narrower groups and ignore other voters.

My model predicts that two-round elections lead to different outcomes because candidates adapt different campaign strategies by offering policy proposals that appeal to a broader group of voters. However, two-round elections may also lead to different outcomes by (i) causing different types of candidates to enter the race or (ii) causing different types of candidates to win the election. While I do not model candidates’ entry decisions in the theoretical framework, I test for these alternative mechanisms in the empirical analysis. I find empirical evidence suggesting that different outcomes in two-round elections are mostly driven by candidates’ strategic responses.

---

7 Please refer to Burden (1997); Lindbeck and Weibull (1987); and Persson and Tabellini (2000) for an overview and review of probabilistic voting models.

8 There is a theoretical literature arguing that higher electoral competition can have the negative effect of incentivizing candidates to focus on narrower groups. Myerson (1993) shows that candidates offer more unequal campaign promises when electoral competition increases, even when voters are homogeneous, in elections using rank-scoring rules. Lizzeri and Persico (2005) extend this model to other electoral rules and introduce a public good, finding a negative effect of political competition on the equality of campaign promises as well as public good provision.
2.1. The environment

Consider an election with three politicians and $J$ localities within a municipality. Politicians are indexed by $c \in \{A, B, C\}$ and localities are indexed by $j \in \{1, 2, \ldots, J\}$ where $J \geq 3$. Each locality has a continuum of voters of mass $1/J$.

Prior to election day, politicians simultaneously announce a platform $q^c = \{q^c_1, q^c_2, \ldots, q^c_J\}$ that describes the fraction of the government budget that will be allocated to each locality if the politician is elected. Thus, the politician’s budget constraint is:

$$\sum_{j=1}^{J} q_j \leq 1$$

Since each locality has the same number of voters, the fraction of the government budget allocated to a locality corresponds to the fraction of the government budget allocated to each voter in that locality. Platforms are binding for politicians both between rounds and after the election.\footnote{This is not unrealistic as the time between rounds is often short compared to the length of the campaign. In Brazil, the second round is held three weeks after the first round. Nevertheless, my model translates into contexts where this assumption is slightly relaxed, as long as there is some continuity between the two rounds. First, if voters’ second round vote depends on a candidate’s policy proposal in both rounds. Second, if candidates are allowed to change their policy proposals between rounds but are constrained in the extent to which their proposals can change.}

To make solving a three-candidate model more tractable, I assume that the third candidate $C$ is a non-strategic candidate with the following platform:

$$q^C = \{0, 0, \ldots, 0, 1\}$$

There are several possible interpretations of candidate $C$. In my model, because voters are partitioned into distinct geographic localities, $C$ is a candidate whose supporters are all located in the same geographic area. However, my model easily translates into other interpretations of candidate $C$. For example, $C$ may be a candidate whose supporters share a common trait and vote for her due to descriptive representation. These traits can include geography but can also include other dimensions such as age or race. Another possibility is that $C$ is a single-issue candidate who attracts voters who only care about that issue. In any of these cases, $C$ should be viewed as a small candidate and $A$ and $B$ as front-runners in relation to $C$. This also empirically matches elections in Brazil. Third placed candidates in Brazil receive on average 11.9% of the vote (11.8% in single-round elections and 13.5% in the
first round of two-round elections), and the vote spread between the second and third placed candidates is on average 23.8% (24.0% in single-round elections and 14.8% in the first round of two-round elections).

Following Genicot et al. (2018), voters in locality \(j\) have preferences \(u_j(q_j)\) over government spending. The function \(u_j(\cdot)\) is strictly increasing and concave in \(q_j\). In addition to the policy component of voters’ preferences, there is an individual shock \(v_i\) and a municipality shock \(\delta\) towards candidate \(A\) which are independently and uniformly distributed:

\[
v_i \sim U\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right] \quad \delta \sim U\left[-\frac{1}{2\gamma}, \frac{1}{2\gamma}\right]
\]

\(v_i\) captures idiosyncratic voter preferences towards candidate \(A\). \(\delta\) captures any political dimensions that swing voters in the municipality as a whole towards candidate \(A\), such as economic shocks or popularity shocks.

Voters cast ballots for the politician who yields the highest payoff. In localities \(j \in \{1, \cdots, J-1\}\), or where candidate \(C\) has not invested, this amounts to voting for either \(A\) or \(B\)\(^{10}\). In locality \(J\), or where candidate \(C\) is dominant, voters randomize between voting for \(C\) with probability \(1-\alpha\) and either for \(A\) or \(B\) with probability \(\alpha\), depending on whether \(A\) or \(B\) offers the higher payoff, where \(0 < \alpha < 1\)\(^{11}\).

Thus, in general, voters will vote for \(A\) if and only if:

\[
u_j(q_j^B) \leq u_j(q_j^A) + v_i + \delta \quad (2.1)
\]

In localities \(j \in \{1, \cdots, J-1\}\), all voters for whom this is true vote for \(A\)\(^{12}\). In locality \(J\), a fraction \(\alpha\) of voters for whom this is true vote for \(A\).

The electoral rules follow those in Brazil. In a single-round system, the candidate with the most votes wins. In a two-round system, if no candidate attains more than 50% of the vote in the first round, the second round occurs with the top two candidates.

\(^{10}\)Since \(u_j(\cdot)\) is strictly increasing, candidates \(A\) and \(B\) will invest a non-zero amount in these localities, so voters will always vote for either \(A\) or \(B\).

\(^{11}\)Assuming \(\alpha > 0\) performs two functions. First, it guarantees that candidate \(C\) always gains strictly less than \(1/3\) in vote share and so never has the most votes. Note that \(C\) will also never have the most votes if \(J > 3\). Second, it guarantees a non-zero first order condition for locality \(J\) in the single-round election, which allows a direct comparison between the single- to the two-round election. This assumption can be relaxed and and will yield the same predictions, see section A.5.

\(^{12}\)For localities \(j \in \{1, \cdots, J-1\}\), we also need that \(u_j(q_j^A) + v_i + \delta \geq u_j(q_j^C)\). Because \(u_j(q_j^B) \geq u_j(0) = u_j(q_j^C)\), condition (2.1) is sufficient for voters to vote for \(A\) in these localities.
2.2. Preliminaries

2.2.a. Vote shares with three candidates.– First, I derive candidates’ vote shares when three candidates are present. This applies to single-round elections and the first round of two-round elections.

Condition (2.1) corresponds to voters for whom \( v_i \geq u_j(q_j^B) - u_j(q_j^A) - \delta \), which allows us to calculate each candidate’s vote share in a locality. As in Genicot et al. (2018), I assume that there are voters to be swung in every locality\(^\text{13}\). Thus, in locality \( j \) during round \( t = 1 \), candidate \( c \)’s vote share \( \pi_{jt}^c \) is given by:

\[
\begin{align*}
\pi_{j1}^A &= \begin{cases} 
\frac{1}{2} + \psi \left( \Delta u_j^{AB} + \delta \right) & \text{if } j \in \{1, \ldots, J-1\} \\
\alpha \left( \frac{1}{2} + \psi \left( \Delta u_j^{AB} + \delta \right) \right) & \text{if } j = J
\end{cases} \\
\pi_{j1}^B &= \begin{cases} 
\frac{1}{2} + \psi \left( \Delta u_j^{BA} - \delta \right) & \text{if } j \in \{1, \ldots, J-1\} \\
\alpha \left( \frac{1}{2} + \psi \left( \Delta u_j^{BA} - \delta \right) \right) & \text{if } j = J
\end{cases} \\
\pi_{j1}^C &= \begin{cases} 
0 & \text{if } j \in \{1, \ldots, J-1\} \\
1 - \alpha & \text{if } j = J
\end{cases}
\end{align*}
\]

where \( \Delta u_j^{cd} \equiv u_j(q_j^c) - u_j(q_j^d) \).

Candidates’ total vote share in the municipality \( \pi_t^c \) is given by:

\[
\begin{align*}
\pi_1^A &= \left( \frac{J-1+\alpha}{J} \right) \left( \frac{1}{2} + \psi \delta \right) + \psi \left( \frac{J-1}{j=1} \sum \Delta u_j^{AB} + \alpha \Delta u_j^{AB} \right) \\
\pi_1^B &= \left( \frac{J-1+\alpha}{J} \right) \left( \frac{1}{2} - \psi \delta \right) + \psi \left( \frac{J-1}{j=1} \sum \Delta u_j^{BA} + \alpha \Delta u_j^{BA} \right) \\
\pi_1^C &= \frac{1-\alpha}{J}
\end{align*}
\]

As in Genicot et al. (2018), I assume that all localities are contestable\(^\text{14}\). Then the

\(^{13}\)This assumption states that \( u_j(q_j^B) - u_j(q_j^A) - \delta \in \left( -\frac{1}{2\psi}, \frac{1}{2\psi} \right) \). See section A.1 for a discussion.

\(^{14}\)This assumption states that \( 0 < Pr(\pi_1^A \geq \theta) < 1 \), or that:

\[
\left( \frac{J-1+\alpha}{J} \right) \left( -\frac{\psi}{2\gamma} + \frac{1}{2} \right) < \theta < \left( \frac{J-1+\alpha}{J} \right) \left( \frac{\psi}{2\gamma} + \frac{1}{2} \right)
\]

and similarly for \( Pr(\pi_1^B \geq \theta) \). See section A.2 for a discussion.
probability of attaining a vote share above $\theta$ is given by:

\[
Pr \left( \pi_1^A \geq \theta \right) = \frac{1}{2} + \frac{\gamma}{\psi} \left[ \frac{1}{2} - \left( \frac{J}{J - 1 + \alpha} \right) \theta + \left( \frac{\psi}{J - 1 + \alpha} \right) \left( \sum_{j=1}^{J-1} \Delta u_j^{AB} + \alpha \Delta u_J^{AB} \right) \right] \tag{2.2}
\]

\[
Pr \left( \pi_1^B \geq \theta \right) = \frac{1}{2} + \frac{\gamma}{\psi} \left[ \frac{1}{2} - \left( \frac{J}{J - 1 + \alpha} \right) \theta + \left( \frac{\psi}{J - 1 + \alpha} \right) \left( \sum_{j=1}^{J-1} \Delta u_j^{BA} + \alpha \Delta u_J^{BA} \right) \right] \tag{2.3}
\]

2.2.b. **Vote shares with two candidates.** Next, I derive vote shares when there are two candidates, $A$ and $B$, present (as mentioned, candidate $C$ never makes it to the second round). This applies to the second round in two-round elections.

Candidate $A$’s vote shares in each locality and the municipality as a whole are given by:

\[
\pi_{j2}^A = \frac{1}{2} + \psi \left( \Delta u_j^{AB} + \delta \right)
\]

\[
\pi_2^A = \frac{1}{2} + \psi \delta + \frac{\psi}{J} \sum_{j=1}^{J} \Delta u_j^{AB}
\]

Assuming again that all localities are contestable (see section A.2), the probability that candidate $A$ attains a vote share above $\theta$ is given by:

\[
Pr \left( \pi_1^A \geq \theta \right) = \frac{1}{2} + \frac{\gamma}{\psi} \left[ \frac{1}{2} - \theta + \frac{\psi}{J} \sum_{j=1}^{J} \Delta u_j^{AB} \right] \tag{2.4}
\]

2.3. **Equilibrium strategies in a single-round election**

In a single-round election, candidates maximize the probability of winning, or the probability that they attain a vote share above $\frac{1}{2} \left( 1 - \frac{1-\alpha}{J} \right)$. Using equation (2.2), for candidate $A$, this is equivalent to solving the following maximization problem:

\[
\max_{q^A=\{q_1^A, \ldots, q_J^A\}} \frac{1}{2} + \left( \frac{\gamma}{J - 1 + \alpha} \right) \left( \sum_{j=1}^{J-1} \Delta u_j^{AB} + \alpha \Delta u_J^{AB} \right) \quad \text{s.t. } \sum_j q_j^A \leq 1
\]
The first order conditions are:

$$
\left( \frac{\gamma}{J - 1 + \alpha} \right) u_j'(q_j^A) = \lambda_{1R} \quad \text{for } j \in \{1, \cdots, J - 1\} \quad (2.5)
$$

$$
\left( \frac{\gamma}{J - 1 + \alpha} \right) \alpha u_j'(q_j^A) = \lambda_{1R} \quad \text{for } j = J \quad (2.6)
$$

where $\lambda_{1R}$ is the Lagrange multiplier of the budget constraint in a single-round system. For localities $j \in \{1, \cdots, J - 1\}$, equation (2.5) implies that localities with a higher marginal utility (or a higher sensitivity to changes in the electoral platform) are allocated more government spending. Of interest here is the ratio in marginal utilities between locality $j$ and $J$, which combining equations (2.5) and (2.6) is given by:

$$
\frac{u_j'(q_j^A)}{u_j'(q_J^A)} = \alpha \quad \forall j \in \{1, \cdots, J - 1\} \quad (2.7)
$$

Since $\alpha < 1$ and $u_j(.)$ is strictly increasing and strictly concave for all $j$, equation (2.7) yields the following prediction:

**Prediction 2.8.** In a single-round system, in equilibrium, candidates promise less resources to locality $J$ in comparison to the other localities. In other words, for all $j \in \{1, \cdots, J - 1\}$, we have that $q_j^A > q_J^A$.

2.4. Equilibrium strategies in a two-round election

In a two-round election, candidates again maximize the probability of winning. Here, this consists of the probability of winning in the first round and, if a second round occurs, the candidate’s probability of winning in the second round:

$$
Pr(A \text{ wins in 1st round}) + Pr(\text{second round occurs}) \cdot Pr(A \text{ wins 2nd round})
$$

$$
= Pr \left( \pi_1^A \geq \frac{1}{2} \right) + \left( 1 - Pr \left( \pi_1^A \geq \frac{1}{2} \right) - Pr \left( \pi_1^B \geq \frac{1}{2} \right) \right) Pr \left( \pi_2^A \geq \frac{1}{2} \right)
$$
Using equations (2.2), (2.3), and (2.4), this corresponds to the following maximization:

$$\max_{q^A = \{q^A_1, \ldots, q^A_J\}} \left( \frac{1}{2} + \gamma \left[ \frac{1}{2} \left( \frac{\alpha - 1}{J - 1 + \alpha} \right) + \left( \frac{\psi}{J - 1 + \alpha} \right) \left( \sum_{j=1}^{J-1} \Delta u^A_j + \alpha \Delta u^A_J \right) \right] \right)$$

$$+ \frac{\gamma}{\psi} \left( \frac{1 - \alpha}{J - 1 + \alpha} \right) \left[ \frac{1}{2} + \frac{\gamma}{J} \sum_{j=1}^{J} \Delta u^A_j \right]$$

s.t. $\sum_j q^A_j \leq 1$

The first order conditions are:

$$\left( \frac{\gamma}{J - 1 + \alpha} \right) \left( 1 + \frac{(1 - \alpha)\gamma}{\psi J} \right) u'(q^A_j) = \lambda_{2R} \quad \text{for } j \in \{1, \ldots, J - 1\} \quad (2.9)$$

$$\left( \frac{\gamma}{J - 1 + \alpha} \right) \left( \alpha + \frac{(1 - \alpha)\gamma}{\psi J} \right) u'(q^A_J) = \lambda_{2R} \quad \text{for } j = J \quad (2.10)$$

where $\lambda_{2R}$ is the Lagrange multiplier of the budget constraint in a two-round system. Combining equations (2.9) and (2.10), the ratio in marginal utilities between locality $j$ and $J$ is:

$$\frac{u'(q^A_j)}{u'(q^A_J)} = \frac{\alpha + \frac{(1 - \alpha)\gamma}{\psi J}}{1 + \frac{(1 - \alpha)\gamma}{\psi J}} \quad \forall j \in \{1, \ldots, J - 1\} \quad (2.11)$$

As in the single-round system, since $\alpha < 1$ and $u_j(\cdot)$ is strictly increasing and strictly concave for all $j$, equation (2.11) yields the following prediction:

**Prediction 2.12.** In a two-round system, in equilibrium, candidates promise less resources to locality $J$ in comparison to the other localities. In other words, for all $j \in \{1, \ldots, J - 1\}$, we have that $q^A_j > q^A_J$.

### 2.5. Comparing single- to two-round systems

In this section, I compare politician’s allocations under the single- and the two-round system. The outcome of interest is how the allocation of resources to locality $J$, where candidate $C$ has committed all of her resources, compares to the allocation to the other localities. Comparing equation (2.7) with (2.11) yields the following prediction:
Prediction 2.13. Candidates promise more to locality $J$ and promise less to the other localities in a two-round system. In other words, $q_j$ is larger in a single-round system compared to a two-round system for all $j \in \{1, \cdots, J-1\}$ and $q_J$ is smaller in a single-round system compared to a two-round system.

The proof is in section A.4. Since $q_j > q_J \forall j \in \{1, \cdots, J-1\}$ (predictions (2.8) and (2.12)), prediction (2.13) implies the following prediction\(^{15}\):

Prediction 2.14. The level of inequality in government resources promised, as measured by the resources promised to locality $J$ compared to the other localities, is lower in two-round systems and higher in single-round systems.

In a single-round system, the presence of a third candidate who appeals to a single locality incentivizes candidates $A$ and $B$ to ignore those voters, as the marginal return to allocating resources to that locality is lower than in other localities. In contrast, the two-round system, while not producing a completely equitable distribution, incentivizes candidates to allocate more resources to that locality. The presence of the second round, where the third candidate is not present, raises the marginal return to allocating resources to that locality. This can be seen in the first order condition in the single-round system (equation (2.6)) compared to that in the two-round system (equation (2.10)), which has the additional term $\frac{(1-\alpha)^\gamma}{\psi J}$. In turn, candidates $A$ and $B$ solicit more votes from locality $J$ in a two-round election.

While my model partitions the electorate into distinct geographic localities, voters can be partitioned along other dimensions, such as income, race, or ideology. My model easily translates into these other settings, and the predictions yield the same interpretation. First, two-round systems raise the marginal return to allocating resources to groups of voters that are heavily targeted by other candidates. This results in candidates offering campaign promises that appeal to these voters. Second, when politicians must commit to their promises if elected, state resources in a two-round system are allocated more equally across groups in the municipality.

\(^{15}\)My model focuses on the difference in allocations between locality $J$ and the other localities. However, depending on the functional form of $u_j(\cdot)$, it is possible that relative allocations between the other $\{1, \cdots, J-1\}$ localities become more unequal in two-round elections. If differences in $u_j(\cdot)$ across the localities are sufficiently small, any increases in inequality between the $\{1, \cdots, J-1\}$ localities will be small relative to the decrease in inequality caused by changes in locality $J$. In the case that $u_j(\cdot) = u(\cdot)$, then all $\{1, \cdots, J-1\}$ localities will receive the same allocation and the only disparity that matters is between $J$ and the other localities.
3. Institutional context

Brazil’s current electoral system was introduced in the 1988 Constitution. Municipalities were established as autonomous governmental entities with an elected executive (mayors, or prefeitos) and legislative body (a council of legislators, or camara de vereadores). Elections for municipal positions are at large and simultaneously held every four years. Mayors in municipalities with less than 200,000 voters are elected through a single-round system, while in larger municipalities they are elected in a two-round system. In the two-round system in Brazil, the threshold for winning is 50%. If no candidate receives at least 50% of the votes in the first round, then a second round is held 3 weeks later with the top two candidates. Legislators are elected through an open-list proportional system. Voters cast separate votes for a mayoral candidate and a council candidate (or a generic vote for the party). Mayors are limited to serving two consecutive terms while there is no term limit for legislators.

Assignment of the electoral rule is based on the number of registered voters in the municipality. State electoral authorities, the Tribunais Electorais Regionais (TRE), are responsible for registering citizens and maintaining electoral rolls. Several features of Brazilian elections facilitate voter turnout on election day. First, voter registration is compulsory and must be completed at least 151 days prior to the election. Second, voting is compulsory for all literate Brazilian citizens between 18-69 years of age. Third, elections are held on the first Sunday in October, a day when few voters are at work.

Brazilian elections are a multi-party system, with over 30 political parties registered in the 2016 municipal elections. Mayoral candidates are associated with a party and often a coalition of parties, which are generally formed prior to the election. Party and coalition affiliations serve as important linkages to the state and federal levels of government (Brollo and Nannicini, 2012). In particular, municipal politicians are viewed as important local operatives during state and federal elections, which are held two years after municipal elections. As a result, local politicians face incentives to deliver votes for their parties at the state and federal levels.

---

16 In my sample, a second round occurs 57% of the time.
17 The registration deadline is specified in Law 9504 Article 91.
18 Compulsory voting is specified in Article 14 of the Constitution. Voters may obtain exemptions by providing justification to a local electoral judiciary office. Absent this justification, voters who fail to vote must pay a small fine and those who fail to vote for three consecutive elections are prevented from accessing public services, such as obtaining a passport or government loans.
19 This is also mandated in Article 77 of the Constitution.
20 Seats for the legislative council are allocated based on the number of votes received by candidates or parties in the coalition.
Once elected, mayors have a broad mandate to provide public goods, particularly in education, health, and local infrastructure projects. Municipal revenue for the municipalities in my sample is a combination of inter-government transfers and local revenues. The majority of the municipal budget is represented by inter-government transfers (on average, 65.6% of the municipal budget)\textsuperscript{21}. The bulk of these transfers come from the state or federal level (on average, 78.0%), and are either constitutional automatic transfers (\textit{Fundo de Participação do Municípios}, or FPM) or discretionary transfers (\textit{convênios}). Municipalities face considerable flexibility in spending these transfers. Among the constitutional automatic transfers, 70% of the funds are unrestricted. While 30% are earmarked, municipalities are only restricted to spending it on health and education.

The allocation of spending on public goods in municipalities occurs through two main channels. First, the majority of public goods are allocated through the annual budgetary process. Second, mayors and legislators can submit bills requesting specific public works or services. While these actions jointly require approval of the mayor and legislature, mayors retain veto power and wield significant power and influence over the process.

As a result, mayors have substantial influence over both the size of the municipal budget and how municipal funds are allocated. This study focuses on public goods provision in municipal education, for two reasons. First, a large fraction of the municipal budget is allocated towards education spending: in 2012, this represented 30.5% of municipal budgets. Second, the provision of elementary education (\textit{Ensino Fundamental}) is one of the few public goods almost entirely under the jurisdiction of the municipality. For other public goods, municipalities typically share joint responsibility with the state or federal government.

4. Empirical strategy

4.1. Econometric framework

The threshold rule for mayoral elections provides a natural candidate for a regression discontinuity design (RDD). I exploit the 200,000 registered voter threshold, which determines a municipality’s electoral rule, to estimate the impact of the electoral rule.

Denote $Y_i(0)$ and $Y_i(1)$ as the potential outcome in municipality $i$ if assigned a single-round

\textsuperscript{21}For municipalities within 50,000 voters of the 200,000 registered voter threshold. Smaller municipalities rely far more on inter-government transfers for revenue, representing on average 90.4% of the municipal budget.
and two-round system, respectively. Assignment of the electoral rule is determined by the running variable, the number of registered voters $X_i$. The assignment variable $D_i$ takes on the value of $D_i = 0$ if $X_i < 200,000$ and $D_i = 1$ if $X_i \geq 200,000$. Define the local average treatment effect as $E[Y_i(1) - Y_i(0)|X_i = 200,000]$. If the conditional expectation function for the potential outcomes is continuous at the threshold, then the local average treatment effect is identified and equal to $E[Y_i|\lim X_i \downarrow 200,000] - E[Y_i|\lim X_i \uparrow 200,000]$.

To estimate the treatment effect at the discontinuity, I use a local linear regression specification:

$$Y_{it} = \beta_1 D_{it} + \beta_2 X_{it} + \beta_3 X_{it} \cdot D_{it} + \gamma_t + \epsilon_{it}$$ (4.1)

where for municipality $i$ in election year $t$, $X_{it}$ is the running variable (the number of registered voters), $D_{it}$ is the assignment variable (which is equal to 1 if $X_{it} \geq 200,000$), $\gamma_t$ is an election-year fixed effect, and $Y_{it}$ is the outcome of interest. $\beta_1$ represents the estimate of the local average treatment effect. Throughout the analysis, standard errors are clustered at the municipality level.

Because the treatment effect is identified only at the threshold, equation 4.1 is estimated on a limited sample of municipalities close to the threshold. The main analysis uses a 50,000 registered voter window, but robustness is provided for other bandwidths as well as bandwidths selected using data driven methods (Calonico et al., 2014 and Imbens and Kalyanaraman, 2012)\textsuperscript{22}.

4.2. Threats to validity

4.2.a. Smoothness.– In order for $\beta_1$ to represent the causal effect of the electoral rule, the conditional expectation of the potential outcomes must be continuous at the threshold. In this section, I discuss various ways this assumption can be violated and show they are unlikely to be violated in my context.

The first way the smoothness assumption can be violated is if the threshold choice is motivated by political or economic factors. There appears to be little evidence for this. The choice of 200,000 registered voters as the threshold was somewhat arbitrary and mainly reflected practical concerns regarding the cost of holding a second round (Chamon et al., 2018; Fujiwara, 2011). In addition, the 200,000 registered voter threshold was set in the

\textsuperscript{22}Because there is an extremely skewed right tail of municipality sizes due to a few extremely large municipalities such as Rio de Janeiro and São Paulo, the data-driven bandwidths would sometimes select bandwidths larger than the support. As a result, the optimal bandwidth is calculated on a subset of elections that lies within the support and is symmetrical around the threshold: 0-400,000 voters.
1988 Constitution and has not changed since then. It is unlikely that politicians chose the threshold anticipating which municipalities would be above or below this threshold in 1996 and later.

Another possibility is that other potential confounds, such as other policies, change discretely at the threshold. While a number of policies in Brazil are implemented using thresholds, most of these thresholds are based off population counts and not the number of registered voters. To the best of my knowledge, there are no other policies at the 200,000 registered voter threshold and most other policy thresholds are far from the 200,000 registered voter threshold. Two exceptions are a constitutional amendment in 2000 that places a salary cap for local legislators at 300,000 inhabitants and a constitutional amendment in 2004 that determines the size of the local legislature at 285,714 inhabitants. Both of these thresholds potentially lie within my bandwidth. I address this in two ways. First, I estimate a placebo regression where the electoral rule is assigned at these population thresholds. I show that there are no discontinuities of a similar size in the mayoral electoral outcomes I examine (Tables C.3 and C.4). However, legislator salaries and legislature size can affect economic outcomes, so I do not estimate placebo regressions for my public goods outcomes. I next test whether the probability of being above or below these thresholds changes discontinuously at the 200,000 voter threshold. I do not find evidence of a discontinuity (Figure C.1)\textsuperscript{23}, indicating that any effect of these policies is balanced across the threshold.

A third way the smoothness assumption can be violated is if municipalities sort across the threshold. Visually, there do not appear to be any discontinuities in the distribution of registered voters at the 200,000 registered voter threshold (Figure 2). To test this more formally, I estimate the size of the discontinuity in the density of the running variable at the threshold (McCrary, 2008). The size of the discontinuity is both small in magnitude (0.169) and insignificant ($p = 0.389$)\textsuperscript{24}.

I test whether potential outcomes vary smoothly around the discontinuity. To do this, I estimate placebo regressions where the electoral rule is assigned at other registered voter thresholds between 170,000 and 230,000 voters. This tests whether the effects I estimate are concentrated at the actual threshold of 200,000 voters. I show that there is no discontinuity in my outcomes at thresholds where the treatment does not change (Figures C.10, C.11, C.12).

\textsuperscript{23}The regression discontinuity estimate on the probability of being above the 300,000 resident threshold is $-0.0350$ ($p = 0.735$). The regression discontinuity estimate on the probability of being above the 300,000 resident threshold is $0.0516$ ($p = 0.667$).

\textsuperscript{24}Because there is an extremely skewed right tail of municipality sizes, the size of the discontinuity was estimated off a sample of municipalities excluding those above the 99.9 percentile of registered voters.
and C.13).

I test whether potential confounds change discretely across the threshold by estimating equation 4.1 on pre-treatment characteristics of municipalities. Since the treatment (here, the two-round election) is determined by the number of registered voters, municipalities can move into the treatment group or be treated multiple times\textsuperscript{25}. As a result, I define “pre-treatment” in two ways: 1) prior to the introduction of the electoral system in the 1988 Constitution, and 2) prior to the most recent election the municipality was untreated or prior to 1996 if the municipality was never untreated\textsuperscript{26}. To measure outcomes for (1), I use the 5% sample of the 1980 census, the most recent census prior to 1988. To measure outcomes for (2), I use outcomes either from the census prior to the most recent year in a single-round system (the 1991, 2000, or 2010 census) or outcomes from the 1991 census, for municipalities who were in a two-round system in 1996.

There is no significant treatment effect on nearly all outcomes measured prior to the 1988 Constitution (Panel A of Table 1) and prior to the most recent election in a single-round system (Panel B of Table 1). The treatment effect is estimated on a number of characteristics, including economic characteristics (such as the unemployment rate or share of low income households), segregation levels\textsuperscript{27} (along income and demographics), and income inequality. To rule out the concern that there are factors that change discontinuously at the threshold and affect which municipalities move into treatment and length of treatment, I test whether population growth is discontinuous at the threshold. If municipalities that grow at slower rates are more likely to be under the threshold and remain there longer, this poses an issue for causal identification. I do not find that population growth is discontinuous at the threshold.

One exception is a large and significant effect on population density. However, there are several reasons to believe this is a false positive. First, dropping one outlier municipality reduces the coefficient by 38.2%, suggesting that a few municipalities are driving this effect, although the estimate is still significant at the 10% level. My main results are all robust to

\textsuperscript{25}Municipalities can also move out of treatment. While municipalities do experience population decline, none moves below the 200,000 voter threshold.

\textsuperscript{26}The earliest electoral data available is 1996, so I cannot observe whether municipalities are treated or untreated prior to 1996. Since there is only one unobserved municipal election after the 1988 Constitution (in 1992) and only 45 municipalities have moved across the 200,000 voter threshold between 1996 and 2016, it is unlikely that many municipalities had their electoral status change between 1992 and 1996.

\textsuperscript{27}This is calculated using the entropy index (see section B.1 for detailed calculation), which measures the extent to which the composition of defined groups within census sectors matches the composition of these groups in the municipality as a whole. This was calculated separately for income and demographics. For income, the groups are defined by bins of income relative to the minimum wage. For demographics, the groups are defined by sex, age, and literacy.
dropping this outlier municipality. In addition, the effect disappears at larger bandwidths (Figures C.4 and C.5). Second, the estimate \( (p = 0.047) \) is not significant after Bonferroni adjusting the significance threshold for the number of hypotheses tested. Third, the regression discontinuity coefficients across the ten outcomes (in the most recent single-round) are not jointly significant \( (p = 0.339) \).

While it is not clear how an imbalance in population density impacts the economic and political outcomes of interest, this does pose an issue if politicians are manipulating the composition and size of the electorate by, say, moving citizens or manipulating municipality borders. This does not appear to be the case for two reasons. First, I find no differences in the urbanization rate (a difference of 0.065\%, where the single-round mean is 95.191\%, \( p = 0.943 \)). Second, this effect is seen in 1980, prior to the introduction of the threshold rule (Panel A of Table 1). Third, there is no evidence that changes in municipality area or population growth change discontinuously across the threshold (Panel B of Table 1). Nevertheless, I include population density as a control in all specifications\(^{28} \).

4.2.b. Registered voters as a running variable.— The fact that treatment (moving into a two-round election) is determined by the number of registered voters generates interesting implications for the interpretation of the regression discontinuity estimate.

As discussed above, municipalities can move into treatment or be treated multiple times. While this does not invalidate causal identification, it does affect whether the treatment is interpreted as a short-term or long-term treatment. In other words, whether the treatment effects I estimate are the result of single or multiple treatments.

To investigate this, I test whether the probability that a municipality’s previous election was a two-round system changes discontinuously at the threshold (Figure C.3). Because the regression discontinuity framework identifies the treatment effect at the threshold, any treatment effects on my outcomes of interest should be interpreted as the result of the change in this probability at the threshold. I find that at the 200,000 registered voter threshold, the probability of having previously been treated is zero and there is no discontinuous jump\(^{29} \). As a result, the treatment effects I estimate identify the effect of moving into a two-round

\[ Y_{it} = \beta_1 D_{it} + \beta_2 X_{it} + \beta_3 X_{it} \cdot D_{it} + \beta_4 Z_{it} + \beta_5 Z_{it} \cdot D_{it} + \gamma_t + \varepsilon_{it} \]

where \( Z_{it} \) is the municipality’s population density in the most recent census prior to the election.

\(^{28}\)I include it linearly and estimate it separately on both sides of the discontinuity:

\(^{29}\)Because no municipalities move under the threshold, all municipalities under the threshold have 0 probability that the previous election was two-round. The regression discontinuity estimate on the probability of being in a two-round system in the previous election is \(-0.00802 (p = 0.870)\).
election for the first time.

4.2.c. Compliance. – While not an issue for causal identification, imperfect compliance with treatment can affect the interpretation of the causal estimates (Angrist et al., 1996). In this context, compliance was perfect (Figure C.2). All municipalities below the threshold or where the top candidate received at least 50% of the vote held one round. All municipalities above the threshold and where the top candidate did not receive at least 50% of the vote held two rounds.

4.3. Data sources

4.3.a. Electoral data. – Data on municipality elections come from Brazil’s electoral authority (Tribunal Superior Eleitoral or TSE). These are available for 6 municipal elections between 1996 and 2016. The electoral results provide information on the candidates running, the party and coalition each candidate belongs to, and votes received. In total, there are 32,767 municipal elections covering 5,568 municipalities. In particular, I leverage the fact that these results are available for each polling station (seção eleitoral), which allows me to observe at a very fine level the number of votes each candidate receives. I use this to measure the level of geographic dispersion of voters within the municipality.

Geographic dispersion of voters. – I construct indices of voter concentration and the dispersion in vote shares. My main indices use votes from the first round of elections, but results are robust to using votes from the second round.

The dispersion in vote shares is a candidate-level measure of geographic dispersion, as it takes the standard deviation of a candidate’s vote share across polling stations in the municipality. Intuitively, if a candidate receives support from more areas of the municipality, she should experience less variation in vote shares across polling stations.

To measure the overall level of geographic dispersion in the municipality, I use indices from the racial segregation literature and calculate two indices which measure the extent to which voters for specific candidates are concentrated in the municipality: the fractionalization index and the entropy index. The fractionalization index is also known as interaction index, and measures the probability that two members within a population chosen at random will be from different groups. The entropy index measures how far the population is from equal representation of all groups. These indices are calculated for each polling station to measure

---

30These indicators are described in more detail in White (1986) and the construction of these indices is described in Section B.1.
the extent to which voters for specific candidates are concentrated in a polling station. The average level across all polling stations is normalized by the overall concentration of voters in the municipality to yield the index I use in my analysis. Higher values of these indices indicate that the composition of polling stations is closer to containing only one voter type, and lower values indicate that the composition of polling stations is closer to the overall composition of voters in the municipality.

4.3.b. Public goods provision in schools.– To measure public goods provision in municipal public schools, I use data provided in the School Census (Censo Escolar). This census is conducted annually by the research arm of the Ministry of Education, Instituto Nacional de Estudos e Pesquisas Educacionais or INEP.

I use the Census to observe the level of resources present in schools offering elementary education (Ensino Fundamental), which is the level of education almost exclusively under the jurisdiction of the municipal government. I calculate a measure of resources present in each school and consider two categories: equipment and infrastructure. Equipment includes movable elements such as the number of computers and availability of air conditioning (see Table B.1 for the full list). Infrastructure includes immovable elements such as the number of classrooms, type of sanitation, and availability of a library (see Table B.2 for the full list). I then construct indices of these resources, separately for equipment and infrastructure, by taking the principal component of these variables and computing the school’s percentile rank within the country in each year.\footnote{The variables collected in the School Census under each category varied from year to year. As a result, this makes it difficult to compare the raw PCA index across years. Calculating a school’s percentile rank across all schools for that year allows comparisons across years.}

5. The effect of the two-round system

I present two main results. First, in section 5.1, I find that candidates in two-round elections receive broader geographical support. Second, in section 5.2, I find that once in office, politicians elected under two-round systems provide more resources to schools and distribute these resources more equitably. These results suggest that politicians in two-round systems are represented by a broader group of voters and that this affects how public goods are provided in the municipality. In section 5.3, I find suggestive evidence that these effects lead to improved economic outcomes in two-round municipalities.
5.1. The geography of votes

Do candidates in two-round elections secure broader bases of support? In this section, I provide evidence that voters for specific candidates are more geographically dispersed in two-round elections, indicating that candidates receive support from a geographically broader group of voters. To quantify the extent to which candidate’s supporters are geographically concentrated or dispersed, I calculate two sets of statistics. First, I calculate measures of concentration of candidate vote shares across polling stations in the municipality to measure an overall level of geographic dispersion. Second, I calculate the standard deviation in a candidate’s vote shares across polling stations in the municipality to measure geographic dispersion at the candidate-level. Across both measures, I find that candidates in two-round elections receive support from geographically broader areas of the municipality.

5.1.a. Geographic concentration of voters. – The overall level of geographic concentration of voters is lower in municipalities holding two-round elections (Figure 3 and Panel A of Table 2). The coefficients on both the entropy index and fractionalization measure are negative, indicating that the composition of voters within polling stations is closer to the composition of voters in the municipality as a whole. Municipalities in two-round systems have a reduction of 0.012 in the fractionalization of voters and a reduction of 0.008 in the entropy index, which corresponds to 44.4% and 26.7%, respectively, of the average level in single-round municipalities within the 50,000 voter bandwidth.

While the concentration indices indicate that voters are overall more geographically dispersed, it does not indicate whether all or only some candidates are obtaining support from geographically broader constituencies. I find that voters for the top two candidates are more dispersed in two-round systems, but that voters for the third and fourth placed candidates in two-round systems are concentrated similarly to voters of the same-placed candidates in single-round systems (Figure 4 and Panel B of Table 2). The estimates for the top two candidates are similar in magnitude: the first placed candidate experiences a 0.017 reduction in variance (21.3% of the single-round mean) and the second placed candidate experiences a 0.014 reduction (18.7% of the single-round mean). The estimates for the third and fourth placed candidates are close to zero and insignificant, suggesting that only the top two candidates garner different bases of support in two-round elections.

These effects are not limited to vote shares in the first round (Table 3). I re-calculate all measures using only vote shares from the top two candidates and compare concentration
in single-round elections with concentration in the second round of two-round elections. Voters in the second round are also less geographically concentrated (Panel A of Table 3). I also compare concentration between the first and second round in two-round municipalities. While these estimates are correlational and not causal, concentration in the second round is lower than in the first round, indicating that candidates consolidate their voter bases between rounds (Panel B of Table 3).

5.1.b. Voter engagement. – Taken together, these results indicate that the candidates elected in two-round elections are supported by a geographically broader group of voters. I also find evidence of voter behavior complementary to these results. More specifically, I find higher rates of voter engagement in two-round elections. While turnout is unaffected (which is to be expected, as turnout is mandatory in Brazil), the number of blank and invalid ballots is significantly lower in two-round municipalities (Table 4). The number of blank and invalid ballots plausibly corresponds to dissatisfied or disinterested voters (Gonzales et al., 2019). The reduction suggests that voters in two-round elections engage in the electoral process at higher rates.

5.2. The allocation of municipal resources

I next investigate the impact of the two-round system on public goods provision. If politicians are securing broader bases of support in two-round elections, they may also provide public goods differently once in office. I provide evidence that two-round elections impact both the level and distribution of municipal resources. Specifically, I estimate the impact on the overall level and variance of the resources that are present in public schools in municipalities.

5.2.a. Municipal schools. – Municipal schools in two-round municipalities have, on average, more resources (Figure 5 and Columns 1 and 2 in Table 5). The coefficients on the level of equipment and infrastructure resources that are present in schools are both positive. A school in a two-round municipality is, on average, 0.081 percentiles and 0.057 percentile higher in the national distribution of equipment resources and infrastructure resources, respectively.

32I use only vote shares from the top two candidates since the second round always has two candidates. This makes the outcomes more comparable.

33Ballots can be invalid or blank for a number of reasons. For example, municipalities with higher numbers of illiterate voters will have more blank and invalid ballots (Fujiwara, 2015). Since the illiteracy rate is not discontinuous across the threshold and all municipalities used electronic voting by 2000 (which substantially reduced the number of unintentional errors), I interpret the difference in the number of blank and invalid ballots as voter engagement. Gonzales et al. (2019) provide empirical evidence for this interpretation as they find that forced electoral participation increases the number of blank and invalid ballots cast.
The coefficient on infrastructure resources is smaller and less significant, which may be due to the fact that infrastructure is more difficult for mayors to manipulate, as building new infrastructure requires significantly more capital and time to implement. This is reflected when looking at the estimates separately by the year in the mayoral term: infrastructure levels are less responsive to the electoral cycle (Column 1 in Table 6).

In addition to differences in the overall levels, there is less variance in the resources present in schools in two-round municipalities (Figure 6 and Columns 1 and 2 in Table 5). The standard deviation in equipment resources is 0.018 percentiles lower in two-round municipalities (15.9% of the single-round mean). Although the estimate on the standard deviation in infrastructure resources is of a similar magnitude (−0.021 percentiles), the difference is not significant. Given the results on the overall level and lower responsiveness of infrastructure resources, it is not surprising that estimates on the variance in infrastructure resources are noisier.

It is instructive to estimate the effects on the level of resources that schools at different parts of the distribution in the municipality have. Intuitively, if the variance in resources is lower in two-round municipalities, we should expect to see that schools with the least (most) resources in the municipality have more (less) resources.

I group schools into quartiles, which are defined by first calculating for each school its percentile in the municipality distribution prior to the election, and then assigning the school to one of four quartiles. For both equipment and infrastructure elements, the increased level of resources is concentrated in schools located at the lower end of the distribution (Table 7). Schools at the bottom 25% of the distribution are 0.082 percentiles higher in equipment resources and 0.116 percentiles higher in infrastructure resources. Schools in the second quartile experience positive, but smaller gains – 0.066 percentiles in equipment resources and 0.102 percentiles in infrastructure resources. There is no significant difference in resources at the top of the distribution.

5.3. Economic outcomes

Sections 5.1 and 5.2 establish that two-round elections lead to the election of politicians who attract a geographically broader group of voters, provide more resources to municipal schools, and distribute these resources more equally across schools in the municipality. I next explore whether these effects lead to different economic outcomes in two-round municipalities. I use the 2000 and 2010 Demographic Censuses, which allows me to observe economic
outcomes in municipalities between 2-10 years after the election\textsuperscript{34}.

I find suggestive evidence that two-round municipalities experience improvements along several indicators (Table 5.3). First, literacy rates are higher among cohorts who are of elementary school age during the mayoral term. While the baseline level of literacy is high, at 91.4\% in single-round municipalities within the 50,000 voter bandwidth, the literacy rate increases by 1.2 percentage points at the threshold. Second, income per capita is higher and the fraction of low income households is lower. This suggests that outcomes are improved both for the average and the poorest households in the municipality, although more modestly for the average household. The difference from the pre-treatment level (in Panel B of Table 1) is statistically significant for the fraction of low income households ($p = 0.007$) and close to statistical significance for income per capita ($p = 0.123$). While I am unable to estimate these results precisely, as outcomes are measured in the decennial censuses and the 2020 Census has not yet been conducted (thus I am unable to measure post outcomes for the 2012 and 2016 elections), the magnitude and direction of the estimates suggest that two-round municipalities do experience improved economic outcomes.

5.4. Robustness of the main results

5.4.a. RDD design. – I investigate the robustness of my results to the regression discontinuity design. First, these results are not driven by the choice of bandwidth, whether fixed or chosen by a data-driven method\textsuperscript{35} (Figures C.6 and C.7 for voter concentration outcomes; Figures C.8 and C.9 for municipal school outcomes). The estimates maintain similar magnitudes and mostly retain significance for bandwidths out to 150,000 voters, although the estimates for the standard deviation of school resources decline at larger bandwidths. Second, dropping controls from the regression, namely population density and election-year fixed effects, does not substantially affect the results, although the results are noisier (Tables C.1 and Table C.2).

5.4.b. Placebo tests. – There are policy thresholds at 300,000 inhabitants, where a cap on local legislator salaries comes into effect, and 285,714 inhabitants, where the size of the local

\textsuperscript{34}For the 1996 elections, outcomes are observed 4 years later in the 2000 Demographic Census. For the 2000, 2004, and 2008 elections, outcomes are observed 10, 6, and 2 years later in the 2010 Demographic Census.

\textsuperscript{35}I use two methods to choose the bandwidth: Imbens and Kalyanaraman (2012) (“IK”) and Calonico et al. (2014) (“MSERD”). Because of the extremely skewed right tail of municipality sizes, IK and MSERD sometimes select bandwidths larger than the support. As a result, the optimal bandwidth is calculated on a subset of elections that lies within the support and is symmetrical around the threshold: 0-400,000 voters.
legislature changes, that fall within my bandwidth and have the potential to confound my results. I show that the probability of being above or below these thresholds does not change discontinuously at the 200,000 registered voter threshold (Figure C.1), indicating that any effects of these policies are balanced across both sides of the threshold. I also re-estimate the mayoral electoral results with this threshold as a placebo. While mayoral electoral outcomes can plausibly be affected by these policies, they less likely to be influenced than public goods outcomes. I do not find a similar effect at these thresholds (Tables C.3 and C.4).

I also show that there are no discontinuities at placebo thresholds in registered voters (170,000; 180,000; 190,000; 210,000; 220,000; 230,000), indicating that the outcomes are relatively continuous at places where the treatment does not change (Figures C.10 and C.11 for voter concentration outcomes; Figures C.12 and C.13 for municipal school outcomes). The treatment effect is isolated to the actual threshold: there are no estimates with the same size and significance as at the actual threshold.

5.4.c. Measuring concentration.– One source of mechanical bias for the concentration indices is the size of the parcels (here, polling stations) used to calculate the indices. First, the number of voters assigned to each polling station is regulated by the TSE so all polling stations should in principle be of similar size (on average, there are 272 valid votes at each polling station). Second, the difference in the number of valid votes at each polling station is small in magnitude (3.605 votes) and insignificant ($p = 0.764$).

A second concern, although potentially a mechanism, is that the reduced variance and concentration in vote shares may be driven by the number of candidates competing: there is a sizeable increase in the number of candidates in two-round elections (Table 4). I do not find evidence that the number of candidates explains the reduction in voter concentration. While the entropy index and fractionalization measure may be affected by the number of candidates, the direction of bias is not monotonic (White, 1986). Nevertheless, I perform three robustness checks.

First, while the number of candidates is a bad control as it is an endogenous outcome, controlling for the number of candidates does not affect the results (Table C.5). Second, I

36Since the number of inhabitants is not the same as the number of registered voters (nor do they map 1:1), to maintain comparability with the baseline estimates, I use a bandwidth of 125,000 inhabitants. This bandwidth was determined by taking half of the population range of municipalities in my 50,000 voter bandwidth (the smallest municipality is 182,082 inhabitants and the largest at 434,474 inhabitants). Since the salary cap was implemented in 2000, I estimate this using elections after 2000. Since the legislature size was implemented in 2004, I estimate this using elections after 2004.
simulate the effect of adding an additional candidate to all single-round elections. For most measures, the estimated bias is small and, for fractionalization, of the wrong sign (the row “Potential bias” in Table 2). However, the bias is substantial for the entropy index. As a result, I re-calculate all the outcomes using only the vote shares from the top two candidates, so as to maintain the same number of candidates across the single- and two-round elections. Doing so does not substantially change the results (Table C.6).

6. Discussion

The results presented indicate that candidates, particularly the top two, in two-round elections are supported by a geographically broader groups of voters. Once in office, these politicians provide higher levels of resources to municipal schools and distribute them more evenly across schools within the municipality. In my model, two-round elections lead to different outcomes because candidates adjust their strategies by offering policies that appeal to a broader group of voters. However, there may be other explanations. Namely, different types of candidates may enter two-round elections or different types of candidates may win two-round elections. In the following section, I explore these explanations and provide suggestive evidence that candidates’ strategic behavioral responses explain a larger part of the effect of the two-round system.

6.1. Selection in candidates

One reason that candidates in two-round elections may have a broader group of supporters is that two-round elections encourage different types of candidates to enter electoral races. For example, two-round elections may incentive candidates who are more competent or

---

37 I do this by adding in a last place candidate to each polling station. I assign that candidate the number of votes the average last place candidate receives (1.5%). I then take away a proportionate number of votes from all other candidates, in order to ensure that the total number of voters remains the same. I then calculate the change between the actual value and the simulated value.

38 The measures are calculated assuming only the top two candidates compete in the election. By construction, the standard deviation in vote shares of the 2nd place candidate will be exactly the same as that of the 1st place candidate, so only the standard deviation in vote shares for the 1st place candidate is shown.

39 A third possibility is that voters behave differently in two-round elections, either through turnout or strategic voting. In Brazil, turnout is not a major factor as turnout is mandatory. Regarding strategic voting, Fujiwara (2011) finds that third placed and lower candidates receive higher vote shares in two-round elections and argues that voters are behaving less strategically in two-round elections. While this paper is not focused on voter behavior, I interpret strategic behavioral responses of candidates as an equilibrium outcome that can arise from the electoral rule directly or indirectly through the electoral rule’s impact on voter behavior.
more relatable to enter the races. I do not find that candidates in two-round elections are significantly different from candidates in single-round elections along observable characteristics. There is no significant difference on age, sex, educational attainment, or state of birth (Panel A of Table 10)\textsuperscript{40}. Candidates also do not have different occupational backgrounds (Panel B of Table 10).

A second possibility is that different types of candidates \textit{win} in two-round elections. I also find no evidence that election winners are observably different in terms of demographics, education, place of birth, and occupational background (Table 11).

I do find differences in political affiliation among candidates who enter the elections. There are more candidates from small parties and who previously ran as mayoral candidates in two-round elections (Panel A of Table 12), but they are not more likely to win the election (Panel B of Table 12)\textsuperscript{41}. Interestingly, candidates with incumbency are less likely to win two-round elections, although this result is not robust to other bandwidths\textsuperscript{42}. Since smaller parties are more likely to appeal to narrower electorates, it is unlikely that the identity of the candidate’s party is explaining the reduced concentration in vote shares.

This lines up with my model and empirical results, where lower placed candidates are small candidates whose strategies do not change between the single- and two-round system. In section 5.1, I find that there is less variance in vote shares among the top two candidates but not among the third and fourth placed candidate. More small candidates enter two-round elections, and there are several reasons they may do so. First, they may be building support for subsequent elections, which is one reason more candidates with previous campaign experience enter two-round elections. Second, these candidates may be seeking to gain positions in the elected administration. Third, and more interestingly, these candidates may want to influence the top candidates’ platforms. In my model, in the two-round election, the small, third candidate indirectly influences the top candidates’ platforms by leading these candidates to offer policies that appeal to her voters. In reality, there may be more direct ways the third candidate can affect these policies, either by strategically shaping her own policy or through direct bargaining.

\textsuperscript{40}Educational attainment is an indicator for whether the candidate completed university-level education. Place of birth is whether the candidate was born in the same state as the election.

\textsuperscript{41}I define a small party as any party that is not one of the top 5 parties in Brazil by national membership.

\textsuperscript{42}Incumbency is defined as whether the candidate held a mayoral position in a previous electoral term.
6.2. Strategic behavioral responses by candidates

I find limited evidence that two-round elections cause different types of individuals to enter or win the elections. Instead, I find more evidence that is consistent with candidates adjusting their strategies during the election, suggesting that candidates’ strategical behavioral responses explain a larger part of the effects I find.

6.2.a. Concentration between the first and second round.– As mentioned in section 5.1, I find that the concentration of voters among the top two candidates decreases between the first and second round of the two-round election (Panel B of Table 3). While these estimates are correlational and not causal, this suggests that candidates consolidate their voter bases between rounds. In my model, candidates do not change their platforms between rounds. However, my model extends to situations where candidates can change their platforms but are either a) constrained in the extent to which they can do so or b) voters’ vote depend on both the first and second round policies. It is more likely that these extensions more closely match to reality. There is empirical evidence that candidates qualifying for the second round rally votes from supporters of the candidates eliminated after the first round (Pons and Tricaud, 2018). The decrease in concentration I find between rounds suggests that candidates adjust their strategies in order to rally these voters.

6.2.b. Campaign financing.– The TSE provides data on campaign finances for elections after 2004, which allows me to observe one aspect of candidate’s strategies: how the campaigns are financed. The data allows me to observe the identity of each candidate’s donors. I find that candidates in two-round elections receive less donations, both on average (Panel A of Table 13) and between the top two candidates (Panel B of Table 13). While the outcomes are noisy, the effects are strongest for donations from corporations: candidates in two-round elections receive fewer donations from corporations.

Unfortunately, I cannot observe the exact behaviors and strategies candidates are implementing, but these results suggest that candidates in two-round elections run their campaigns differently. The reduction in donations from corporations could arise from several possible behaviors; for example, candidates could be exerting more effort on door-to-door campaigns or offering policies that appeal to individuals rather than corporations.

Donors are classified as corporations or individuals, depending on whether a CPF (individual identification number) or CNPJ (corporate identification number) was filed for the donation.
6.2.c. Rewarding voters.– Differences in public goods provision in two-round elections may be a result of politicians facing different re-election incentives, politicians offering and implementing different policy platforms, or politicians rewarding their supporters differently. I do not find evidence that politicians face different re-election incentives: there is no significant difference in the treatment effect on municipal school resources between mayors in their first term (eligible for re-election) or second term (Table 8).

One possibility is that mayors in two-round municipalities provide public goods differently because their supporters are less geographically concentrated and they subsequently reward their supporters in a less geographically concentrated fashion. In Brazil, mayors face strong incentives to reward their supporters after the election, even those who are not eligible for re-election. Municipal mayors are viewed as important local operatives who deliver votes for their parties at the state and federal levels, and it is not uncommon for mayors to seek office themselves at the state and federal levels. In my sample, 46.2% of mayors stand as candidates in a state or federal election after their term and 83.7% of mayors have stood at this level ever\textsuperscript{44}. This is reflected when looking at the treatment effect on municipal school resources for equipment by year in the term: the discontinuity is largest in the second and third year of the term, which coincides with the timing of state and federal elections (Table 6).

However, the treatment effects in different years are not significantly different from each other, with the exception of the standard deviation in equipment resources where the third year is significantly less than the fourth year ($p = 0.085$). As a result, I cannot fully distinguish between a) mayors offering and implementing different policy platforms from b) mayors rewarding their supporters differently. It is likely that two-round elections incentivize a combination of the two, resulting in higher and broader public goods provision.

6.3. Why not two-round systems everywhere?

If two-round systems lead to positive outcomes along both public goods provision (section 5.2) and downstream economic indicators (section 5.3), why is the two-round system not more widely used? In this section, I discuss possible, although not an exhaustive set of, trade-offs that countries may face.

First, two-round elections are costly. The main reason Brazil did not implement two-round

\textsuperscript{44}Among municipalities within the 50,000 registered voter bandwidth. In larger cities, the position of the municipal mayor is often viewed as a stepping stone for higher office. Running in state and federal elections is less common in smaller municipalities.
systems in all municipalities, resulting in the threshold rule, was a concern for the monetary cost of holding two-round elections. In addition, this imposes a cost on voters who have to vote twice in a short span of time. In fact, I find that turnout is lower in the second round. While these estimates are correlational, the turnout rate is 2.48 percentage points lower in the second round compared to the first round ($p < 0.001$), suggesting that there is indeed a cost to voters to voting a second time.

A second reason is that my results show that in two-round systems, individuals at lower parts of the distribution benefit. Specifically, I find that schools at the lower part of the distribution receive more resources (Table 7) and the fraction of low-income households is lower (Table 9). As a result, there may be opposition by richer households or the elite to implementing two-round systems.

A final reason is that two-round elections may only result in better outcomes when the electorate is composed of many small groups. Brazil is a multi-party system with over 30 parties registered in the 2016 elections and the average single-round election has 4.6 candidates running. As a result, an electoral rule that incentivizes candidates to incorporate these smaller groups in the coalition may lead to better outcomes. However, these results may not translate to contexts where the electorate is composed of two large groups. In these contexts, the two-round system is unlikely to incentivize politicians to incorporate both groups in her coalition, resulting in further entrenchment within one of the groups.

7. Conclusion

A majority of countries use two-round systems to choose their leaders, and an increasing number of countries are adopting this system over time. This paper studies how the electoral rule may lead to the election of politicians who represent a broader or narrower groups of voters and distribute state resources differently. To identify the effect of the two-round system, I leverage a unique rule in Brazilian municipal elections: municipalities above a threshold of registered voters hold two rounds whereas municipalities below this threshold hold a single round. I find that candidates in two-round municipalities are represented by a geographically broader group of voters and that voters are less concentrated in these elections. The reduction in overall concentration is $26.7 - 44.4\%$ of the average level in municipalities holding single-round elections. Once in office, mayors elected under two-round systems provide more equipment and infrastructure resources to municipal schools and distribute
these resources more evenly across schools. Schools in two-round municipalities are 0.081 and 0.057 percentiles higher in the national distribution of equipment and infrastructure resources, respectively, and the variance in equipment and infrastructure resources is 0.018 and 0.021 percentiles lower, respectively.

I present a simple model to highlight why two-round elections may lead to these empirical results. The main intuition is that two-round elections perform two functions. First, they require a candidate to attain at least 50% of the vote in order to win. Second, the second round effectively limits the number of candidates to two. In my model, the presence of a third candidate in a single-round system leads candidates to ignore the voters that the third candidate is appealing to. In contrast, candidates in a two-round system must look ahead to the possibility of a second round, where the third candidate is not present, occurring. As a result, in a two-round system, the top two candidates promise a higher allocation of resources to those voters.

My model proposes that two-round elections lead to different outcomes because candidates adjust their behavior. I find evidence suggesting that two-round systems cause candidates to adjust their behavior rather than cause different types of candidates to enter the races. Candidates entering two-round elections are not significantly different from those in single-round elections along observable characteristics. Candidates in two-round races are more likely to come from a small party and to have run for office previously, but these candidates are not more likely to win. Instead, I find suggestive evidence that the main differences are candidates’ strategic responses. First, geographic concentration in voters decreases between the the first and second round, suggesting that candidates consolidate their bases of support. Second, candidates in two-round elections rely less on corporate donors for their campaigns.

Differences in politicians’ voter bases affect how mayors distribute public goods once in office. A limitation of my paper is that I am unable to observe other aspects of politicians’ behavior, including those of lower placed candidates, during their campaigns nor do I address voters’ behavior. Future work that is able to map out politicians’ behavior more fully and integrate strategic behaviors by voters would greatly advance our current understanding of electoral rules.
References


Figure 1  Electoral systems around the world

*Single-round plurality* (light red) and *Two-round majority* (dark red) correspond to the single- and two-round system in Brazil. Vertical axis corresponds to the number of countries under the specified electoral system. Data encompasses 497 presidential elections across 71 countries with democratic regimes, between 1946-2016. *Source:* Bormann and Golder (2013).
Figure 2  Density of elections around the 200,000 registered voter cutoff

Plot includes only elections with between 50,000 and 400,000 registered voters (6.0% of the universe of elections). Size of the discontinuity in the density of elections uses the McCrary test and is estimated off all elections below the 99.9 percentile in size. An “election” is defined as a municipality-election year. Bin sizes are 10,000 voter bins.
Table 1  Regression discontinuity estimates on municipality pre-characteristics

**Panel A: Characteristics measured prior to the 1988 Constitution**

<table>
<thead>
<tr>
<th></th>
<th>% illiterate</th>
<th>Unempl. rate</th>
<th>% low income</th>
<th>Pop. density</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>-1.499</td>
<td>0.044</td>
<td>-0.751</td>
<td>-397.678**</td>
</tr>
<tr>
<td></td>
<td>(1.057)</td>
<td>(0.145)</td>
<td>(1.148)</td>
<td>(184.347)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>19.029</td>
<td>1.978</td>
<td>50.182</td>
<td>516.067</td>
</tr>
<tr>
<td>Observations</td>
<td>293</td>
<td>293</td>
<td>293</td>
<td>293</td>
</tr>
</tbody>
</table>

**Panel B: Characteristics measured prior to most recent single-round**

<table>
<thead>
<tr>
<th></th>
<th>Muni. area change (%)</th>
<th>Pop. growth (%)</th>
<th>Pop. density</th>
<th>Income seg.</th>
<th>Dem. seg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>-0.601</td>
<td>-0.468</td>
<td>-663.035**</td>
<td>-0.003</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(2.296)</td>
<td>(1.260)</td>
<td>(332.555)</td>
<td>(0.004)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>-3.003</td>
<td>2.863</td>
<td>1,013.396</td>
<td>0.090</td>
<td>0.028</td>
</tr>
<tr>
<td>Observations</td>
<td>231</td>
<td>295</td>
<td>295</td>
<td>231</td>
<td>231</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>% illiterate</th>
<th>Income per capita</th>
<th>% low income</th>
<th>Unempl. rate</th>
<th>Gini coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>0.014</td>
<td>-18.964</td>
<td>1.211</td>
<td>-0.652</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.590)</td>
<td>(33.328)</td>
<td>(2.290)</td>
<td>(1.114)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>7.409</td>
<td>646.917</td>
<td>36.123</td>
<td>11.671</td>
<td>0.540</td>
</tr>
<tr>
<td>Observations</td>
<td>295</td>
<td>295</td>
<td>295</td>
<td>295</td>
<td>295</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Panel A: outcomes from the 1980 census. Panel B: outcomes either from the census prior to the most recent election in a single-round system or from the 1991 census. F-stat for all treatment effects in Panel B jointly significant: 1.146 (p = 0.339). Muni. area change is the percentage change in municipality area from the prior census. Pop. growth is the percentage change in population from the prior census. Pop. density is the population density, per km$^2$. Income seg. and Dem. seg. refer to income and demographic segregation, respectively, of census tracts (measured using the entropy index). Income per capita is the average monthly household income per capita, in reais. % low income is the fraction of households earning between 0 and 50% of the minimum wage. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Standard errors clustered at the municipality level. Source: 1980, 1991, 2000, and 2010 Demographic Census.
Figure 3  Regression discontinuity plots of concentration indices

(a) Fractionalization

(b) Entropy index

*Fractionalization* (a) and *Entropy index* (b) are indices of concentration of candidates’ supporters, using vote counts from polling stations. Vote shares are from the first round. Bin sizes are 10,000 voter bins. Diameter of the points is proportional to the number of observations. Confidence intervals represent the 95% confidence intervals of a linear regression with standard errors clustered at the municipality level.
Figure 4  Regression discontinuity plots of the standard deviation in votes received

(a) 1st place candidate

(b) 2nd place candidate

Vertical axis is the standard deviation in a candidate’s vote counts across polling stations, for the 1st place (a) and 2nd place (b) candidate. Vote shares are from the first round. Bin sizes are 10,000 voter bins. Diameter of the points is proportional to the number of observations. Confidence intervals represent the 95% confidence intervals of a linear regression with standard errors clustered at the municipality level.
Table 2  Regression discontinuity estimates on the geographic dispersion of voters

Panel A: Concentration indices of voters for specific candidates

<table>
<thead>
<tr>
<th></th>
<th>Fractionalization</th>
<th>Entropy index</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>−0.012**</td>
<td>−0.008*</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Potential bias</td>
<td>0.0002</td>
<td>−0.0270</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.027</td>
<td>0.030</td>
</tr>
<tr>
<td>Observations</td>
<td>264</td>
<td>264</td>
</tr>
</tbody>
</table>

Panel B: Dispersion in vote shares for each candidate

<table>
<thead>
<tr>
<th></th>
<th>1st place candidate</th>
<th>2nd place candidate</th>
<th>3rd place candidate</th>
<th>4th place candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>−0.017**</td>
<td>−0.014*</td>
<td>−0.005</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.007)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Potential bias</td>
<td>−0.0011</td>
<td>−0.0011</td>
<td>−0.0004</td>
<td>−0.0002</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.080</td>
<td>0.075</td>
<td>0.042</td>
<td>0.023</td>
</tr>
<tr>
<td>Observations</td>
<td>264</td>
<td>264</td>
<td>251</td>
<td>216</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Panel A: concentration indices (fractionalization and entropy index) of candidates’ supporters, using vote counts from polling stations. Panel B: dispersion in a candidate’s vote shares (for the 1st-4th place candidate), measured as the standard deviation in vote counts across polling stations. Potential bias is the simulated effect on the outcome from having an additional candidate in every single-round election. Vote shares are from the first round. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table 3  Estimates on the geographic dispersion of voters across different rounds of elections

<table>
<thead>
<tr>
<th></th>
<th>Fractionalization</th>
<th>Entropy index</th>
<th>Std Dev of 1st place candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: 1st round in single-round compared to 2nd round in two-round</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td>−0.023***</td>
<td>−0.017***</td>
<td>−0.019**</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.006)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.038</td>
<td>0.029</td>
<td>0.088</td>
</tr>
<tr>
<td>Observations</td>
<td>216</td>
<td>216</td>
<td>216</td>
</tr>
<tr>
<td><strong>Panel B: 1st round in two-round compared to 2nd round in two-round</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2ndRound</td>
<td>−0.012***</td>
<td>−0.010***</td>
<td>−0.013***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>First round mean</td>
<td>0.049</td>
<td>0.037</td>
<td>0.103</td>
</tr>
<tr>
<td>Observations</td>
<td>432</td>
<td>432</td>
<td>432</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Fractionalization and Entropy index are indices of concentration of candidates’ supporters, using vote counts from polling stations. Standard deviation of 1st place candidate is the standard deviation in the candidate’s vote counts across polling stations. All outcomes use only vote shares from the top two candidates. Panel A compares the 1st round results (in single-round elections) with the 2nd round results (in two-round elections) and presents the regression discontinuity estimates. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level. Panel B compares the 1st round results (in two-round elections) with the 2nd round results (in two-round elections), using the full sample of elections that held two rounds. Estimation method: Standard regression with 2ndround as the regressor and election fixed effects. Standard errors clustered at the municipality level.
Table 4  Regression discontinuity estimates on other electoral outcomes

<table>
<thead>
<tr>
<th></th>
<th>Turnout</th>
<th>Blank/invalid ballots</th>
<th># candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>0.006</td>
<td>−3.821**</td>
<td>1.273***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(1.670)</td>
<td>(0.339)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.843</td>
<td>16.524</td>
<td>4.604</td>
</tr>
<tr>
<td>Observations</td>
<td>296</td>
<td>296</td>
<td>296</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Turnout is the fraction of eligible voters who cast a ballot in the election. Blank/invalid ballots is the sum of ballots (in thousands) that were either blank or voided, and is also equal to turnout minus valid ballots.

Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

Table 5  Regression discontinuity estimates on resources in municipal schools

<table>
<thead>
<tr>
<th></th>
<th>Mean level of resources</th>
<th>Standard deviation in resources</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Equipment</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.081**</td>
<td>0.057*</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.738</td>
<td>0.731</td>
</tr>
<tr>
<td>Observations</td>
<td>820</td>
<td>912</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Equipment and Infrastructure are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. The first two columns (Mean level of resources) has as the dependent variable the mean index level across schools in the municipality. The last two columns (Standard deviation in resources) has as the dependent variable the standard deviation in the index across schools in the municipality. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Figure 5  Regression discontinuity plots of the overall level of resources in municipal schools

(a) Equipment, mean level of resources

(b) Infrastructure, mean level of resources

Equipment and Infrastructure are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. Mean level of resources is the mean index level across schools in the municipality for equipment (a) and infrastructure (b). Bin sizes are 10,000 voter bins. Diameter of the points is proportional to the number of observations. Confidence intervals represent the 95% confidence intervals of a linear regression with standard errors clustered at the municipality level.
Figure 6  Regression discontinuity plots of the distribution of resources in municipal schools

(a) Equipment, standard deviation in resources

(b) Infrastructure, standard deviation in resources

*Equipment* and *Infrastructure* are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. *Standard deviation in resources* is the standard deviation in the index across schools in the municipality for equipment (a) and infrastructure (b). Bin sizes are 10,000 voter bins. Diameter of the points is proportional to the number of observations. Confidence intervals represent the 95% confidence intervals of a linear regression with standard errors clustered at the municipality level.
Table 6  Regression discontinuity estimates on resources in municipal schools, by year in the term

<table>
<thead>
<tr>
<th></th>
<th>Mean level of resources</th>
<th>Standard deviation in resources</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Equipment</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>Panel A: First year in the electoral term</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.077**</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Observations</td>
<td>197</td>
<td>227</td>
</tr>
<tr>
<td>Panel B: Second year in the electoral term</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.091**</td>
<td>0.059*</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Observations</td>
<td>197</td>
<td>228</td>
</tr>
<tr>
<td>Panel C: Third year in the electoral term</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.087**</td>
<td>0.059*</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>Observations</td>
<td>197</td>
<td>228</td>
</tr>
<tr>
<td>Panel D: Fourth year in the electoral term</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.070**</td>
<td>0.059*</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>Bandwidth size</td>
<td>50,000</td>
<td>50,000</td>
</tr>
<tr>
<td>Observations</td>
<td>229</td>
<td>229</td>
</tr>
</tbody>
</table>

* p < 0.10, ** p < 0.05, *** p < 0.01

Mayoral terms are for four years. Each panel displays the estimate separately for the 1st year (Panel A), 2nd year (Panel B), 3rd year (Panel C) and 4th year (Panel D) of the term. Equipment and Infrastructure are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. The first two columns (Mean level of resources) has as the dependent variable the mean index level across schools in the municipality. The last two columns (Standard deviation in resources) has as the dependent variable the standard deviation in the index across schools in the municipality. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table 7  Regression discontinuity estimates on resources in municipal schools, by quartile in the municipal distribution

<table>
<thead>
<tr>
<th>Mean level of resources in schools at different quartiles</th>
<th>1st quartile</th>
<th>2nd quartile</th>
<th>3rd quartile</th>
<th>4th quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Bottom 25%)</td>
<td>(Top 25%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel A: Equipment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.082**</td>
<td>0.066*</td>
<td>0.069*</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.037)</td>
<td>(0.042)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.652</td>
<td>0.733</td>
<td>0.781</td>
<td>0.856</td>
</tr>
<tr>
<td>Observations</td>
<td>700</td>
<td>728</td>
<td>760</td>
<td>748</td>
</tr>
</tbody>
</table>

**Panel B: Infrastructure**

| TwoRound                                                 | 0.116**      | 0.102**      | 0.056        | 0.013        |
|                                                          | (0.046)      | (0.047)      | (0.035)      | (0.021)      |
| Single round mean                                        | 0.540        | 0.689        | 0.814        | 0.914        |
| Observations                                             | 776          | 764          | 784          | 780          |

*p < 0.10, **p < 0.05, ***p < 0.01

Equipment and Infrastructure are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. Dependent variables are the mean index level of equipment (Panel A) and infrastructure (Panel B) elements, separately by quartiles. Quartiles are defined by the school’s percentile in the municipal distribution in the year prior to the election. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table 8  Regression discontinuity estimates on resources in municipal schools, by possibility of re-election

<table>
<thead>
<tr>
<th></th>
<th>Mean level of resources</th>
<th>Standard deviation in resources</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Equipment</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.113**</td>
<td>0.088**</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>TwoRound * FirstTerm</td>
<td>-0.034</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.738</td>
<td>0.731</td>
</tr>
<tr>
<td>Bandwidth size</td>
<td>50,000</td>
<td>50,000</td>
</tr>
<tr>
<td>Observations</td>
<td>789</td>
<td>789</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Heterogeneous treatment effects by re-election incentives. FirstTerm is a dummy indicating whether the mayor is a first-term mayor (eligible for re-election). Equipment and Infrastructure are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. The first two columns (Mean level of resources) has as the dependent variable the mean index level across schools in the municipality. The last two columns (Standard deviation in resources) has as the dependent variable the standard deviation in the index across schools in the municipality. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

Table 9  Regression discontinuity estimates on municipal economic outcomes

<table>
<thead>
<tr>
<th></th>
<th>Elem. literacy</th>
<th>Income per capita</th>
<th>Low income rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>1.199*</td>
<td>64.667</td>
<td>-5.186*</td>
</tr>
<tr>
<td></td>
<td>(0.710)</td>
<td>(61.782)</td>
<td>(3.079)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>91.445</td>
<td>762.417</td>
<td>27.929</td>
</tr>
<tr>
<td>Observations</td>
<td>177</td>
<td>177</td>
<td>177</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Elem. literacy is the literacy rate of cohorts who are of elementary school age during the mayoral term. Low income rate is the fraction of households earning between 0 and 50% of the minimum wage. Outcomes measured either in the 2000 or 2010 Demographic Census. Estimation method: Local linear regression with a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table 10  Regression discontinuity estimates on candidate characteristics

<table>
<thead>
<tr>
<th>Panel A: Demographic characteristics of candidates</th>
<th>Age</th>
<th>Female</th>
<th>Univ. degree</th>
<th>Born same state</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>0.470</td>
<td>−0.065</td>
<td>−0.045</td>
<td>−0.070</td>
</tr>
<tr>
<td></td>
<td>(1.514)</td>
<td>(0.040)</td>
<td>(0.048)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>49.955</td>
<td>0.129</td>
<td>0.796</td>
<td>0.784</td>
</tr>
<tr>
<td>Observations</td>
<td>264</td>
<td>263</td>
<td>263</td>
<td>263</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Candidate's previous occupation</th>
<th>Public sector</th>
<th>Technical</th>
<th>Business</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>−0.008</td>
<td>0.014</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.059)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.466</td>
<td>0.388</td>
<td>0.047</td>
</tr>
<tr>
<td>Observations</td>
<td>263</td>
<td>263</td>
<td>263</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Outcomes are the average characteristics of candidates in an election. Panel A contains demographic characteristics. Univ. degree is the fraction of candidates whose highest educational attainment is university or higher. Born same state is the fraction of candidates who were born in the same state as the election. Panel B contains Previous occupation, which is the fraction of candidates with that previous occupation. Public sector includes Previous occupation, which is the fraction of candidates with that previous occupation. Technical includes occupations such as scientists, technicians, and artists. Business includes occupations such as administrative positions, workers in commerce and services, and business owners. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table 11  Regression discontinuity estimates on characteristics of the winner

<table>
<thead>
<tr>
<th>Panel A: Demographic characteristics of winners</th>
<th>Age</th>
<th>Female</th>
<th>Univ. degree</th>
<th>Born same state</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>0.160</td>
<td>-0.027</td>
<td>0.023</td>
<td>-0.041</td>
</tr>
<tr>
<td></td>
<td>(2.763)</td>
<td>(0.087)</td>
<td>(0.100)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>51.608</td>
<td>0.112</td>
<td>0.832</td>
<td>0.789</td>
</tr>
<tr>
<td>Observations</td>
<td>264</td>
<td>263</td>
<td>263</td>
<td>263</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Winner’s previous occupation</th>
<th>Public sector</th>
<th>Technical</th>
<th>Business</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>-0.088</td>
<td>0.069</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td>(0.131)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.534</td>
<td>0.348</td>
<td>0.043</td>
</tr>
<tr>
<td>Observations</td>
<td>263</td>
<td>263</td>
<td>263</td>
</tr>
</tbody>
</table>

* * * p < 0.01

Outcomes are the characteristics of the candidate who won the election. Panel A contains demographic characteristics. Univ. degree is an indicator for whether the winner’s highest educational attainment is university or higher. Born same state is an indicator for whether the winner was born in the same state as the election. Panel B contains Previous occupation, which is an indicator for the winner’s previous occupation. Public sector includes occupations such as elected positions, judiciary, and workers in public administration. Technical includes occupations such as scientists, technicians, and artists. Business includes occupations such as administrative positions, workers in commerce and services, and business owners. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table 12  Regression discontinuity estimates on candidates’ political affiliation

<table>
<thead>
<tr>
<th></th>
<th>Previous candidacy</th>
<th>Incumbency</th>
<th>Small party</th>
<th>PT party</th>
<th>Governor’s party</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: All candidates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.539**</td>
<td>−0.182</td>
<td>0.721**</td>
<td>0.038</td>
<td>−0.084</td>
</tr>
<tr>
<td></td>
<td>(0.259)</td>
<td>(0.155)</td>
<td>(0.334)</td>
<td>(0.100)</td>
<td>(0.136)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>1.652</td>
<td>0.752</td>
<td>2.535</td>
<td>0.636</td>
<td>0.584</td>
</tr>
<tr>
<td>Observations</td>
<td>263</td>
<td>263</td>
<td>296</td>
<td>296</td>
<td>263</td>
</tr>
<tr>
<td><strong>Panel B: Winner only</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td>−0.144</td>
<td>−0.201*</td>
<td>−0.008</td>
<td>−0.022</td>
<td>−0.012</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.115)</td>
<td>(0.127)</td>
<td>(0.094)</td>
<td>(0.120)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.621</td>
<td>0.410</td>
<td>0.369</td>
<td>0.187</td>
<td>0.242</td>
</tr>
<tr>
<td>Observations</td>
<td>263</td>
<td>263</td>
<td>296</td>
<td>296</td>
<td>263</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

*Previous candidacy* is whether the candidate stood as a candidate in a previous mayoral election. *Incumbency* is whether the candidate held the position of mayor in a previous term. *Small party* defined as any party that is not one of the top 5 parties, by national membership. *PT party* is whether the candidate is from the *Partido dos Trabalhadores*. *Governor’s party* is whether the candidate is from the party of the incumbent state governor. Dependent variables are either the number of candidates with that characteristic (*Panel A*) or an indicator for the winner having that characteristic (*Panel B*). *Estimation method*: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table 13  Regression discontinuity estimates on campaign donations received

<table>
<thead>
<tr>
<th></th>
<th>Donation amounts received by candidates</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>From individuals</td>
<td>From corporations</td>
</tr>
<tr>
<td>Panel A: Average donations per candidate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td></td>
<td>-0.225</td>
<td>-0.491</td>
<td>-0.673$^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.286)</td>
<td>(0.310)</td>
<td>(0.404)</td>
</tr>
<tr>
<td>Single round mean</td>
<td></td>
<td>12.844</td>
<td>10.742</td>
<td>11.782</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>154</td>
<td>154</td>
<td>154</td>
</tr>
<tr>
<td>Panel B: Total donations among top two candidates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoRound</td>
<td></td>
<td>-0.074</td>
<td>-0.614</td>
<td>-1.023$^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.335)</td>
<td>(0.492)</td>
<td>(0.541)</td>
</tr>
<tr>
<td>Single round mean</td>
<td></td>
<td>14.053</td>
<td>11.717</td>
<td>12.846</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>154</td>
<td>154</td>
<td>154</td>
</tr>
</tbody>
</table>

$^*$p < 0.10, **p < 0.05, ***p < 0.01

Panel A: Outcomes are log average donation levels, in reais, received by candidates (total donations divided by the number of candidates). Panel B: Outcomes are log total donations, in reais, received by the top two candidates. Donors identified as Individual and Corporation depending on whether the donor provided a CPF (individual identification number) or CNPJ (corporate identification number). Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
A. Theory Appendix

A.1. Swing-able voters

Equation (2.1) characterizes the voters who prefer A’s proposal to B’s proposal. In order for $0 \leq \pi^A_{j1} \leq 1$, we need that:

$$ u_j(q^B_j) - u_j(q^A_j) - \delta \in \left( -\frac{1}{2\psi}, \frac{1}{2\psi} \right) $$

or that there are swing-able voters in the locality. This assumption is satisfied if:

$$ \delta \in \left( u(1) - u(0) - \frac{1}{2\psi}, u(1) - u(0) + \frac{1}{2\psi} \right) $$

$$ \iff \frac{1}{2\gamma} + u(1) - u(0) < \frac{1}{2\psi} $$

In other words, that swings in municipality vote shares are smaller than the variation in individual preferences. Note that this implies that $\gamma > \psi$ since $u(1) - u(0) > 0$.

A.2. Contestability of localities

For equations (2.2) and (2.3) to be non-degenerate, or that $0 < Pr(\pi^A_1 \geq \theta) < 1$, we need the following to be true:

$$ \frac{1}{\psi} \left[ \left( \frac{J}{J - 1 + \alpha} \right) \theta - \frac{1}{2} - \left( \frac{\psi}{J - 1 + \alpha} \right) \left( \sum_{j=1}^{J-1} \Delta u^A_{j} + \alpha \Delta u^A_{J} \right) \right] \in \left( -\frac{1}{2\gamma}, \frac{1}{2\gamma} \right) $$

$$ \frac{1}{\psi} \left[ \frac{1}{2} - \left( \frac{J}{J - 1 + \alpha} \right) \theta + \left( \frac{\psi}{J - 1 + \alpha} \right) \left( \sum_{j=1}^{J-1} \Delta u^B_{j} + \alpha \Delta u^B_{J} \right) \right] \in \left( -\frac{1}{2\gamma}, \frac{1}{2\gamma} \right) $$

which corresponds to the following condition for the first round:

$$ \theta \in \left( \left( \frac{J - 1 + \alpha}{J} \right) \left( -\frac{\psi}{2\gamma} + \frac{1}{2} \right), \left( \frac{J - 1 + \alpha}{J} \right) \left( \frac{\psi}{2\gamma} + \frac{1}{2} \right) \right) \quad (A.1) $$

Similarly, for equation (2.4) to be non-degenerate, or that $0 < Pr(\pi^A_2 \geq \theta) < 1$, we need
that:

\[
\frac{1}{\psi} \left( \theta - \frac{1}{2} - \frac{\psi}{J} \sum_{j=1}^{J} \Delta u_{j}^{AB} \right) \in \left( -\frac{1}{2\gamma}, \frac{1}{2\gamma} \right)
\]

\[
\frac{1}{\psi} \left[ \frac{1}{2} - \theta + \frac{\psi}{J} \sum_{j=1}^{J} \Delta u_{j}^{BA} \right] \in \left( -\frac{1}{2\gamma}, \frac{1}{2\gamma} \right)
\]

which corresponds to the following condition for the second round:

\[
\theta \in \left( -\frac{\psi}{2\gamma} + \frac{1}{2}, \frac{\psi}{2\gamma} + \frac{1}{2} \right)
\]

(A.2)

Next, I show that conditions (2.2), (2.3), and (2.4) are non-degenerate in single- and two-round elections.

**Claim A.3**. \( \frac{1}{2} \left( 1 - \frac{1 - \alpha}{J} \right) \) satisfies condition (A.1).

Both the upper and lower inequalities are satisfied because:

\[-\frac{\psi}{2\gamma} + \frac{1}{2} < \frac{1}{2} < \frac{\psi}{2\gamma} + \frac{1}{2}\]

**Claim A.4**. \( \frac{1}{2} \) satisfies condition (A.1).

Note that the lower inequality is satisfied:

\[
\left( \frac{J - 1 + \alpha}{J} \right) \left( -\frac{\psi}{2\gamma} + \frac{1}{2} \right) < \frac{1}{2}
\]

because \( \frac{J - 1 + \alpha}{J} < 1 \) and \( -\frac{\psi}{2\gamma} + \frac{1}{2} < \frac{1}{2} \).

I next show that the upper inequality is satisfied. The upper inequality is equivalent to:

\[
\frac{1}{2} < \left( \frac{J - 1 + \alpha}{J} \right) \left( \frac{\psi}{2\gamma} + \frac{1}{2} \right)
\]

\[
\iff J > (1 - \alpha) \left( \frac{\gamma + \psi}{\psi} \right)
\]

which is true so long as \( J \) is large enough and \( \gamma/\psi \) is not too large.
Claim A.5. \( \frac{1}{2} \) satisfies condition (A.2).

Both the upper and lower inequalities are satisfied because:

\[-\frac{\psi}{2\gamma} + \frac{1}{2} < \frac{1}{2} < \frac{\psi}{2\gamma} + \frac{1}{2}\]

A.3. C never makes it to the second round

For C to never make it to the second round, the probability that candidates A and B attain vote shares above candidate C’s must be 1. In other words, we need that \( \pi_1^C \) does not satisfy condition (A.1):

\[
\frac{1 - \alpha}{J} \leq \left( \frac{J - 1 + \alpha}{J} \right) \left( -\frac{\psi}{2\gamma} + \frac{1}{2} \right) \quad \text{or} \quad \frac{1 - \alpha}{J} \geq \left( \frac{J - 1 + \alpha}{J} \right) \left( \frac{\psi}{2\gamma} + \frac{1}{2} \right)
\]

The first inequality is equivalent to:

\[
J \geq (1 - \alpha) \left( \frac{2\gamma}{\gamma - \psi} + 1 \right)
\]

The second inequality is equivalent to:

\[
J \leq (1 - \alpha) \left( \frac{2\gamma}{\gamma + \psi} + 1 \right)
\]

The first inequality is much more likely to be satisfied, which is true so long as \( J \) is large enough and \( 2\gamma/(\gamma - \psi) \) is not too large.

A.4. Prediction (2.13)

Denote the equilibrium allocations in a single-round system as \( q_j^{1R} \) and in a two-round system as \( q_j^{2R} \).

First, observe that the ratio of marginal utilities in the single-round system (equation (2.7)) is smaller than that in the two-round system (equation (2.11)):

\[
\frac{u'_j(q_j^{1R})}{u'_j(q_j^{1R})} < \frac{u'_j(q_j^{2R})}{u'_j(q_j^{2R})} \iff \alpha < \frac{\alpha \psi + (1 - \alpha)\gamma/J}{\psi + (1 - \alpha)\gamma/J}
\]

which is true because \( \alpha < 1 \).
For both single- and two-round elections, from equations (2.5) and (2.9), we have that
\[
u_j'(q_j) = u_{j'}'(q_{j'}) \quad \forall j, j' \in \{1, \ldots, J - 1\}
\]
Thus, we have the following corollaries:

**Corollary A.6.** If \( q_j^1 > q_j^2 \) for one \( j \neq J \) then \( q_j^1 > q_j^2 \) for all other \( j \in \{1, \ldots, J - 1\} \).

**Corollary A.7.** If \( q_j^1 > q_j^2 \) for one \( j \neq J \), then \( q_j^1 < q_j^2 \).

Corollary (A.6) follows because \( u_j(\cdot) \) is strictly concave. Corollary (A.7) follows from corollary (A.6) and the budget constraint.

With these preliminaries established, I prove the following claim:

**Claim A.8.** \( \frac{u_j'(q_j^1)}{u_j'(q_j^2)} < \frac{u_j'(q_j^2)}{u_j'(q_j^2)} \iff u_j'(q_j^1) < u_j'(q_j^2) \).

*Proof.* Assume the contrary. If \( u_j'(q_j^1) \geq u_j'(q_j^2) \), then \( q_j^1 \leq q_j^2 \). Corollary (A.7) then implies that \( q_j^1 \geq q_j^2 \) which in turn implies that \( u_j'(q_j^1) \leq u_j'(q_j^2) \). But then it cannot be true that \( \left(u_j'(q_j^1)/u_j'(q_j^2)\right) < \left(u_j'(q_j^2)/u_j'(q_j^2)\right) \) so we must have that \( u_j'(q_j^1) < u_j'(q_j^2) \).

Given that \( u_j'(q_j^1) < u_j'(q_j^2) \), this implies that \( q_j^1 > q_j^2 \). Then we have prediction (2.13):
\[
q_j^1 > q_j^2 \quad \forall j \in \{1, \ldots, J - 1\} \quad \text{corollary (A.6)}
\]
\[
q_j^1 < q_j^2 \quad \text{corollary (A.7)}
\]
A.5. Relaxing the assumption that $\alpha > 0$

Assume $\alpha = 0$ so that candidate $C$ receives all the votes in locality $J$. Then vote shares in the first round, or when all 3 candidates are present, are given by:

\[
\pi^A_1 = \left(\frac{J-1}{J}\right) \left(\frac{1}{2} + \psi \delta\right) + \frac{\psi}{J} \sum_{j=1}^{J-1} \Delta u^A_{j}
\]

\[
\pi^B_1 = \left(\frac{J-1}{J}\right) \left(\frac{1}{2} - \psi \delta\right) + \frac{\psi}{J} \sum_{j=1}^{J-1} \Delta u^B_{j}
\]

\[
\pi^C_1 = \frac{1}{J}
\]

and the probability of attaining a vote share above $\theta$ is given by:

\[
Pr(\pi^A_1 \geq \theta) = \frac{1}{2} + \frac{\gamma}{\psi} \left[1 - \left(\frac{J-1}{J}\right) \theta + \left(\frac{\psi}{J-1}\right) \sum_{j=1}^{J-1} \Delta u^A_{j}\right]
\]

\[
Pr(\pi^B_1 \geq \theta) = \frac{1}{2} + \frac{\gamma}{\psi} \left[1 - \left(\frac{J-1}{J}\right) \theta + \left(\frac{\psi}{J-1}\right) \sum_{j=1}^{J-1} \Delta u^B_{j}\right]
\]

Vote shares for the second round, or when candidates $A$ and $B$ are present, are the same as when $\alpha > 0$.

A.5.a. Equilibrium strategies in a single-round election.– The single-round maximization is given by:

\[
\max_{q^A = \{q^A_1, \ldots, q^A_J\}} \frac{1}{2} + \frac{\gamma}{J-1} \sum_{j=1}^{J-1} \Delta u^A_{j} \quad \text{s.t.} \quad \sum_{j} q^A_{j} \leq 1
\]

For $j \neq J$, the first order conditions imply, which is the same as when $\alpha > 0$:

\[u'(q^A_j) = u'(q^A_{j'}) \quad \forall j, j' \in \{1, \cdots, J-1\}\]

The first-order condition with respect to $q^A_j$ is 0, as it does not enter the maximization. I show that in equilibrium, the optimal strategy is to allocate $q^A_j = 0$.

Say $q^A_j > 0$. Consider the following deviation: $(q^A_j)' = q^A_j - \epsilon$ and $(q^A_k)' = q^A_k + \epsilon$ for some $k \neq J$ and $\epsilon > 0$. Note that candidate $C$’s vote share is unchanged, so the threshold for
winning remains \( \frac{1}{2} \left( 1 - \frac{1}{J} \right) \). Thus, the net change in the probability of winning is given by:

\[
\left[ \frac{1}{2} + \frac{\gamma}{J-1} \left( \Delta (u_k^{AB})' + \sum_{j\neq k,J} \Delta u_j^{AB} \right) \right] - \left[ \frac{1}{2} + \frac{\gamma}{J-1} \left( \Delta u_k^{AB} + \sum_{j\neq k,J} \Delta u_j^{AB} \right) \right] = \frac{\gamma}{J-1} \left[ u_k \left( (q_k^A)' \right) - u_k \left( q_k^A \right) \right] > 0
\]

where the last line follows because \( \epsilon > 0 \) and \( u_k(\cdot) \) is strictly increasing. Thus, there is a deviation which strictly increases the probability of winning. As a result, any \( q_j^A > 0 \) cannot be optimal so the optimal strategy is to allocate \( q_j^A = 0 \).

**A.5.b. Equilibrium strategies in a two-round election.** – The two-round maximization is given by:

\[
\max_{q^A \in \{q_1^A, \ldots, q_J^A\}} \left( \frac{1}{2} + \frac{\gamma}{\psi} \left[ - \left( \frac{1}{J-1} \right) \frac{1}{2} + \left( \frac{\psi}{J-1} \right) \sum_{j=1}^{J-1} \Delta u_j^{AB} \right] \right) + \frac{\gamma}{\psi} \left( \frac{1}{J-1} \right) \left[ \frac{1}{2} + \frac{\gamma}{J} \sum_{j=1}^{J} \Delta u_j^{AB} \right]
\]

s.t. \( \sum_j q_j^A \leq 1 \)

The first-order conditions are:

\[
\left( \frac{\gamma}{J-1} \right) \left( 1 + \frac{\gamma}{\psi,J} \right) u'(q_j^A) = \lambda'_{2R} \quad \text{for } j \in \{1, \ldots, J-1\}
\]

\[
\left( \frac{\gamma}{J-1} \right) \left( \frac{\gamma}{\psi,J} \right) u'(q_j^A) = \lambda'_{2R} \quad \text{for } j = J
\]

where \( \lambda'_{2R} \) is the Lagrange multiplier of the budget constraint in a two-round system when \( \alpha = 0 \). From this, we can see that \( q_j^A > 0 \), and so we have that \( q_j^{1R} < q_j^{2R} \). Since corollary (A.6) still holds and because of the budget constraint, then we must have that \( q_j^{1R} > q_j^{2R} \forall j \in \{1, \ldots, J-1\} \) and we still have prediction (2.13).
B. Data Appendix

B.1. Formulas for concentration indices

This section provides the calculations used to construct the concentration indices. For a
discussion of these indicators as well as their properties, please refer to White (1986).

The equations use the following notation. A “group” $k$ is defined as voters for a specific
candidate. The geographic unit $i$ is a polling station in municipality $m$. $p_k$ denotes the
fraction of voters for candidate $k$.

First, a polling station-level measure of diversity is calculated:

Fractionalization: $f_{im} = 1 - \sum_k (p_{imk})^2$

Entropy: $h_{im} = -\sum_k p_{imk} \ln p_{imk}$

These are then aggregated to the municipality level by taking a weighted average of the
polling station measures, weighted by the number of voters in each polling station, $n_{im}$:

$\bar{f}_m = \frac{\sum_i n_{im} f_{im}}{\sum_i n_{im}}$  $\bar{h}_m = \frac{\sum_i n_{im} h_{im}}{\sum_i n_{im}}$

These are also calculated for the municipality as a whole:

$\hat{f}_m = 1 - \sum_k (p_k)^2$  $\hat{h}_m = -\sum_k p_k \ln p_k$

The indices used in the analysis are the municipality averages normalized by the overall level
in the municipality:

$f_m = \frac{\hat{f}_m - \bar{f}_m}{\hat{f}_m}$  $h_m = \frac{\hat{h}_m - \bar{h}_m}{\hat{h}_m}$

When each polling station has the same concentration as the municipality ($\hat{f}_m = \bar{f}_m$ and
$\hat{h}_m = \bar{h}_m$), then the indices take a value of 0. When each polling station contains only one
type of voter ($\bar{f}_m = 0$ and $\bar{h}_m = 0$), then the indices take a value of 1.
### B.2. Variables from the *Censo Escolar*

#### Table B.1 Variables used to construct the equipment index

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VCR/DVD</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>TV</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>satellite dish</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>satellite dish with internet</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>overhead projector</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>projector</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fax</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>copier</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>stereo/sound system</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>camera/camcorder</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>drinking fountain</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>special needs accom.</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>fan</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>air conditioning</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>computers</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>printer</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>local network</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>internet</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>broadband</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

61
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>principal office</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>secretary office</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>teacher lounge</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>teacher housing</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>library</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>reading room</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>video library / room</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>toy library</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>auditorium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>solarium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>science lab</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>computer lab</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>other lab</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kitchen</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>food pantry</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>cafeteria</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>warehouse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>schoolyard</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>green area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>sports field</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>pool</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gymnasium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>playground</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>laundry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>sanitation</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>special needs accomm.</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>classrooms</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>
C. Additional Figures and Tables

**Figure C.1** Probability of falling above/below other policy thresholds

The vertical axis is the fraction of elections above the 300,000 resident threshold (a) and the 285,714 resident threshold (b). At 300,000 residents, a salary cap for municipal legislators comes into effect. At 285,714 residents, the size of the legislature changes. Bin sizes are 5,000 voter bins. Diameter of the points is proportional to the number of observations. Confidence intervals represent the 95% confidence intervals of a linear regression with standard errors clustered at the municipality level.
Figure C.2  Compliance with treatment assignment

The vertical axis is the fraction of elections that held two rounds of elections. The horizontal axis is the vote share of the top candidate in the first round. Municipalities below the 200,000 registered voter threshold are denoted by red circles. Municipalities above the 200,000 registered voter threshold are denoted by blue triangles. Bin sizes are 10% vote share bins.
Figure C.3  Probability of treatment in the previous election

The vertical axis is the probability that the previous election was a two-round election. The horizontal axis is the number of registered voters in the current election. Bin sizes are 5,000 voter bins. Diameter of the points is proportional to the number of observations. Confidence intervals represent the 95% confidence intervals of a linear regression with standard errors clustered at the municipality level. Note that because all observations to the left of the threshold are 0, there are no standard errors.
Figure C.4  Regression discontinuity plots of pre-treatment population density

(a) Measured prior to the 1988 Constitution

(b) Measured prior to the most recent single-round

Population density measured in the 1980 census (a) or from the census prior to the most recent year in a single-round system or from the 1991 census (b). Bin sizes are 10,000 voter bins. Diameter of the points is proportional to the number of observations. Confidence intervals represent the 95% confidence intervals of a linear regression with standard errors clustered at the municipality level.
Figure C.5  Regression discontinuity coefficients on pre-treatment population density at different bandwidths

(a) Measured prior to the 1988 Constitution

(b) Measured prior to the most recent single-round

Population density measured in the 1980 census (a) or from the census prior to the most recent year in a single-round system or from the 1991 census (b). The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. Estimation method: Local linear regression with a 50,000 voter bandwidth. Standard errors clustered at the municipality level. Source: 1980, 1991, 2000, and 2010 Demographic Census.
Figure C.6  Regression discontinuity coefficients on concentration indices at different bandwidths

(a) Fractionalization

(b) Entropy index

Fractionalization (a) and Entropy index (b) are indices of concentration of candidates’ supporters, using vote counts from polling stations. Vote shares are from the first round. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Figure C.7  Regression discontinuity coefficients on the dispersion in votes at different bandwidths

(a) Dispersion in votes for 1st place candidate

(b) Dispersion in votes for 2nd place candidate

Dispersion in votes for the 1st place (a) and 2nd place (b) candidate is the standard deviation in a candidate’s vote counts across polling stations. Vote shares are from the first round. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Figure C.8  Regression discontinuity coefficients on overall level of resources in municipal schools at different bandwidths

(a) Equipment, mean level of resources

(b) Infrastructure, mean level of resources

*Equipment* and *Infrastructure* are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. *Mean level of resources* is the mean index level across schools in the municipality for equipment (a) and infrastructure (b). The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. *Estimation method*: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Figure C.9  Regression discontinuity coefficients on distribution of resources in municipal schools at different bandwidths

(a) Equipment, standard deviation in resources

(b) Infrastructure, standard deviation in resources

Equipment and Infrastructure are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. Standard deviation in resources is the standard deviation in the index across schools in the municipality for equipment (a) and infrastructure (b). The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table C.1  Regression discontinuity estimates on the geographic dispersion of voters, without controls

<table>
<thead>
<tr>
<th>Panel A: Concentration indices of voters for specific candidates</th>
<th>Fractionalization</th>
<th>Entropy index</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>-0.010**</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.027</td>
<td>0.030</td>
</tr>
<tr>
<td>Observations</td>
<td>264</td>
<td>264</td>
</tr>
</tbody>
</table>

Panel B: Dispersion in vote shares for each candidate

<table>
<thead>
<tr>
<th>1st place candidate</th>
<th>2nd place candidate</th>
<th>3rd place candidate</th>
<th>4th place candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>-0.012*</td>
<td>-0.011</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.008)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.080</td>
<td>0.075</td>
<td>0.042</td>
</tr>
<tr>
<td>Observations</td>
<td>264</td>
<td>264</td>
<td>251</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Panel A: concentration indices (fractionalization and entropy index). Panel B: standard deviation in a candidate’s vote shares across polling stations (for the 1st-4th place candidate). Vote shares are from the first round. Estimation method: Local linear regression with a 50,000 voter bandwidth. Standard errors clustered at the municipality level.
Table C.2  Regression discontinuity estimates on resources in municipal schools, without controls

<table>
<thead>
<tr>
<th></th>
<th>Mean level of resources</th>
<th></th>
<th>Standard deviation in resources</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Equipment</td>
<td>Infrastructure</td>
<td>Equipment</td>
</tr>
<tr>
<td>TwoRound</td>
<td>0.068*</td>
<td>0.036</td>
<td>−0.019*</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.029)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.738</td>
<td>0.731</td>
<td>0.121</td>
</tr>
<tr>
<td>Observations</td>
<td>821</td>
<td>916</td>
<td>821</td>
</tr>
</tbody>
</table>

*\(p < 0.10\), **\(p < 0.05\), ***\(p < 0.01\)

*Equipment* and *Infrastructure* are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. The first two columns (*Mean level of resources*) has as the dependent variable the mean index level across schools in the municipality. The last two columns (*Standard deviation in resources*) has as the dependent variable the standard deviation in the index across schools in the municipality. *Estimation method:* Local linear regression with a 50,000 voter bandwidth. Standard errors clustered at the municipality level.
Table C.3  Placebo regression discontinuity estimates on the geographic dispersion of voters, at 300,000 inhabitant threshold

<table>
<thead>
<tr>
<th></th>
<th>Fractionalization</th>
<th>Entropy index</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.024</td>
<td>0.027</td>
</tr>
<tr>
<td>Observations</td>
<td>471</td>
<td>471</td>
</tr>
</tbody>
</table>

**Panel B: Dispersion in vote shares for each candidate**

<table>
<thead>
<tr>
<th></th>
<th>1st place candidate</th>
<th>2nd place candidate</th>
<th>3rd place candidate</th>
<th>4th place candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>0.001</td>
<td>0.0001</td>
<td>0.001</td>
<td>−0.003</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.075</td>
<td>0.072</td>
<td>0.040</td>
<td>0.021</td>
</tr>
<tr>
<td>Observations</td>
<td>471</td>
<td>471</td>
<td>444</td>
<td>373</td>
</tr>
</tbody>
</table>

*p < 0.10, ** p < 0.05, *** p < 0.01

At 300,000 inhabitants, a 2000 constitutional amendment placing a cap on local legislator salaries comes into effect. Panel A: concentration indices (fractionalization and entropy index). Panel B: standard deviation in a candidate’s vote shares across polling stations (for the 1st-4th place candidate). Vote shares are from the first round. Includes only elections after 2000. Estimation method: Local linear regression with election-year fixed effects and a 125,000 inhabitant bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table C.4  Placebo regression discontinuity estimates on the geographic dispersion of voters, at 285,714 inhabitant threshold

Panel A: Concentration indices of voters for specific candidates

<table>
<thead>
<tr>
<th></th>
<th>Fractionalization</th>
<th>Entropy index</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>0.005</td>
<td>0.007*</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.024</td>
<td>0.027</td>
</tr>
<tr>
<td>Observations</td>
<td>423</td>
<td>423</td>
</tr>
</tbody>
</table>

Panel B: Dispersion in vote shares for each candidate

<table>
<thead>
<tr>
<th></th>
<th>1st place candidate</th>
<th>2nd place candidate</th>
<th>3rd place candidate</th>
<th>4th place candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>0.005</td>
<td>0.003</td>
<td>0.009*</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.075</td>
<td>0.071</td>
<td>0.040</td>
<td>0.022</td>
</tr>
<tr>
<td>Observations</td>
<td>424</td>
<td>423</td>
<td>400</td>
<td>331</td>
</tr>
</tbody>
</table>

*<p < 0.10, **<p < 0.05, ***<p < 0.01

At 285,714 inhabitants, a 2004 constitutional amendment determining the size of the local legislature comes into effect. Panel A: concentration indices (fractionalization and entropy index). Panel B: standard deviation in a candidate’s vote shares across polling stations (for the 1st-4th place candidate). Vote shares are from the first round. Includes only elections after 2004. Estimation method: Local linear regression with election-year fixed effects and a 125,000 inhabitant bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Figure C.10  Regression discontinuity coefficients on concentration indices at different thresholds

Fractionalization (a) and Entropy index (b) are indices of concentration of candidates' supporters, using vote counts from polling stations. Vote shares are from the first round. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
**Figure C.11**  Regression discontinuity coefficients on the dispersion in votes at different thresholds

![Diagram](image)

(a) Dispersion in votes for 1st place candidate

(b) Dispersion in votes for 2nd place candidate

Dispersion in votes for the 1st place (a) and 2nd place (b) candidate is the standard deviation in a candidate’s vote counts across polling stations. Vote shares are from the first round. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Figure C.12  Regression discontinuity coefficients on overall level of resources in municipal schools at different thresholds

(a) Equipment, mean level of resources

(b) Infrastructure, mean level of resources

Equipment and Infrastructure are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. Mean level of resources is the mean index level across schools in the municipality for equipment (a) and infrastructure (b). The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Figure C.13  Regression discontinuity coefficients on distribution of resources in municipal schools at different thresholds

(a) Equipment, standard deviation in resources

(b) Infrastructure, standard deviation in resources

Equipment and Infrastructure are indices constructed by taking the first principal component of a school’s equipment and infrastructure elements, then calculating the school’s percentile in the national distribution. Standard deviation in resources is the standard deviation in the index across schools in the municipality for equipment (a) and infrastructure (b). The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.
Table C.5  Regression discontinuity estimates on the geographic dispersion of voters, with number of candidates as a control

<table>
<thead>
<tr>
<th>Panel A: Concentration indices of voters for specific candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fractionalization</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>TwoRound</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Single round mean</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Dispersion in vote shares for each candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st place candidate</td>
</tr>
<tr>
<td>TwoRound</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Single round mean</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

Panel A: concentration indices (fractionalization and entropy index). Panel B: standard deviation in a candidate’s vote shares across polling stations (for the 1st-4th place candidate). Vote shares are from the first round. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Number of candidates included as a control. Standard errors clustered at the municipality level.
Table C.6  Regression discontinuity estimates on the geographic dispersion of voters, using vote shares from the top two candidates only

<table>
<thead>
<tr>
<th></th>
<th>Fractionalization</th>
<th>Entropy index</th>
<th>Std Dev of 1st place candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwoRound</td>
<td>$-0.018^{**}$</td>
<td>$-0.013^{**}$</td>
<td>$-0.015$</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Single round mean</td>
<td>0.038</td>
<td>0.029</td>
<td>0.088</td>
</tr>
<tr>
<td>Observations</td>
<td>264</td>
<td>264</td>
<td>264</td>
</tr>
</tbody>
</table>

*p < 0.10, ** p < 0.05, *** p < 0.01

*Fractionalization, Entropy index and Standard deviation of 1st place candidate calculated using only vote shares from the top two candidates. Vote shares are from the first round. Estimation method: Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.*