

ELECTORAL RULES, POLITICAL COMPETITION AND FISCAL SPENDING: REGRESSION DISCONTINUITY EVIDENCE FROM BRAZILIAN MUNICIPALITIES[§]

MARCOS CHAMON[†], JOÃO M. P. DE MELLO[‡] AND SERGIO FIRPO[¥]

First Version: March 2008

This Version: January 2009

Abstract

We exploit a discontinuity in Brazilian municipal election rules to investigate whether political competition has a causal impact on policy choices. In municipalities with less than 200,000 voters mayors are elected with a plurality of the vote. In municipalities with more than 200,000 voters a runoff election takes place among the top two candidates if neither achieves a majority of the votes. In a first stage, we show that the possibility of runoff increases political competition. In a second stage, we use the discontinuity as a source of exogenous variation to infer causality from political competition to fiscal policy. Our second stage results suggest that political competition induces more investment and less current spending, particularly personnel expenses. Furthermore, the impact of political competition is larger when incumbents can run for reelection, suggesting incentives matter insofar as incumbents can themselves remain in office.

KEY WORDS: Electoral Systems; Strategic Voting; Political Competition; Regression Discontinuity; Fiscal Spending.

JEL CODES: H72; D72; C14; P1

[§]We would like to thank Helios Herrera, Leonardo Rezende, Filipe Campante, Alexandre Samy, Claudio Ferraz, and seminar participants at EESP-FGV, PUC-Rio and IPEA-RJ for useful comments. Edson Macedo, Mariano Lima, João Pedro Borges and Julia Folescu provided excellent research assistance. Usual disclaimer applies.

[†]International Monetary Fund and BREAD: mchamon@imf.org. The views expressed in this paper are those of the authors and should not be attributed to the International Monetary Fund.

[‡]Departamento de Economia, PUC-Rio: jmpm@econ.puc-rio.br.

[¥]Escola de Economia de São Paulo, FGV: sergio.firpo@fgv.br.

“How an excess of political stability can get in the way of good government”

The Economist, May 17th, 2008 commenting on the mishaps of the *Concertación*, the Chilean long-standing governing coalition.

1. Introduction

It is well established that electoral rules have strong implications for the political process. For example, plurality voting favors a two-party system (“Duverger’s Law”, Duverger, 1954). By affecting party formation, different electoral rules induce different levels of electoral competition. However, the effects of political competition on policy choices are not well understood empirically. It is not surprising that the empirical link from political competition to policy making can be elusive, as the two are simultaneously determined. While barriers to entry can lead to low competition and bad policies, low competition can also coexist with good policies, if for example a highly capable incumbent discourages entry by challengers. Although a growing body of evidence supports the view that competition improves policy making,¹ too little a barrier to entry may lead to instability, fragmentation and worse policies.²

In this paper we explore a unique discontinuity in the rules for Municipal elections in Brazil, which provides a sharp identification of how lower political entry-costs can affect policy outcomes. Our results indicate greater competition improves fiscal policy. Mayoral elections in Brazil take place every four years, with the election rules varying depending on the size of the electorate. Voting is mandatory. In municipalities with more than 200,000 voters elections are in a two-round system. A runoff between the first-round winner and the runner-up takes place if the former receives less than 50% of valid votes. In municipalities with less than 200,000 voters there is only one round with the winner being the one who gains the most votes. The 200,000 threshold rule provides an exogenous and abrupt change in

¹ See, for example, Besley et al (2005), Besley and Case (1995), Rodgers and Rodgers (2000), and Besley and Case (2003).

² See, for example, Campante et al (2008).

the voting system as a function of the electorate size. As long as the electoral rules cause “political market” structure, it can be a source of exogenous variation in the degree of competitiveness of the “political market”. This discontinuity arguably provides the sharpest identification for the causal effect of political competition on fiscal policy outcomes.

The link from electoral rules to political market structure is well established in the literature.³ Most comparisons contrast majority and proportional systems, and why the latter favors multi-party structures. Nevertheless, similar arguments can be made for comparing one-round with two-round majority elections. For example, consider a one-round election and suppose that 60 percent of the electorate is left-leaning. If there is one left-leaning and one right-leaning party contesting the election, the former should easily win. But if there are two competing left-leaning parties, the right-leaning one may be able to achieve a plurality of the vote. In this case, the third candidate would be a “spoiler,” and in a well functioning system the two left-leaning parties should form a coalition and launch a single candidate. In a two-round election, the presence of the third candidate should not affect the final outcome and therefore we should expect a larger supply of candidates under that system.^{4,5}

Methodologically, we use an IV procedure relying on a regression-discontinuity design (RDD)⁶, which is known for its very high internal validity as it exploits the exogenous variation that occurs around the discontinuity point. Thus, it dispenses with concerns about unobserved heterogeneity driving results. We show a discrete and sizeable jump in voting concentration as function of the electorate size and that this jump occurs at the 200,000 voter threshold. That is, there is an abrupt increase in political competition for municipalities where the second round is present. Following the Political Science literature, we measure concentration both as the number of effective candidates as well as the share of votes going

³ For example, Duverger’s Law, which is formally proved in Palfrey (1989).

⁴ Note that the presence of a runoff would not necessarily rule out a right-leaning party victory in this example. Suppose there are four left-leaning parties each of which receives 15 percent of the vote, and two right-leaning parties that receive 20 percent of the vote. Then, a runoff would take place with the two right-leaning parties.

⁵ The presence of a runoff is more likely to affect the outcome of the election when voters choose their first choice of candidate (sincere voting) than under strategic voting. The political economy literature has interest in sincere versus strategic voting because of their different implications for modeling. Some empirical evidence is available, mainly using structural modeling strategies. See Degan and Merlo (2007) and Merlo (2006).

⁶ RDD was introduced by Thistlewaite and Campbell (1960) but has been widely diffused in the empirical economics literature since the work of Van der Klaauw (2002), Angrist and Lavy (1999), Hahn, Todd and Van der Klaauw (2001), among others. See also the recent work of Imbens and Lemieux (2008) for an extensive survey on RDD.

to the third-placed or lower candidates.⁷ Thus, the runoff reduces the “political market concentration” by encouraging more parties to enter and/or inducing sincere voting in the first round, thus turning the political regime more competitive.⁸

Under the identifying assumption that the presence of the second round does not affect policy choices directly, it provides exogenous variation to estimate the causal impact of political competition on policy choices. Although one can always conceive circumstances in the presence of the second will affect, *per se*, policy choices, they are rather far-fetched,⁹

Since population determines the presence of the second round, our identification assumption is more likely to hold around the 200,000 voter threshold. Thus, weights play an important role in our IV strategy by augmenting the importance of observations around the discontinuity, thus increasing the credibility of the identification assumption.¹⁰ In summary, we use a weighted instrumental variables procedure in which: i) the dependent variables are policy outcomes (capital and current expenses and construction of schools); ii) the endogenous regressor is a measure of political competition; iii) the instrument is a dummy variable that equals one if a municipality has more than 200,000 voters and zero otherwise; iv) finally, we give more weight to municipalities whose electoral size is closer to 200,000. Our results indicate that a higher degree of political competition causes more capital spending, less spending in current expenses, and more construction of schools. Most of the

⁷ Starting with Laakso and Taagera (1979), the Political Science literature has used the number of effective parties as their main measure of political competition. The number of effective candidates is the inverse of the sum of squared vote shares, i.e., the inverse of the Herfindahl-Hirschman Index of vote shares multiplied by 10,000.

⁸ While revising this paper we came across Fujiwara (2009), which is an independent paper exploring the same discontinuity in electoral rules. His paper only uses the discontinuity to test sincere voting, unlike our paper which uses the RD design as source of exogenous variation in political competition which is then used to test the effect of differential political competition on policy outcomes.

⁹ One possible scenario is the following: When a second round is necessary, new coalitions are formed with first-round losers and thus additional compromises on policies have to be made. This critique does not threaten our identification strategy because the future cannot cause the past: we use political competition at the end to the mandate to proxy for expected political competition. Nevertheless, we compared fiscal policy in cities in which the election went to the second by a small margin to cities whose election was decided in the first round by a small margin, in the spirit of Lee et al (2004). We found no significant difference.

¹⁰ In fact, controlling for different polynomials of the electoral size at each side of the discontinuity, we show in the Appendix an equivalence result between the weighted IV regression and the Wald estimator using local polynomial regression. Such result was suggested by Imbens and Lemieux (2008, p. 627) and proved for a special case (local constant, rectangular weights) by Hahn, Todd and Van der Klaauw (2001).

estimated reduction in current expenditure takes place through lower payroll spending. To the extent that Brazilian politicians tend to heavily favor current expenditures and underprovide public investment, this shift is assumed to be welfare enhancing (especially given an under-developed infra-structure) and political competition to be beneficial.¹¹ In contrast with previous works, we do not estimate the effect of political competition on the size of government since in most municipalities the vast majority of expenditures are financed by federal and state transfers, which accounted for almost all of the revenues in our main sample of municipalities with 125,000-275,000 voters. Federal transfers are determined by a formula, as a function of the municipality's population and the per capita income in its state. There are no discontinuities with respect to population in our main sample.¹³

Interestingly, our results are much stronger when we only consider races in which the incumbent could run for reelection. Thus, political competition is more beneficial when the incumbent has a higher stake in his or her party's future prospects, a result consistent with both theory and previous empirical studies.¹⁴

Our results are also consistent with previous studies documenting the beneficial effects of political competition at the sub-national level.¹⁵ For example, Besley et al, 2005 use the Voting Rights Act of 1965 to associate an increase in political competition in American Southern states with an improved economic performance measured by income per capita. They show convincingly that the Voting Rights Act of 1965 increased political competition. However, the link from political competition to income per capita remains

¹¹ It is consensual among Brazilian economists (within the mainstream of the profession) that current spending is excessive while public investment is at a sub-optimal level. See, for example, Werneck (2008).

¹³ The last discontinuity in the formula takes place at 156,216 habitants, which involves municipalities smaller than those with 125,000-275,000 voters.

¹⁴ Using US state-level data, Besley and Case (1995) compare the behavior of governors who face a bidding term limit with those who can run for reelections, and find that a bidding term limit affects policy choices. Using Brazilian municipal data, Ferraz and Finan (2008) document that mayors in their second term (who cannot run for reelection) are more corrupt than first-term mayors, who can run for reelection. Finally, a seemingly unrelated result is in Ferreira and Gyourko (2008), who find that party affiliation is not relevant for policy making in US municipalities, which suggests that politicians are constrained by voter preferences and political competition, at least at the local level.

¹⁵ It is possible that, by reducing the probability of remaining in office, higher political competition worsens policies by shortening the incumbent's horizon. Campante et al (2008) show that stability can have non-monotonic effects on the quality of policies, which is empirically supported in their cross-country analysis.

somewhat suggestive, as it is difficult to isolate other factors that may have caused income to rise in the South over a long horizon. In contrast, our paper focuses on policy choices instead of economic outcomes such as income per capita. By focusing on actual policy choices, whose change can be measured in the short-run, we provide a cleaner identification of the effect of political competition to welfare.

It is interesting to relate our results to the findings of cross-country comparisons of electoral rules on fiscal policy outcomes. Persson and Tabellini (2004) show that presidential regimes and majoritarian rules lead to smaller governments than parliamentary regimes and proportional representation. Majoritarian rules also tilt the composition towards less transfer expenditures than proportional representation. This last result was also presented and formalized in Milesi-Ferretti et al (2002). It is difficult to draw comparisons with our setting, since much of these results focused on the distinction between majoritarian and proportional representation, whereas our analysis is limited to the presence of a runoff in a local election. But taking the results at face-value would suggest that increasing political competition through a runoff election can lower current expenditures, whereas increasing political competition through proportional representation (where entry barriers for parties are lower) can have the opposite effect.

The rest of the paper is organized as follows. In section 2 we describe the institutional background and the data used. In section 3 we show some graphical evidence that the 200,000 rule is exogenous, in the sense that a “no-manipulation” condition is satisfied. In section 4, we discuss in detail how our weighted IV regressions exploit the discontinuity in the voting system as a function of the electorate size in order to identify the causal effect of political competition on fiscal spending. In section 5 we present and discuss our main findings. Finally, in section 6, we conclude.

2. Institutional Background and Data Description

In Brazil, runoff elections were introduced by the 1988 Constitution. The system for municipal elections is legislated by article 29, chapter 4. Little hard evidence is available on the motivation behind instituting two-round elections. Anecdotes suggest a desire to ensure “legitimate” outcomes by avoiding the risk that a candidate wins a one-round elections with a

small share of the votes (this Constitution was written at a time when Brazil was transitioning from twenty years of military rule towards becoming a consolidated democracy). The presidential and all gubernatorial elections have two-rounds. The 200,000 threshold for municipal elections was driven by cost considerations. Since voting is mandatory, it is safe to assume that the intensity of political competition will not affect whether or not a municipality is above or below the threshold (which would not be the case if voter registration was voluntary).¹⁶

The first round election takes place sometime in the beginning of October, and the second round sometime between the end of October and early November.¹⁷ Where the election has only one round, it takes place the same day as the first round. The state-level electoral authority is in charge of counting the number of registered voters per city to define where second round may take place. The electoral authority rests with the Electoral Justice System, which is composed of a federal entity, *Tribunal Superior Eleitoral* (TSE), and 27 state entities, the *Tribunais Eleitorais Regionais* (TREs). Although formally a member of the judicial system, the TSE not only judges but also performs executive and legislative tasks. It enacts specific legislation for elections and is co-responsible for the actual execution of the elections (presidential, gubernatorial and mayoral elections). The TREs are responsible for the execution of gubernatorial and mayoral elections. Among the executive tasks are registering voters, resolving litigation among candidates, enforcing electoral legislation, and running the actual voting process. The fact that voters' headcount is done by the state-level TREs dramatically reduces the scope for small municipalities to manipulate their electorate size. Moreover, since voting and voter registration are compulsory in Brazil one would have to orchestrate large scale document fraud to manipulate the municipal-level electorate size, something rather far-fetched.

¹⁶ The electorate is composed of three groups. All citizens between 18 and 64 years are automatically registered, and voting is mandatory for registered voters. Second, between 16 and 18 registration is optional, but voting is mandatory once registered. Finally, voting is optional for registered voters older than 64 years. Besides fines, sanctions for not voting include becoming ineligible for public sector jobs, passport issuance and, more importantly, government transfers.

¹⁷ In 1996, the first round took place on October 3rd, and the second on November 15th. In 2000, it took place on October 1st and October 29th. In 2004, on October 3rd and October 31st.

Election data are published by the *Tribunal Superior Eleitoral* (TSE), the federal-level electoral authority. Election results, as well the number of registered voters, are electronically available for a total of 16,498 first-round races over three election cycles: 1996, 2000 and 2004.¹⁸ The first two-round municipal election took place in 1992 (the first after the 1988 Constitution). Unfortunately, electronic data are not available for 1992.¹⁹

As measures of fiscal policy, we consider four variables: capital, current and payroll expenses as proportions of total spending aggregated over the administration cycle, and the number of schools built net of schools closed throughout the administration cycle. Fiscal data come from the *Secretaria do Tesouro Nacional*, the National Treasury, which is subordinated to the *Ministério da Fazenda*.²⁰ From the *Tesouro* we have annual data on current spending of all Brazilian municipalities for the 1996-2005 period. The number of schools at the municipal level is from the *Censo Escolar*, a universal census of schools conducted annually by the Ministry of Education.

Although the size of the government would also be of interest, the vast majority of expenditures in small Brazilian municipalities are financed by transfers from the federal and state governments. This makes the size of municipal governments almost exogenous to the municipal-level political process.²¹

Finally, Brazilian electoral institutions are such that it is quite difficult to see plausible channels for the electoral rule to have a direct effect on fiscal policies, which makes

¹⁸ The total number of municipalities in Brazil is a little over 5,000. This figure oscillates slightly because of new municipalities, which are normally created by dismembering from another municipality. No municipality around the discontinuity (between 125,000 and 275,000) was dismembered during our sample period, so our results are not affected.

¹⁹ Data is available after the electronic ballot was introduced by Law # 9.100, from 1995 onwards. The 1996 municipal elections were the first to have electronic ballot in the vast majority of races.

²⁰ *Ministério da Fazenda* is the Brazilian equivalent of the Ministry of Finance.

²¹ In our main sample (municipalities between 125,000 and 275,000 voters), overall transfers represent on average roughly 69% of revenue. Taxes and fees amounted to another 18%, and capital revenues were the remaining 13%. Transfers are constitutionally mandated shares on state and federal-level taxes, and are therefore exogenously determined. The two sources of municipal-level sources of income are an urban property tax (IPTU, roughly 4.7% of total income) and a tax on services (ISS, with 5.8% of total income). The former is highly dependent on property values, and the later on economic activity. Although small changes in tax rates are possible, total income on IPTU and ISS are largely not under the control of incumbents. Finally, except for very large municipalities, Brazilian municipalities do not have access to debt markets, arms' length or banking. Thus, the only of capital revenues is the sale of physical assets, which has clear limits.

the rule a source of exogenous variation to estimate the impact of political competition on policy making.²²

The exogeneity of the runoff is confirmed by the actual distribution of electoral size. Figure 1 shows the histogram and kernel density estimate of the electorate size. A significant discontinuity at 200,000 would raise the suspicion that municipalities were manipulating the electorate size. As expected, the histogram shows that the frequency drops almost monotonically with electorate size. The histogram shows a slight drop from bin [186,000 ; 200,000] to bin [200,000 ; 214,000], but it is not particularly pronounced compared with other fluctuations in the figure. Still, given the drop in the histogram, we further investigate the possibility of manipulation by estimating the density below and above the discontinuity point 200,000, a procedure inspired in McCrary (2008).²³

[insert Figure 1 here]

Figure 2A shows a small discontinuity at 200,000, already suggested by the histogram in Figure 1. This tiny discontinuity is neither practically nor statistically significant.²⁴ In Figure 2B, we repeat the procedure at 150,000. The “discontinuity” is larger now, despite the absence of any reason for the electorate distribution to have any discrete change at 150,000.

[insert Figure 2 here]

3. Summary Statistics

Tables 1A-1C contain some relevant descriptive statistics. In Table 1A we can see that the vast majority of municipalities in Brazil are small: half of the 16,674 races occurred

²² We could conjecture only two possible objections to this assumption, none of which seems relevant. One is that, in anticipation to the possibility of a second round, incumbents would invest more to inaugurate public works between rounds. The runoff takes place approximately three weeks after the first round, so this channel seems quite farfetched, particularly since electoral law forbids inauguration for a period before elections. One could also argue that incentives for accepting lobbying money from contractors are higher in two-round elections because one has to finance a longer campaign. However, since rounds are so close in time, the additional campaigning comes at relatively low cost above and beyond that of first-round campaigning. TV advertising is allocated in a centralized manner and is free of charge. Thus, little room is left to spend campaign money between rounds.

²³ The procedure consists of two parts. The first stage estimates the histogram in Figure 1. The second stage consists in estimating two local linear regressions, above and below the discontinuity point. The percentage of observations in each bin is treated as the dependent variable, and the midpoint of the bins as regressors. See McCrary (2008) for further details.

²⁴ We compute the t -statistic based on the test proposed in McCrary (2008).

in municipalities whose electorate was smaller than 7,066 voters. As expected, municipalities with more than 125,000 voters are quite different from the average municipality: years of schooling and income per capita increase with population. It is interesting to note that income inequality within larger cities is not different from the rest of the country. Finally, there is no substantial difference between municipalities to left and to right of the discontinuity, as panels B and C reveal, i.e. the differences around 200,000 are neither large in practice nor statistically significant.²⁵

[insert Table 1 here]

The background of the first stage appears when we also look at Table 1B. The size of electorate and the number of candidates are positively related, which is expected as the size of the political market induces entry. The number of candidates increases considerably around the discontinuity threshold: from an average of 4.67 in municipalities whose electorate is between 125,000 and 200,000, to 5.45 in municipalities with electorate between 200,000 and 275,000. Same pattern arises for the median.

Following the political economy literature,²⁶ we use two different measures of political competition: the number of effective candidates, which is the inverse of the Herfindahl-Hirschman Index (HHI), and the percentage of votes for all candidates except the first and second placed candidates in the first round. The HHI is the sum of squared market shares (in this case, vote shares).. Finally, note the only consider races where at least three candidates ran since there is no reason why the presence of the second round should make any difference if there are one or two first-round candidates.

Electoral competition as measured by the number of effective candidates similarly follows the pattern of number of candidates. We can see that the number of effective candidates increases with electoral size. Around the discontinuity point 200,000, the number of effective candidates increases from 2.65 to 2.81. The same result arises when concentration is measured by the percentage of votes received by the 3rd placed candidate or

²⁵ We checked whether there are discontinuities for income per capita, income inequality or for years of schooling. Our results indicate that the municipalities slightly below and above 200,000 are indeed similar.

²⁶ Starting with Laakso and Taagera (1979), the literature has used the number of effective parties as their main measure of political competition in the Political Science literature.

lower. This percentage rises from 18.35% in municipalities with electorate size between 125,000 and 200,000 to 22.41% in municipalities with electorate size between 200,000 and 275,000. Again, both are significantly higher than in the whole sample (14.71% on average).

Figure 3 provides additional graphical evidence of the behavior of concentration around the discontinuity threshold, which shows that the first stage regression is not weak. Imbens and Lemieux (2008) propose a histogram-type procedure. We construct 8 bins by dividing the [100,000;200,000] interval into five mutually exclusive equal-sized sub-intervals of width 20,000, and by dividing the [200,000;300,000] interval into three intervals: (200,000;225,000], (225,000;255,000] and (255,000;300,000]. The asymmetry is due to the rapidly decreasing number of observations for larger electorate sizes. The larger bin width in the (100,000,200,000] interval guarantees at least 20 observations per bin. For each bin, we compute the average number of effective candidates and % of votes received by the 3rd or lower placed candidates, and attribute this number for all values in the bin width.

[insert Figure 3 here]

While in the [100,000 ; 200,000] interval, the number of effective candidates remains roughly constant around 2.53, it jumps to 2.77 in the (200,000 ; 225,000] interval, and fluctuates around this level over the (200,000,300,000] interval. Results are even stronger for the % of votes received by the 3rd or lower placed candidates. Of course, these are unconditional differences.

Table 1C shows some statistics for four fiscal variables: investment as proportion of total spending, current spending as proportion of total spending, payroll as proportion of total spending and increase (in %) in the number of municipal public schools. We restricted our attention to municipalities in the 125,000-275,000 voters range. In 2004, 83 cities were in this range. While this is a small share of the total number of Brazilian municipalities, the combined population covered in our sample is fairly large, with about 14.5 million voters in 2004, corresponding to 12% of the Brazilian electorate. We divided the sample into four groups: races where the incumbent could run for reelection, below and above the discontinuity; and races where, given the inexistence of the reelection or the impossibility of a reelection (given by the rule), the incumbent could not run for reelection, again below and above the discontinuity.

Before 1997 incumbent mayors, governors and presidents could not run for reelection. In January 1997, Congress amended the Constitution to allow reelection, with at most two consecutive terms. Hence, while incumbent mayors could not run for reelection in 1996, all incumbents could in 2000. In our sample for 2000, 77% of the incumbents actually ran for reelection. In 2004, only 52% of the incumbents were in their first terms, and could run. Out of those, 91% ran for reelection.

We can see from Table 1C that for reelection races, the fiscal variables follow the expected pattern: average investment and investment in education (number of schools) are larger and average current and payroll spending are smaller for municipalities to the right of the discontinuity (200,000-275,000) when compared to municipalities to left of the discontinuity (125,000-200,000). The same pattern, however, cannot be found for municipalities where the mayor could not run for reelection.

4. Weighted Instrumental Variables Regression

We are primarily interested in the parameter β_1 given by the following equation:

$$FISCAL_{it} = \beta_0 + \beta_1 E[POLCOMP_{i\tau} | \tau > t] + \phi(ELECT_{it}) + BC_{it} + \varepsilon_{it} \quad (1)$$

where $FISCAL_{it}$ is a fiscal policy outcome in municipality i at an year t prior to the election year τ ($t < \tau$); $POLCOMP_{i\tau}$ is the level of political competition in the next election that is *expected* by the incumbent when making policy decisions over the administration cycle. Political competition is measured by concentration of vote shares, and β_1 is the causal effect of the expected competition at election year τ on policy outcome variables. $ELECT$ is the size of electorate (number of registered voters). Fiscal policy may change systematically with the city size, and the empirical strategy hinges on using the 200,000 rule. Thus, the inclusion of $\phi(.)$ - a flexible function of electorate - is crucial for identifying causality. Finally, C_{it} is a vector of controls such as year dummies, a polynomial of the number of candidates, and the fragmentation in the city council.²⁷

²⁷ The reason we include the number of candidates and the fragmentation in the city council as controls, is because they are likely to directly affect fiscal policy in our second stage regression. If the threshold rule affects
(continued)

As mentioned previously, the intensity of political competition is likely to be affected by the quality of policies (reverse causation), so $Cov[\varepsilon_{it}, POLCOMP_{it}] \neq 0$, and a simple OLS estimation strategy would fail to recover the causal impact of *POLCOMP* on policy outcomes. Moreover, political competition is measured with error by construction. Ideally, that variable should be defined as the incumbent's *expectation* of how competitive the political environment will be in next election. Unfortunately, that expectation is not observable.²⁸ The alternative often used in the literature on political cycles (which we emulate) is to use actual, realized political competition. In other words:

$$POLCOMP_{it} = E[POLCOM_{it} | \tau > t] + v_{it} \quad (2)$$

where v_{it} is uncorrelated with $E[POLCOM_{it} | \tau > t]$. In this case, we expect that measurement error causes attenuation bias, which would work against finding an impact of political competition on policy choices when OLS is used.

Therefore, we have to estimate (1) by a two-stage least squares procedure. The first stage consists of estimating the expectation of the *actual* *POLCOMP* conditional on *DUM200*, a flexible polynomial of electorate size and other controls:

$$\begin{aligned} E[POLCOMP_r | ELECT_r, DUM200_r, C_r] \\ = \gamma_0 + \gamma_1 DUM200_r + \phi_1(ELECT_r) + HC_r \end{aligned} \quad (3)$$

where r is electoral race, i.e., a municipality in an election year and $\phi_1(.)$ is a flexible polynomial of electorate of electorate that may have different functional forms above and below the threshold point.

Note that the number of candidates increases with electoral size, and concentration falls with the number of candidates. Thus, although descriptive statistics and some visual analysis suggested that the concentration of voting drops around 200,000, we could not be conclusive only with these pieces of evidence given the mechanical relationship between the

these two controls, and the latter are omitted in the second stage, then the error ε would be correlated with our instrument, compromising the identification. All of our results remain similar if these two controls are dropped.

²⁸ It is conceivable that one could use opinion polls during the administration cycle. However, these polls are not conducted at a sufficient number of mid-sized municipalities to implement any quantitative empirical procedure.

number of candidates and the effective number of candidates.²⁹ Equation (3) will thus be useful for checking whether the pattern suggested by Table 1 survives after we control for a polynomial of the number of candidates, a function of the electoral size and other controls.

The outcome equation (Equation 1) can be rewritten using the actual political competition instead of the expected one. We are interested in estimation of the parameter β_1

$$FISCAL_r = \beta_0 + \beta_1 POLCOMP_r + \phi_2(ELECT_r) + BC_r + \zeta_r, \quad (4)$$

where $\phi_2(.)$ is a flexible function of electorate size that may behave differently below and above the threshold value.

In order to identify β_1 , one may proceed in two different ways. The “conventional” one imposes that ζ_r is uncorrelated with $DUM200_r$ (the instrumental variable), which, by its turn, is correlated with $POLCOMP$ given $ELECT_r$ and C_r . A local version of it imposes that such stochastic relationships should hold only in a neighborhood of $ELECT_r = 200,000$.

We follow the local identification strategy as it is more plausible to assume that any non-controlled factor will be randomly assigned to municipalities immediately before and after the 200,000 point. In other words, there is nothing systematic that could affect political competition when we compare municipalities on both sides of the threshold, the only difference being the difference in the voting rule.

Hahn, Todd and van der Klaauw (2001) have shown that local IV assumptions yield the same identification result as imposing two continuity assumptions: Given $ELECT_r$ and C_r , the conditional expectation of $POLCOMP$ and the conditional expectation of $FISCAL$ are continuous at the $ELECT_r = 200,000$. Thus, the only way $FISCAL_r$ can respond to changes in the $E[POLCOMP_r | DUM200_r, ELECT_r, C_r]$ around the discontinuity (fixing C_r and $ELECT_r = 200,000$) is through changes from the left to right of the discontinuity. Of

²⁹ Candidacy decisions may be endogenous to the electoral system. In a two-round system, candidacy is “cheap” in the first-round. In this case, the second-round system induces party fragmentation not only because it induces sincere voting, but because it induces more candidacy. See next section for further discussion.

course, given that we had fixed $ELECT_r = 200,000$, this is a counterfactual exercise that only makes sense in a close neighborhood of the threshold.

In order to obtain a consistent estimator of the causal effect in a regression-discontinuity design, one typically runs non-parametric regressions (local polynomial regressions) just before and after the cutoff point. Intuitively, such procedure works because it heavily exploits the data around the discontinuity, giving lesser importance to observations further away. Thus, an estimation procedure that produces “local” estimates at the discontinuity should put weights that decrease with the distance to the cutoff point, measured along the running variable, $ELECT$.

With that in mind we proceed with estimation of β_1 on the following ways. First we use a simple 2SLS regression for the restricted sample within a window around the 200,000 voters. A 2SLS procedure like that should identify β_1 , as argued by the equivalence result in Imbens and Lemieux (2008) if the window is sufficiently narrow. They argue that (i) setting $\phi_1(\cdot)$ and $\phi_2(\cdot)$ to be linear splines and (ii) discarding data points that are outside a window of size h to the left and to the right of the discontinuity, a unweighted 2SLS will be algebraically equivalent to the local linear Wald estimator proposed by Hahn, Todd and van der Klaauw (2001) if all estimates use the same rectangular kernel and bandwidth h .

We generalize (and prove) Imbens and Lemieux (2008) result to polynomial splines and general weighting functions. We provide in the appendix a generalizing equivalence result for the case that kernel is not necessarily rectangular and the polynomial function may not be linear.

Therefore, our second estimation strategy uses a weighted two-stage least squares procedure that (i) gives weights, W , that decrease as a function of the distance to the cutoff point and (ii) controls for polynomial splines. In fact, $\phi_1(\cdot)$ and $\phi_2(\cdot)$ are functions of $G(k)$, an order k polynomial interacted with the instrument in the following way:³⁰

³⁰ An equivalent procedure to the local linear regression sets $k=1$.

$$G(k) = \begin{bmatrix} ELECT - 200thd \\ (ELECT - 200thd)^2 \\ \vdots \\ (ELECT - 200thd)^k \\ (ELECT - 200thd) * DUM200 \\ (ELECT - 200thd)^2 * DUM200 \\ \vdots \\ (ELECT - 200thd)^k * DUM200 \end{bmatrix}.$$

Thus, our estimator corresponds specifically to the coefficient of *POLCOMP* in a weighted IV regression using *FISCAL* as dependent variable, *POLCOMP* as the endogenous regressor, $G(k)$ and other elements of vector C as exogenous regressors, *DUM200* as the instrumental variable and W as the weighing function. In the appendix we present some equivalent ways to algebraically express $\hat{\beta}_1$, our weighted 2SLS estimator of β_1 . Finally, in the case where the weights are appropriate functions of the sample size, our weighted procedure shares the non-parametric nature of the local polynomial regressions used in Hahn, Todd and van der Klaauw (2001).

5. Results

5.1. First Stage Results

In the last section we argued that consistent estimation of the structural parameter β_1 involved finding exogenous variation in political competition. In section 3 we presented some graphical evidence that this was true. We now present our first stage results.

Table 2 reports results from six regressions each one using two different measures of political competition as dependent variable. We obtain results that are very similar, using either the number of effective candidates or the percentage of votes of all candidates placed third or lower.³¹

³¹ Since we only consider races with more than two candidates, there is no direct link between the number of candidates and the percentage of votes for third and lower placed candidates.

In all of the six regressions, the right-hand side variables are *DUM200*, *ELECT* and its square, number of candidates, squared number of candidates, and year dummies. Column (1) reports the first stage for all races, using no weights. We can see that using the whole sample, it seems that there is nothing special with the threshold of 200,000 in terms of inducing political competition. Even in column (2), restricting the sample to the interval [125,000;275,000], we still get no statistical significant impact below and after the threshold, once we control for a quadratic function of the electorate size. In all subsequent specifications we restrict the sample to the interval [125,000;275,000].

[insert Table 2 here]

Results dramatically change when, as appropriate, we introduce weights. We tried two different types of weights W_1 and W_2 , which yield similar results. Our first weighting scheme weighs observations by the inverse of the absolute distance to the cutoff point. Specifically, $W_1 = (|ELECT_r - 200thd|)^{-1}$. An alternative weighing scheme is a Gaussian

kernel: $W_2 = \exp\left(-\left(\frac{ELECT_r - 200thd}{h}\right)^2\right)$, where h was chosen to be 15,000. Column (3)

reports results with W_1 . We can see that the effect of “crossing” the threshold is now significant and with the correct sign: existence of a two-round system increases political competition by decreasing the number of effective candidates and increasing the vote share of third and lower candidates. The 200,000 rule induces a reduction of 34% in the HHI of vote concentration and a 132 percent increase in the proportion of votes given to the third and lower placed candidates. When evaluated at the mean of these dependent variables, the effects correspond to 1.46 and 0.50 standard deviation changes, respectively.

In column (4) we use a higher-order polynomial of electorate size and obtain similar results in terms of statistical significance. In column (5), in addition to the specification of column (3), we include the interaction of *DUM200* with *ELECT* and its square.³² This specification (for the case of $k=2$) was the focus of an extensive discussion in the last section.

³² In column (5) the interaction is actually between *DUM200* and a quadratic function of centered (at 200,000) *ELECT*.

It is important to notice two aspects of that column. First, the coefficient of *DUM200* remains statistically significant, for both dependent variables. Second, in both models, the gain from introducing these interaction terms seems to be relatively small, which can be seen after a simple comparison between R-squared's from column (5) and column (3). Nevertheless, we can reject the null hypothesis that the interactions of *DUM200* with *ELECT* and its square are jointly insignificant.

Finally, in column (6) we test for the relevance of the instrument. We can see at column (6) that exclusion of *DUM200* from column (3) impacts negatively and soundly on the R-squared. This constitutes piece of evidence for instrumental variable relevance.

5.2. Second Stage Results

In the previous subsection we showed that *DUM200* (a dummy for municipalities with electorate above 200,000) increases political competition, that is, we have a strong first-stage regression. Under the identifying assumption that *DUM200* only impacts fiscal policy through its effect on political competition, *DUM200* is a source of exogenous variation to estimate β_1 .³³ Therefore we estimate equation (4) using *DUM200* as an instrument.³⁴

Three fiscal dependent variables are considered: the log of the share in total expenditures of investment, of current spending, and of payroll expenditures. Since yearly data is rather noisy, the dependent variables are the total share over the administration cycle. Additionally, we also measure the impact of political competition on physical investment in schools, measured as the change in the number of schools. For the fiscal variables, three cycles are considered: 1993-1996, 1997-2000 and 2001-2004. Unfortunately, data on the number of schools are only available for the 1997-2000 and 2001-2004 cycles. Finally,

³³ The municipalities do not bear the costs of the election, so there is no reason why crossing the 200,000 voter threshold should impact fiscal policy other than through political competition.

³⁴ In fact, we are also using the dummy *DUM200* interacted with electorate and squared electorate as instruments. The reasons why these two terms appear only on first stage and not on second stage are that (i) we cannot reject the null hypothesis that their respective associated coefficients are both null at the second stage and (ii) as we will see on Table 5, their inclusion to the second stage enlarges standard errors of the coefficient of interest.

political competition is measured by the two first-stage concentration measures previously presented: the Herfindahl-Hirschman Index and the share of votes for the third placed candidate or lower.

Among the controls, we include a set of year dummies, quadratic functions of the number of candidates and of the electorate size and the Herfindahl-Hirschman Index of concentration in the city council. Councilors are elected by direct ballot in an open-list proportional system. The race takes place concurrently with the mayor race, and evidence suggests that the presence of mayoral candidate may “pull” votes for her party’s councilors (see Samuels, 2000).³⁵ In this case, we might expect fragmentation in the city council to jump at 200,000 voters, with consequences for our identification strategy.³⁶ The macro literature on fiscal policy and party structure suggests that fiscal adjustment is harder when the parliament is fragmented (e.g., Milesi-Ferretti et al, 2002, and Persson and Tabellini, 2004).

Table 3 presents the results of our weighted IV procedure. We report the coefficients associated with the endogenous regressor for two subsamples: races in which the mayor could run and races in which reelection was prohibited by legislation. In Panel A we use log the number of effective candidates as endogenous regressor and in Panel B we use log of share of votes for the third placed candidate or lower. In all procedures observations are clustered at the state level because unobserved shocks to fiscal revenue are likely to be correlated among cities within a certain state.³⁷

[insert Table 3 here]

In races in which the mayor could run for reelection (panel A), a 1% increase in the number of effective candidates causes a statistically significant increase of 0.889% of investment as a proportion of total expenses. Results are similar when the endogenous regressor is log of the percentage of votes received by the 3rd placed candidates or lower. A 1% increase in the proportion of votes given to the third and lower placed candidates (increase of political competition) causes a statistically significant increase of 0.157% of

³⁵ Technically, the relevant unit for the list is the coalition.

³⁶ Indeed, it does seem that the possibility of the second round reduces the HHI at the city council by some 8%, as expected. However, the impact is not statistically significant. Results are available upon request.

³⁷ One of the main sources of revenue is the participation funds on *Imposto de Circulação de Mercadorias e Serviços* (ICMS), a state-level value-added tax.

investment as a proportion of total expenses. It is important to emphasize that for races in which the incumbent mayor could not run there is no evidence of the same phenomenon, which seems strong evidence of political competition on politicians' behavior.

One possible concern is whether investment is the type of expenditure with most electoral appeal. It may be that, as a response to an increase in competition, incumbents would increase transfers, which are part of current spending. We cannot fully address this concern because we do not have data on transfers. However, municipal transfers are very small relative to federal transfers. We do have data on payroll spending, which does not directly benefit most voters.³⁸ When the share of payroll spending in total expenditures is used as the dependent variable, the results show that the reduction in current spending is due to a reduction in personnel spending. The magnitude of the estimated impact of competition is an order of magnitude larger for payroll than for ordinary current expenses. Again, the results arise only for the sample of reelection races.

In races in which the mayor could run, a 1% increase in the number of effective candidates causes a statistically significant increase of 0.579% of payroll as a proportion of total expenses; and a 1% increase in the proportion of votes given to the third and lower placed candidates (increase of political competition) causes a statistically significant decrease of 0.108% of payroll as a proportion of total expenses.

Finally, we further investigate the nature of the investments made. We have data on the number of municipal schools built and closed down during the administration cycle, which we also use as a dependent variable in Table 3. The results suggest that increases in political competition increase investment in education, at least as measured by physical assets.³⁹ The same pattern for reelection and non-reelection data persists: the impact of political competition on schools built arises only in the sample of reelection races.

Table 4 shows estimates of the same models as in table 3, except that $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ are now linear, not quadratic splines. Results are quite similar to those in table 3, which

³⁸ This is certainly true for the middle-sized municipalities in our sample. In small poor cities public sector payroll can represent a substantial proportion of local income, and thus it is hard to argue that payroll spending does not have electoral appeal.

³⁹ It would be interesting to see how payroll spending on teachers behaves as a function of political competition. Unfortunately only total payroll spending is only available.

shows their robustness to the particular form of $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$, as long as we allow some flexibility.

[insert Table 4 here]

Table 5 shows several robustness tests for the quadratic model, our preferred model. For conciseness we present only the results for the sample of reelection races.⁴⁰ Panel A contains tests that concern the form of the functions $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$. First, we include the quadratic spline in the second stage, i.e., the function $\varphi_2(\cdot)$ is a quadratic spline at 200,000 voters. Although we lose statistical significance, estimated coefficients are quite similar to those in Table 3, showing that the exclusion of the interactions between *ELECT* and squared *ELECT* with *DUM200* in the second stage does not cause bias. Second, we run a parametric model in which both $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ are a sixth degree polynomial of electorate size, but with no splines. Results are similar to those using a quadratic function (Table 3), both in terms of estimated coefficients and statistical significance. This confirms that estimates are not sensitive to the particular form of $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$, as long as we allow some flexibility.

[insert Table 5 here]

We use a weighting procedure that puts a significant weight on observations close to 200,000. The advantage of this procedure is that we emulate as best we can the ideal experiment of comparing cities around the discontinuity (199,999 and 200,001 voters). The reduced number of observations and the weighting schemes that weigh very heavily observations close to 200,000 may raise concern about outliers. For this reason we present robustness checks in which the weights are not used, and the number of observations is expanded to include electorate sizes between 100,000 and 300,000.⁴¹

In panel B we assess whether results are sensitive to the weighting procedure. First, we implement a Gaussian weighting scheme, which also weighs more heavily observations around 200,000 but less so than our scheme. Results are similar. Then, we do not weigh observation at all, treating all cities in the [125,000; 275,000] interval equally. Except for investment, no estimated coefficient is statistically significant now. However, this is driven

⁴⁰ Estimates for the subsample of non-reelection races are never significant.

⁴¹ One could further expand the sample. However, this would only exacerbate the problems raised by comparing cities of very different sizes.

by reduced precision: the estimated coefficients are *larger* than their weighted counterparts. Thus, weighting heavily observations around 200,000, if anything, reduces the point estimates for the impact of political competition on policy choices. Panel C shows that our results are not particular to the [125,000; 275,000] sample. We get similar results when the sample is expanded to [100,000 ; 300,000], which increases the number of observations by 30%.

6. Conclusion

This paper exploited a discontinuity in Brazilian electoral rules to show that runoff elections are associated with more candidates and sharper political competition than majoritarian elections. This result is in line with a large body of theoretical and empirical evidence on electoral rules and electoral competition. An important contribution of our paper is to exploit a quasi-natural experiment that exogenously changes the electoral rule. Thus, among the existing papers on this subject, our design arguably provides the cleanest identification setup for capturing the effect of electoral rules on electoral competition.

Our most interesting result, however, is related to the effect of lower entry costs for political competition on fiscal outcomes. In theory that effect can be ambiguous and lower entry costs may improve or worsen fiscal policy. Also, incumbent politicians can make policy choices that directly affect political competition, which could create a reverse causality problem.

However, by taking advantage of the discontinuity in the electoral rule as a function of the electorate size, we can unequivocally identify the causal effect of political competition on fiscal outcomes. Our results suggest that lower political entry costs shift public expenditures from current expenditures towards investment, which can be perceived as welfare improving.

Despite the sharp identification provided by the discontinuity we explore, there are valid concerns relating to external validity. It is likely that the net effect depends on the particular features of the setting. For example, higher competition likely affects young and consolidated democracies differently, and municipalities have less “fiscal levers” to “play

with” than national governments. But with these caveats in mind, this paper does suggest that lower costs of political entry in a multi-party democracy are beneficial.

References

Angrist, J. and V. Lavy, “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement,” *Quarterly Journal of Economics*, Vol. 114, pp. 533–576, 1999.

Besley, T. and A. Case, “Incumbent Behavior: Vote-Seeking, Tax-Setting and Yardstick Competition,” *American Economic Review*, Vol. 85, pp. 25-45, 1995.

Besley, T. and A. Case, “Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits,” *Quarterly Journal of Economics*, Vol. 110, pp. 769-798, 1995.

Besley, T. and A. Case, “Political Competition and Policy Choices: Evidence from the United States,” *Journal of Economic Literature*, Vol. 41, pp. 7-73, 2003.

Besley, T., T. Persson and D. Sturm, “Political Competition and Economic Performance: Theory and Evidence from the United States,” NBER Working Paper Series 11484, 2005.

Black, S., “Do Better Schools Matter? Parental Valuation of Elementary Education,” *Quarterly Journal of Economics*, Vol. 114, pp. 577–599, 1999.

Campante, F., D. Chor and Q. Do, “Instability and the Incentives for Corruption,” SSRN Working Paper, 2008.

Degan, A. and A. Merlo, “A Structural Model of Turnout and Voting in Multiple Elections,” PIER Working Paper 07-011, 2007.

Duverger, Maurice *Political Parties*, New York: Wiley, 1954.

Fan, J., Gijbels, I., *Local Polynomial Modeling and its Applications*, Chapman & Hall, London, 1996.

Ferraz, C. and F. Finan, “Exposing Corrupt Politicians: The Effects of Brazil’s Publicly Released Audits on Electoral Outcomes,” *Quarterly Journal of Economics*, Vol. 123, pp. 703-745, 2008.

Ferreira, Fernando and Joseph Gyourko, “Do Political Parties Matter? Evidence from US Cities,” Forthcoming, *Quarterly Journal of Economics*.

Fujiwara, T. “Does Plurality Rule Favors the Two Party System? A Regression Discontinuity Test of Strategic Voting and Duverger’s Law”, Unpublished manuscript, 2009.

Hahn, J., P. Todd and W. Van der Klaauw, “Identification and Estimation of Treatment Effects with Regression Discontinuity Design,” *Econometrica*, Vol. 69, pp. 201-209, 2001.

Imbens, G. and T. Lemieux, "Regression Discontinuity Designs: a Guide to Practice," *Journal of Econometrics*, Vol. 142, pp. 615-635, 2008.

Jones, M. *Electoral Laws and the Survival of Presidential Democracies*, Notre Dame: University of Notre, 1995.

Laakso, M. and Taagepera, R. "The "Effective" Number of Parties: a Measure with Application to Western Europe," *Comparative Political Studies*, Vol 12, pp. 3-28, 1979.

Lee, D., "Randomized Experiments from non-Random Selection in U.S. House Elections," *Journal of Econometrics*, Vol. 142, pp. 675-697, 2008.

Lee, D., E. Moretti and M. Butler, "Do Voters Affect or Elect Policies? Evidence from the US House," *Quarterly Journal of Economics*, Vol. 103, pp. 807-859, 2004.

Lemieux, T. and M. Milligan, "Incentive Effects of Social Assistance: a Regression Discontinuity Approach," *Journal of Econometrics*, Vol. 142, pp. 807-828, 2008.

McClintock, C., "Plurality versus Runoff Rules for the Election of the President in Latin America: Implications for Democracy," *Annals of the Meeting of the American Political Science Association*, 2007.

McCrary, J. "Manipulation of the Running Variable in the Regression Discontinuity Design: a Density Test," *Journal of Econometrics*, Vol. 142, pp. 698-714, 2008.

Merlo, A., "Whither Political Economy? Theories, Facts and Issues," in *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, eds. Blundell, R., W. Newey and T. Persson, Cambridge: Cambridge University Press, 2006.

Milesi-Ferretti, R. Perotti and M. Rostagno, "Electoral Systems and Public Spending," *Quarterly Journal of Economics*, Vol. 107, pp. 609-657, 2002.

Palfrey, T. "A Mathematical Proof of Duverger's Law," in *Models of Strategic Choice in Politics*, ed. Ordeshook, P., Ann Arbor: The University of Michigan Press, 1989.

Persson, T. and G. Tabellini, "Constitutional Rules and Fiscal Policy Outcomes," *American Economic Review*, Vol. 94, p. 25-45, 2000.

Rodgers, D. and J. Rodgers, "Political Competition and State Government Size: Do Tighter Produce Looser Budgets?" *Public Choice*, Vol. 105, pp. 1-20, 2000.

Rogoff, K., "Equilibrium Political Budget Cycles," *American Economic Review*, Vol. 80, pp. 21-36, 1990.

Samuels, D., "The Gubernatorial Coattails Effect: Federalism and Congressional Elections in Brazil," *The Journal of Politics*, Vol. 62, pp. 240-253, 2000.

Thistlewaite, D. and D. Campbell, "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment," *Journal of Educational Psychology*, Vol. 51, pp. 309-317, 1960.

Van Der Klaauw, W., "Estimating the effect of financial aid offers on college enrollment: a regression-discontinuity approach," *International Economic Review*, Vol. 43, pp. 1249-1287, 2002.

Werneck, R. "The Rocky Road to Countercyclical Fiscal Policy in Brazil", Manuscript, PUC-Rio, Brazil, Department of Economics.

Table 1.A, Descriptive Statistics for Municipal Election Races: City Demographics

	Mean	Std Dev	Median	Num. Obs.
Panel A: All Races				
Size of Electorate	19,036	114,228	7,028	16,674
Income per Capita ^a	169.23	97.37	157.63	16,674
Gini Coefficient	0.56	0.08	0.55	16,674
Average Years of Schooling ^b	4.00	1.34	4.05	16,674
Panel B: Races With 125,000-200,000 Voters				
Size of Electorate	157,523	22,446	154,090	140
Income per Capita ^a	321.73	112.89	308.79	140
Gini Coefficient	0.56	0.05	0.56	140
Average Years of Schooling ^b	6.45	0.80	6.48	140
Panel C: Races With 200,000-275,000 Voters				
Size of Electorate	233,086	22,953	229,751	54
Income per Capita ^a	382.36	143.07	369.64	54
Gini Coefficient	0.56	0.05	0.56	54
Average Years of Schooling ^b	6.96	0.93	6.99	54

Source: Tribunal Superior Eleitoral, Secretaria do Tesouro Nacional and Instituto Brasileiro de Geografia e Estatística.

^a In 2000 reais

^b Years of schooling for the population between 15 and 64 years old.

Table 1.B, Descriptive Statistics for Municipal Election Races: First Stage Variables

	Mean	Std Dev	Median	Num. Obs.
Panel A: All Races				
Number of Candidates	2.79	1.12	2.00	16,500
Effective Number of Candidates ^a	2.51	11.83	2.40	8,144
Share of Votes of Third Placed and Lower ^b	14.71	11.83	12.40	8,144
Panel B: Races With 125,000-200,000 Voters				
Number of Candidates	4.67	1.42	4.00	134
Effective Number of Candidates ^a	2.65	0.60	2.50	134
Share of Votes of Third Placed and Lower ^b	18.35	11.85	16.66	134
Panel C: Races With 200,000-275,000 Voters				
Number of Candidates	5.45	1.58	5.00	55
Effective Number of Candidates ^a	2.81	0.70	2.67	55
Share of Votes of Third Placed and Lower ^b	22.41	13.48	21.66	55

Source: Tribunal Superior Eleitoral, Secretaria do Tesouro Nacional and Instituto Brasileiro de Geografia e Estatística

^a The number of effective candidates is the inverse of the Herfindahl-Hirschman Index (HHI) multiplied by 10,000. Only races with more than 2 candidates included

^b Percentage of votes received by candidates placed third or lower in the first round. Only races with more than 2 candidates included

Table 1.C, Descriptive Statistics for Municipal Election Races: Second Stage Variables

	Mean	Std Dev	Median	Num. Obs.
Panel A: Races With 125,000-200,000 Voters, Reelection Races				
Investment as % of Total Spending ^c	12.88	6.67	11.13	59
Current Spending as % of Total Spending ^c	75.13	8.38	77.02	58
Payroll Spending as % of Total Spending ^c	52.51	14.88	49.59	58
% Change in Number of Schools	11.37	21.72	6.87	74
Panel B: Races With 200,000-275,000 Voters, Reelection Races				
Investment as % of Total Spending ^c	13.01	6.64	11.11	22
Current Spending as % of Total Spending ^c	74.86	9.56	77.44	21
Payroll Spending as % of Total Spending ^c	50.57	16.21	46.72	21
% Change in Number of Schools	14.83	34.43	9.05	28
Panel C: Races With 125,000-200,000 Voters, Non-Reelection Races				
Investment as % of Total Spending ^c	17.16	6.93	17.13	52
Current Spending as % of Total Spending ^c	69.76	12.04	70.06	52
Payroll Spending as % of Total Spending ^c	47.73	17.38	44.07	40
% Change in Number of Schools	-0.73	9.67	0.00	25
Panel D: Races With 200,000-275,000 Voters, Non-Reelection Races				
Investment as % of Total Spending ^c	15.50	6.53	14.84	23
Current Spending as % of Total Spending ^c	71.15	9.25	73.83	23
Payroll Spending as % of Total Spending ^c	53.84	13.82	50.15	23
% Change in Number of Schools	3.08	16.44	3.20	13

Source: Tribunal Superior Eleitoral, Secretaria do Tesouro Nacional, Instituto Brasileiro de Geografia e Estatística and Ministério da Educação

^a Percentage of votes received by candidates placed third or lower in the first round. Only elections with more than 2 candidates included

^b Herfindahl-Hirschman Index, the sum of the squares of the voting shares of all candidates times 10,000. Only elections with more than 2 candidates included

^c Only positive values includes (zeros excluded)

Table 2. OLS Regressions For Vote Share Concentration

	All Races	125,000- 275,000	125,000- 275,000	125,000- 275,000	125,000- 275,000	125,000- 275,000
Panel A: Dependent Variable: Log(Number of Effective Candidates)						
	(1)	(2)	(3) ^(a)	(4) ^(b)	(5) ^{(a) (c)}	(6) ^(a)
Dummy For 200,000 or More Voters	-0.029 (0.033)	0.074 (0.051)	0.342 (0.027)***	0.15 (0.060)**	0.386 (0.030)***	-
Electorate Size	-2.16E-07 (6.79E-08)***	5.43E-07 (3.08E-06)	5.64E-06 (4.43E-06)	0.002 (0.006)	1.01E-05 (2.16E-06)***	1.21E-06 (3.98E-06)
Electorate Size Squared	3.46E-14 (9.23E-15)***	-4.44E-12 (8.50E-12)	-2.57E-11 (1.28E-11)*	-3.32E-08 (8.11E-08)	-1.20E-10 (3.39E-11)***	-3.95E-12 (9.87E-12)
Number of Candidates	0.207 (0.013)***	0.201 (0.059)***	0.181 (0.080)**	0.199 (0.055)***	0.157 (0.084)*	0.368 (0.117)***
Number of Candidates Squared	-0.011 (0.001)***	-0.01 (0.005)*	-0.011 (0.008)	-0.01 (0.004)***	-0.009 (0.008)	-0.024 (0.011)**
Quadratic Spline	No	No	No	No	Yes	No
R ²	0.253	0.308	0.644	0.317	0.669	0.399
Number of Observations	8092	187	187	187	187	187
Panel B: Dependent Variable: Log(100-Vote Share of Top Two Candidates)						
Dummy For 200,000 or More Voters	-0.094 (0.205)	0.273 (0.323)	1.316 (0.295)***	0.584 (0.332)*	1.5 (0.300)***	-
Electorate Size	-4.17E-07 (2.36E-07)*	-4.61E-06 (8.17E-06)	-9.29E-06 (2.89E-05)	0.007 (0.020)	5.93E-05 (2.10E-05)***	-2.63E-05 (2.75E-05)
Electorate Size Squared	1.13E-13 (3.73E-14)***	2.83E-12 (2.07E-11)	-2.17E-11 (6.95E-11)	-1.14E-07 (2.69E-07)	-7.09E-10 (2.47E-10)***	6.18E-11 (6.86E-11)
Number of Candidates	1.222 (0.085)***	0.96 (0.218)***	1.487 (0.801)*	0.952 (0.207)***	1.372 (0.791)*	2.206 (0.912)**
Number of Candidates Squared	-0.076 (0.008)***	-0.059 (0.017)***	-0.105 (0.064)	-0.059 (0.016)***	-0.097 (0.063)	-0.156 (0.074)
Quadratic Spline	No	No	No	No	Yes	No
R ²	0.137	0.305	0.617	0.313	0.635	0.475
Number of Observations	8092	187	187	187	187	187

Notes: Only races with more than 2 candidates considered. Standard errors in parenthesis, and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. All regressions include year dummies and the HHI index of votes for the city council.

^(a): Weighted Least Squares. Weight = 1/[200,000 - Electorate]. ^(b): 6-degree polynomial of electorate-200,000 included. ^(c): The terms involving electorate are centered around 200,000 (electorate minus 200,000 voters).

Table 3. Second Stage Regressions of Political Competition on Fiscal Outcomes in Municipalities with 125,000-275,000 Registered Voters

Panel A: Explanatory Variable is Log(Number of Effective Candidates)				
Fiscal Outcome	Sample	Effect of Political Competition		Num. Obs.
Log(Investment/Expenditures)	Reelection Races	0.889	(0.366)**	76
	Non-Reelection Races	0.08	(2.874)	73
Log(Current Expenditures/Expenditures)	Reelection Races	-0.124	(0.081)	74
	Non-Reelection Races	-0.597	(0.904)	73
Log(Payroll/Expenditures)	Reelection Races	-0.579	(0.185)***	74
	Non-Reelection Races	-1.357	(1.694)	73
Percentage Change in Number Schools	Reelection Races	93.148	(8.845)***	97
	Non-Reelection Races	-47.086	(288.238)	35
Panel B: Explanatory variable is Log(Percentage of Votes of Candidates Placed Third and Lower)				
Fiscal Outcome	Sample	Effect of Political Competition		Num. Obs.
Log(Investment/Expenditures)	Reelection Races	0.157	(0.086)*	76
	Non-Reelection Races	0.004	(0.241)	73
Log(Current Expenditures/Expenditures)	Reelection Races	-0.019	(0.019)	74
	Non-Reelection Races	0.061	(0.137)	73
Log(Payroll/Expenditures)	Reelection Races	-0.108	(0.062)*	74
	Non-Reelection Races	-0.133	(0.167)	73
Percentage Change in Number Schools	Reelection Races	19.571	(10.297)*	97
	Non-Reelection Races	-3.445	(8.173)	35

Notes: Regressions weighted by the inverse of the distance of electorate size to 200,000. Standard errors in parenthesis, and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. All regressions include year dummies, the log of the HHI in the city council, a second degree polynomial of the number of candidates and a second degree polynomial of the size of the electorate. Instruments include Dummy for Electorate above 200,000, electorate size, electorate size squared, and the dummy for Electorate above 200,000 interacted with electorate - 200,000 and its square. Standard errors clustered at the state level. Sample restricted to races with 125,000-275,000 voters and at least 3 candidates.

Table 4. Second Stage Regressions of Political Competition on Fiscal Outcomes in Municipalities with 125,000-275,000 Registered Voters

Linear Spline				
Model 1: Explanatory Variable: Log(Number of Effective Candidates)				
Fiscal Outcome	Sample	Effect of Political Competition		Num. Obs.
Log(Investment/Expenditures)	Reelection Races	0.917	(0.258)***	76
	Non-Reelection Races	3.448	(27.437)	73
Log(Current Expenditures/Expenditures)	Reelection Races	-0.043	(0.117)	74
	Non-Reelection Races	4.813	(43.764)	73
Log(Payroll/Expenditures)	Reelection Races	-0.544	(0.178)***	74
	Non-Reelection Races	-3.581	(24.237)	73
Percentage Change in Number Schools	Reelection Races	98.727	(9.795)***	97
	Non-Reelection Races	97.721	(313.315)	35
Model 2: Log(Percentage of Votes of Candidates Placed Third and Lower)				
Fiscal Outcome	Sample	Effect of Political Competition		Num. Obs.
Log(Investment/Expenditures)	Reelection Races	0.142	(0.057)**	76
	Non-Reelection Races	-0.001	(0.590)	73
Log(Current Expenditures/Expenditures)	Reelection Races	-0.001	(0.020)	74
	Non-Reelection Races	0.189	(0.267)	73
Log(Payroll/Expenditures)	Reelection Races	-0.088	(0.049)*	74
	Non-Reelection Races	-0.136	(0.308)	73
Percentage Change in Number Schools	Reelection Races	17.462	(9.107)*	97
	Non-Reelection Races	0.618	(12.486)	35

Notes: Regressions weighted by the inverse of the distance of electorate size to 200,000. Standard errors in parenthesis, and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. All regressions include year dummies, the log of the HHI in the city council, a second degree polynomial of the number of candidates and the size of the electorate. Instruments include Dummy for Electorate above 200,000, electorate size, and the dummy for Electorate above 200,000 interacted with electorate - 200,000. Standard errors clustered at the state level. Sample restricted to races with 125,000-275,000 voters and at least 3 candidates.

Table 5. Second Stage Regressions of Political Competition on Fiscal Outcomes in Municipalities
with 125,000-275,000 Registered Voters

Robustness Checks, Model 1: Explanatory Variable: Log(Number of Effective Candidates)				
Panel A: Different Forms for the $\varphi_1()$ and $\varphi_2()$ functions				
Fiscal Outcome	Sample	Effect of Political Competition		Num. Obs.
Log(Investment/Expenditures)	Spline 1 st Stage ^(a)	0.778	(0.905)	76
	Parametric ^(b)	0.763	(0.235)***	76
Log(Current Expenditures/Expenditures)	Spline 1 st Stage ^(a)	-0.426	(0.253)	74
	Parametric ^(b)	-0.095	(0.047)*	74
Log(Payroll/Expenditures)	Spline 1 st Stage ^(a)	-0.486	(0.555)	74
	Parametric ^(b)	-0.503	(0.107)***	74
Percentage Change in Number Schools	Spline 1 st Stage ^(a)	72.567	(23.049)***	97
	Parametric ^(b)	73.726	(30.001)***	97
Panel B: Benchmark Quadratic Spline Model with Gaussian and no weights				
Fiscal Outcome	Sample	Effect of Political Competition		Num. Obs.
Log(Investment/Expenditures)	Gaussian ^(c)	1.331	(0.550)**	76
	No-weight	1.647	(0.753)	76
Log(Current Expenditures/Expenditures)	Gaussian ^(c)	-0.137	(0.102)	74
	No-weight	-0.199	(0.299)	74
Log(Payroll/Expenditures)	Gaussian ^(c)	-0.442	(0.190)**	74
	No-weight	-0.789	(0.905)	74
Percentage Change in Number Schools	Gaussian ^(c)	61.075	(18.779)***	97
	No-weight	98.123	(91.672)	97
Panel C: Expanded Sample: 100thd-300thd				
Fiscal Outcome	Sample	Effect of Political Competition		Num. Obs.
Log(Investment/Expenditures)	Benchmark	0.889	(0.279)***	111
	Bechmark no-weight	2.447	(0.883)***	110
Log(Current Expenditures/Expenditures)	Benchmark	-0.077	(0.048)	109
	Bechmark no-weight	-0.135	(0.265)	109
Log(Payroll/Expenditures)	Benchmark	-0.490	(0.111)***	109
	Bechmark no-weight	-0.734	(0.529)	109
Percentage Change in Number Schools	Benchmark	72.784	(1.321)***	144
	Bechmark no-weight	120.878	(157.896)	144

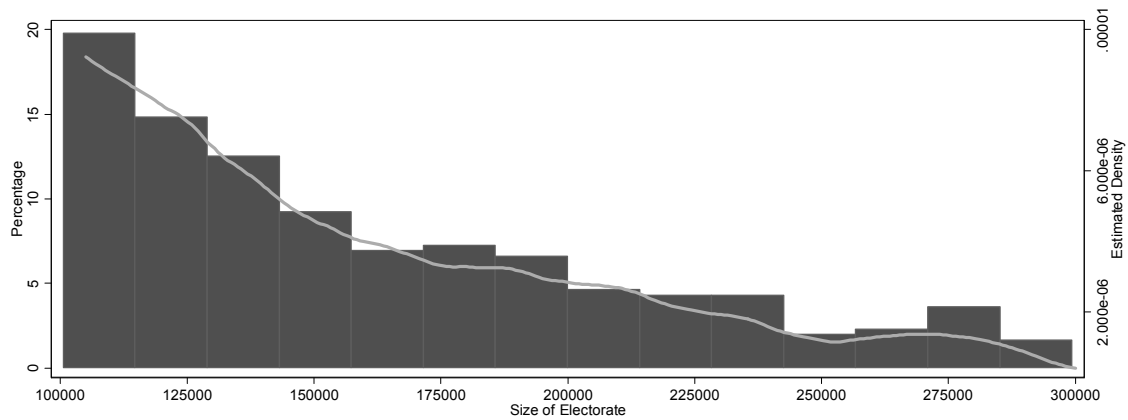
Notes: Regressions weighted by the inverse of the distance of electorate size to 200,000. Standard errors in parenthesis, and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. All regressions include year dummies, the log of the HHI in the city council, a second degree polynomial of the number of candidates. Benchmark model as in table 3 with the sample restricted to races with 125,000-275,000 voters and at least 3 candidates.

(a): Second stage also includes the dummy for Electorate above 200,000 interacted with electorate and its square.

(b): Only dummy for electorate above 200,000 used as an instrument. First and second stages include a polynomial of degree 6 of electorate

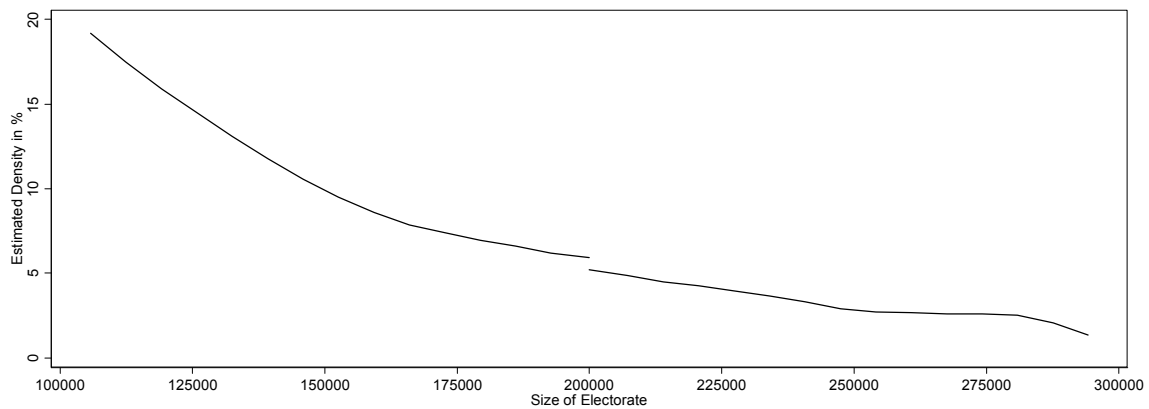
(c): Weight = $\exp(-(Electorate-200,000)/12,500)^2$

Figure 1. Distribution of Electorate Size
Histogram and Estimated Density of Electorate in 100,000-300,000 Voter Range



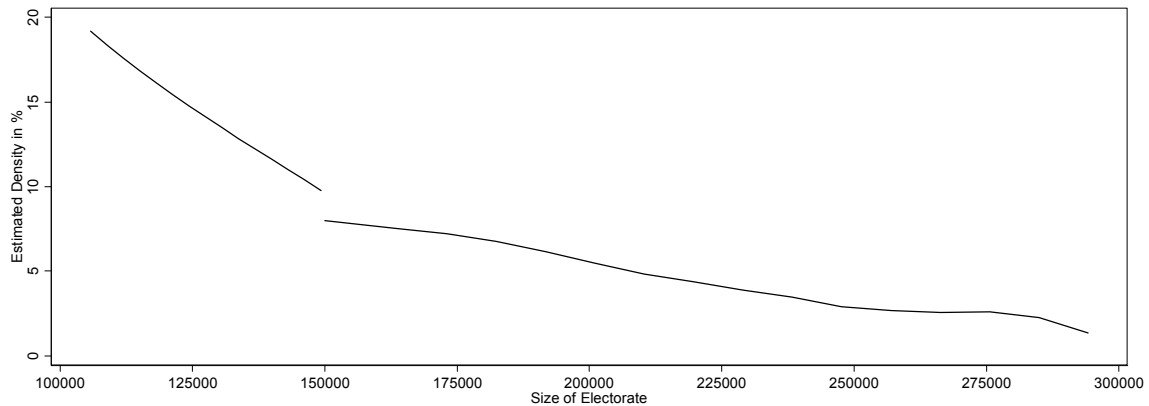
Note: Kernel Density Estimated with Epanechnikov Kernel

Figure 2. Discontinuities in the Estimated Density of Electorate
2.A Discontinuity at 200,000 Voters



Notes: Using McCrary's Local Linear Procedure with Triangular Kernel (width of 30,000 and bin of 14,000).

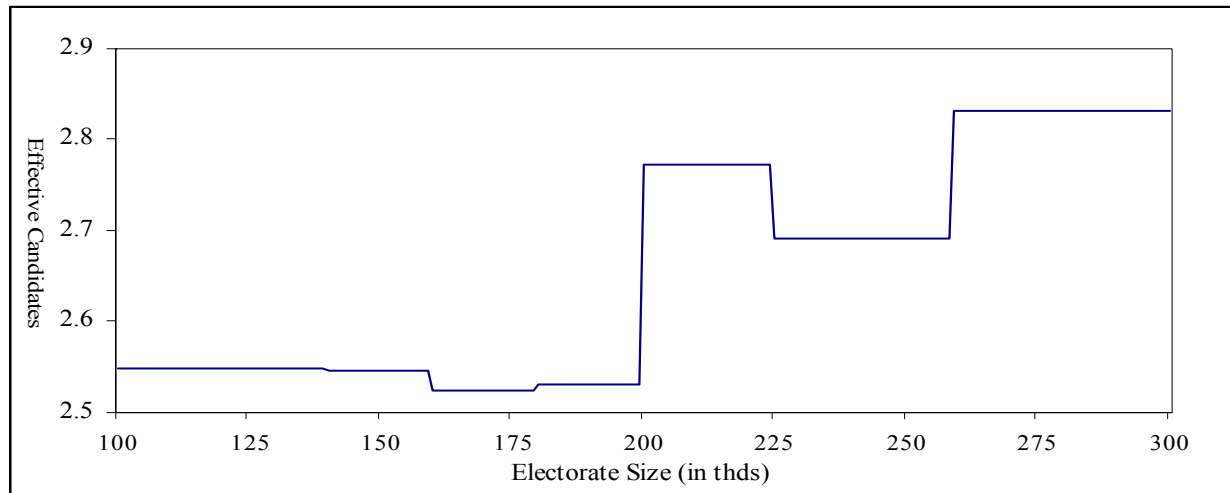
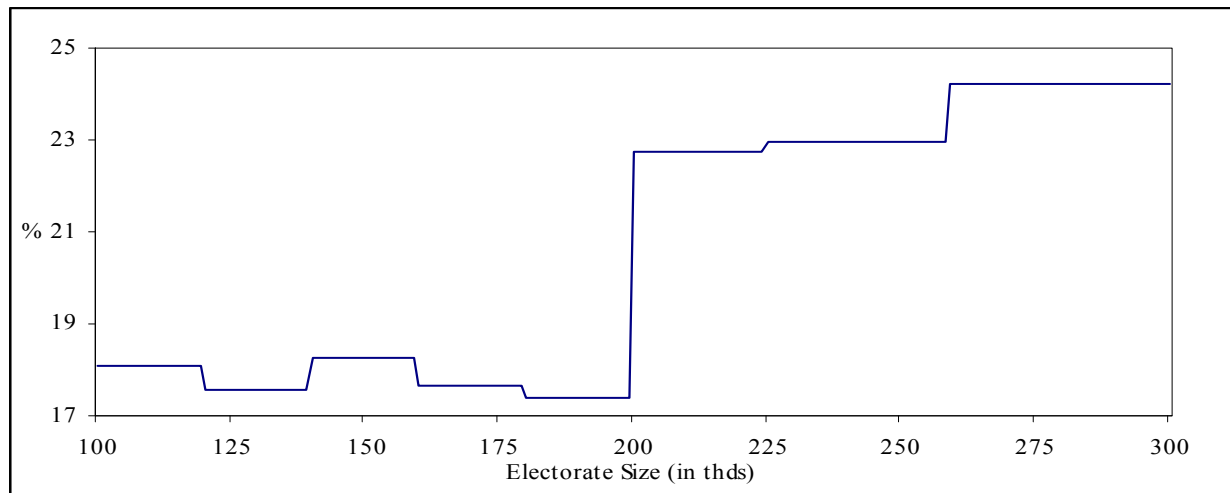
2.B. Discontinuity at 150,000 Voters.



Notes: Using McCrary's Local Linear Procedure with Triangular Kernel (width of 30,000 and bin of 14,000).

Figure 3. Vote Share Concentration By Electoral Size

3.A. Number of Effective Candidates

3.B. Total Percentage of Votes for Candidates Placed 3rd of Lower

Appendix: Algebraic equivalence between weighted 2SLS with polynomial spline and Local Polynomial Wald Estimator

We first set notation. Define Y and X as the R by 1 vectors of outcome variable and endogenous regressor, respectively. Note that R is the number of races we observe. Now define

$$Z_r = \begin{bmatrix} ELECT_r - 200thd \\ (ELECT_r - 200thd)^2 \\ \vdots \\ (ELECT_r - 200thd)^k \end{bmatrix}$$

and the matrix Z stacks the R observations of Z_r^T . G is the R by $2k$ matrix of polynomials of Z , and interactions with the dummy variable $DUM200_r$ (previously defined in the text). We also define \mathbf{D} as the R by R diagonal matrix, in which the entry $\mathbf{D}_{r,r}$ is the dummy variable $DUM200_r$ (notationally simplified here by D_r). The R by R diagonal matrix of weights is \mathbf{W} in which the entry $\mathbf{W}_{r,r}$ is the weighting variable W_r (previously defined in the text). Finally, $\mathbf{1}$ is the R by 1 vector of ones (summer vector) and $V=[\mathbf{1}, G]$.

We also define the following estimators:

$$\hat{\beta}_{Y,1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T * \arg \min_{\beta_{Y,1}, \theta_{Y,1}} \sum_{r=1}^R \left(Y_r - \beta_{Y,1} - \theta_{Y,1}^T Z_r \right)^2 * W_r * D_r .$$

Similarly we define

$$\begin{aligned} \hat{\beta}_{Y,0} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T * \arg \min_{\beta_{Y,0}, \theta_{Y,0}} \sum_{r=1}^R \left(Y_r - \beta_{Y,0} - \theta_{Y,0}^T Z_r \right)^2 * W_r * (1 - D_r) \\ \hat{\beta}_{X,1} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T * \arg \min_{\beta_{X,1}, \theta_{X,1}} \sum_{r=1}^R \left(X_r - \beta_{X,1} - \theta_{X,1}^T Z_r \right)^2 * W_r * D_r \\ \hat{\beta}_{X,0} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T * \arg \min_{\beta_{X,0}, \theta_{X,0}} \sum_{r=1}^R \left(X_r - \beta_{X,0} - \theta_{X,0}^T Z_r \right)^2 * W_r * (1 - D_r) \end{aligned}$$

And define

$$\hat{\beta}_Y = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T * \arg \min_{\beta_Y, \delta_Y} \sum_{r=1}^R (Y_r - \beta_Y D_r - \delta_Y^T V_r)^2 * W_r$$

$$\hat{\beta}_X = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T * \arg \min_{\beta_X, \delta_X} \sum_{r=1}^R (X_r - \beta_X D_r - \delta_X^T V_r)^2 * W_r$$

Proposition 1: $\hat{\beta}_Y = \hat{\beta}_{Y,1} - \hat{\beta}_{Y,0}$ and $\hat{\beta}_X = \hat{\beta}_{X,1} - \hat{\beta}_{X,0}$.

Proof of Proposition 1: We only show that $\hat{\beta}_Y = \hat{\beta}_{Y,1} - \hat{\beta}_{Y,0}$ as the second equality follows trivially by analogy. We first show that the first order conditions that define $\hat{\beta}_{Y,1}$ and $\hat{\beta}_{Y,0}$ can be written as

$$\mathbf{1}^T \mathbf{W} \mathbf{D} (Y - \mathbf{1}(\hat{\beta}_Y + \hat{\beta}_{Y,0}) - Z(\hat{\theta}_Y + \hat{\theta}_{Y,0})) = 0 \quad (\text{A.1})$$

$$Z^T \mathbf{W} \mathbf{D} (Y - \mathbf{1}(\hat{\beta}_Y + \hat{\beta}_{Y,0}) - Z(\hat{\theta}_Y + \hat{\theta}_{Y,0})) = 0 \quad (\text{A.2})$$

$$\mathbf{1}^T \mathbf{W} (\mathbf{I}_R - \mathbf{D})(Y - \mathbf{1} \hat{\beta}_{Y,0} - Z \hat{\theta}_{Y,0}) = 0 \quad (\text{A.3})$$

$$Z^T \mathbf{W} (\mathbf{I}_R - \mathbf{D})(Y - \mathbf{1} \hat{\beta}_{Y,0} - Z \hat{\theta}_{Y,0}) = 0, \quad (\text{A.4})$$

where $\hat{\theta}_Y = \hat{\theta}_{Y,1} - \hat{\theta}_{Y,0}$ and \mathbf{I}_R is the identity matrix of size R . Summing A.1 with A.3 and A.2 with A.4, we have:

$$\mathbf{1}^T \mathbf{W} (Y - \mathbf{1} \hat{\beta}_{Y,0} - Z \hat{\theta}_{Y,0} - \mathbf{D} \mathbf{1} \hat{\beta}_Y - \mathbf{D} Z \hat{\theta}_Y) = 0 \quad (\text{A.5})$$

$$Z^T \mathbf{W} (Y - \mathbf{1} \hat{\beta}_{Y,0} - Z \hat{\theta}_{Y,0} - \mathbf{D} \mathbf{1} \hat{\beta}_Y - \mathbf{D} Z \hat{\theta}_Y) = 0. \quad (\text{A.6})$$

And letting $\hat{\delta}_Y = [\hat{\beta}_{Y,0}, \hat{\theta}_{Y,0}^T, \hat{\theta}_Y^T]^T$ we note that A.1, A.2, A.5 and A.6 can be written respectively as

$$\mathbf{1}^T \mathbf{D} \mathbf{W} (Y - V \hat{\delta}_Y - \mathbf{D} \mathbf{1} \hat{\beta}_Y) = 0 \quad (\text{A.1}')$$

$$Z^T \mathbf{D} \mathbf{W} (Y - V \hat{\delta}_Y - \mathbf{D} \mathbf{1} \hat{\beta}_Y) = 0 \quad (\text{A.2}')$$

$$\mathbf{1}^T \mathbf{W} (Y - V \hat{\delta}_Y - \mathbf{D} \mathbf{1} \hat{\beta}_Y) = 0 \quad (\text{A.5}')$$

$$Z^T \mathbf{W} (Y - V \hat{\delta}_Y - \mathbf{D} \mathbf{1} \hat{\beta}_Y) = 0 \quad (\text{A.6}')$$

which are the first order conditions for $\min_{\beta_Y, \delta_Y} \sum_{r=1}^R (Y_r - \beta_Y D_r - \delta_Y^T V_r)^2 * W_r$ since $V=[\mathbf{1}, Z, \mathbf{D}Z]$.

Thus, $\hat{\beta}_Y = \hat{\beta}_{Y,1} - \hat{\beta}_{Y,0}$ and $\hat{\delta}_Y = \hat{\delta}_{Y,1} - \hat{\delta}_{Y,0}$ equal $\arg \min_{\beta_Y, \delta_Y} \sum_{r=1}^R (Y_r - \beta_Y D_r - \delta_Y^T V_r)^2 * W_r$. *Q.E.D.*

We now define $\hat{\beta}_1$, which is the weighted 2SLS estimator of Y on X controlling for a polynomial spline of electorate. This corresponds to a two-stage least square estimator that uses \mathbf{W} as a matrix of weights and a polynomial function given by the matrix G as controls.⁴² We can formally define $\hat{\beta}_1$ as:

$$\hat{\beta}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T * \arg \min_{\beta_Y, \delta_Y} (Y^* - X^* \beta - V^* \delta)^T P_{V^*, D^*} (Y^* - X^* \beta - V^* \delta)$$

where $Y^* = \mathbf{W}^{1/2} Y$, $X^* = \mathbf{W}^{1/2} X$, $V^* = \mathbf{W}^{1/2} V$, $P_{V^*, D^*} = P_{V^*} + M_{V^*} \mathbf{D} \mathbf{1}^* (\mathbf{1}^{*T} \mathbf{D} M_{V^*} \mathbf{D} \mathbf{1}^*)^{-1} \mathbf{1}^{*T} \mathbf{D} M_{V^*}$, $\mathbf{1}^* = \mathbf{W}^{1/2} \mathbf{1}$, $P_{V^*} = V^* (V^{*T} V^*)^{-1} V^{*T}$, $M_{V^*} = \mathbf{I}_R - P_{V^*}$ and where $\mathbf{W}^{1/2}$ is a square-root matrix of $\mathbf{W}^{1/2}$, that is, $\mathbf{W}^{1/2} \mathbf{W}^{1/2} = \mathbf{W}$ the diagonal matrix of weights.

A solution to the above minimization problem is the well known formula for the 2SLS estimator using, of course, the Frisch-Waugh formula:

$$\hat{\beta}_1 = (X^{*T} P_{V^*, D^*} M_{V^*} P_{V^*, D^*} X^*)^{-1} X^{*T} P_{V^*, D^*} M_{V^*} P_{V^*, D^*} Y^*.$$

Proposition 2: $\frac{\hat{\beta}_Y}{\hat{\beta}_X} = \hat{\beta}_1$.

Proof of Proposition 2: We first note that

$$P_{V^*, D^*} M_{V^*} P_{V^*, D^*} = M_{V^*} \mathbf{D} \mathbf{1}^* (\mathbf{1}^{*T} \mathbf{D} M_{V^*} \mathbf{D} \mathbf{1}^*)^{-1} \mathbf{1}^{*T} \mathbf{D} M_{V^*}$$

and that X and $\mathbf{1}$ are R by 1 matrices. Thus, we have

$$\begin{aligned} \hat{\beta}_1 &= (X^{*T} M_{V^*} \mathbf{D} \mathbf{1}^* (\mathbf{1}^{*T} \mathbf{D} M_{V^*} \mathbf{D} \mathbf{1}^*)^{-1} \mathbf{1}^{*T} \mathbf{D} M_{V^*} X^*)^{-1} X^{*T} M_{V^*} \mathbf{D} \mathbf{1}^* (\mathbf{1}^{*T} \mathbf{D} M_{V^*} \mathbf{D} \mathbf{1}^*)^{-1} \mathbf{1}^{*T} \mathbf{D} M_{V^*} Y^* \\ &= (\mathbf{1}^{*T} \mathbf{D} M_{V^*} X^*)^{-1} \mathbf{1}^{*T} \mathbf{D} M_{V^*} Y^*. \end{aligned}$$

Now, we rewrite equations (A.1'), (A.2'), (A.5') and (A.6') and solve them for $\hat{\beta}_Y$:

$$0 = \mathbf{1}^T \mathbf{D} \mathbf{W} (Y - V \hat{\delta}_Y - \mathbf{D} \mathbf{1} \hat{\beta}_1) \Rightarrow 0 = \mathbf{1}^{*T} \mathbf{D} (Y^* - V^* \hat{\delta}_Y - \mathbf{D} \mathbf{1}^* \hat{\beta}_1)$$

⁴² In the text we use other controls beyond the polynomial of electorate. Equivalence results, however, do not change.

$$0 = V^T \mathbf{W}(Y - V \hat{\delta}_Y - \mathbf{D} \mathbf{1} \hat{\beta}_Y) \Rightarrow 0 = V^{*T} \mathbf{D}(Y^* - V^* \hat{\delta}_Y - \mathbf{D} \mathbf{1}^* \hat{\beta}_Y),$$

and therefore

$$\hat{\beta}_Y = (\mathbf{1}^{*T} \mathbf{D} M_{V^*} \mathbf{1}^*)^{-1} \mathbf{1}^{*T} \mathbf{D} M_{V^*} Y^*.$$

By analogy

$$\hat{\beta}_X = (\mathbf{1}^{*T} \mathbf{D} M_{V^*} \mathbf{1}^*)^{-1} \mathbf{1}^{*T} \mathbf{D} M_{V^*} X^*$$

And therefore $\frac{\hat{\beta}_Y}{\hat{\beta}_X} = \hat{\beta}_1 \cdot Q.E.D.$

Our result shows the algebraic equivalence between a Local Polynomial Wald estimator of the type $(\hat{\beta}_{Y,1} - \hat{\beta}_{Y,0})/(\hat{\beta}_{X,1} - \hat{\beta}_{X,0})$ that uses the same kernel function and bandwidth for all four estimators and a weighted 2SLS that uses kernel as weights, that is,

$$W_r = \frac{1}{h} K\left(\frac{ELECTORATE_r - 200thd}{h}\right)$$

and controls as those given by C_r . Therefore, a weighted 2SLS (that uses weights that shrink to the discontinuity point as R goes to infinity) will be consistent even if the instrument is not valid outside the discontinuity point, as long as the continuity assumptions (described in Hahn, Todd and van der Klaauw, 2001) on conditional expectations of potential outcomes and regressors hold.

The algebraic equivalence result presented above generalizes the results described in Hahn, Todd and van der Klaauw (2001) and in Imbens and Lemieux (2008). Hahn, Todd and van der Klaauw (2001) show the algebraic equivalence between a Local Constant Wald Estimator that uses rectangular kernels (which is a Wald estimator based on Nadaraya-Watson estimators to the left and to the right of the discontinuity) and a 2SLS that discards data that are further than h away from the discontinuity point. Imbens and Lemieux (2008) generalize that result for a Local Linear Wald, showing that such estimator could also be written as a 2SLS under the specific case of rectangular kernel. In the regression discontinuity setting, local linear regressions are preferable than local constant because of the bias inherent in using Nadaraya-Watson on the boundary point.

Our generalization occurs in two fronts: first it allows us to use any kernel function and not only the rectangular one; second it allows us to use polynomial functions of the

continuous variable (electorate in our case) as a way to reduce asymptotic bias. Finally, given the algebraic equivalence between our procedure and the one proposed by Hahn, Todd and van der Klaauw (2001), under the same regularity conditions and for the case that X is binary, our estimator will have the same asymptotic distribution as the one derived by them.

In general, asymptotic distribution of the weighted IV will depend on how weights are specified. If weights are such that they give higher importance to observations closer to the discontinuity point as the sample size increases, then the estimator will be essentially non-parametric. Consistency will depend on the continuity assumptions described by Hahn, Todd and van der Klaauw (2001).

However, if one uses a fixed bandwidth, then although the estimator will converge at the parametric rate, consistency will be achieved only if usual IV assumptions are satisfied. Also, if bandwidth is fixed, asymptotic distribution of our method is straightforward to derive and usual standard errors reported by statistical packages will be valid.