Universidad de San Andrés

Departamento de Economía

Maestría en Economía

Electoral Rules and Political Competition

Regression Discontinuity Evidence from Brazilian Municipalities

Lautaro Carrizo

40208573

Mentor: Christian Ruzzier

Vito Dumas 284

13 de Noviembre, 2020

Tesis de Maestría en Economía de

Lautaro Carrizo

"Reglas Electorales y Competencia Política"

Resumen

Este paper usa la variabilidad exógena en esquemas de votación para estimar el efecto causal de diferentes reglas electorales sobre la competencia política, utilizando la asignación de ballotage simple y ballotage doble en las elecciones a alcalde en Brasil. Esta asignación lo hace un escenario ideal para utilizar el esquema cuasi-experimental de regresión discontinua. Los resultados indican que la regla electoral de doble ballotage produce un incremento de la competencia política, disminuyendo no solo la probabilidad de que el alcalde sea reelecto, sino también el porcentaje de votos del candidato ganador. En línea con resultados teóricos, encontramos evidencia consistente con un incremento en el número de partidos y candidatos como un posible mecanismo detrás estos efectos.

Palabras clave: reglas electorales, competencia política, ballotage, regresión discontinua.

"Electoral Rules and Political Competition"

Abstract

This paper uses exogenous variation in voting schemes to estimate the causal effect of different electoral rules on political competition, exploiting a regression discontinuity design in the assignment of single-ballot and dual-ballot plurality systems in Brazilian mayoral races. The results indicate that dual-ballot plurality rule causes an increase in political competition, decreasing the probability of reelection of the incumbent mayor (the personal incumbency advantage) and the vote share of winning candidate. In line with theoretical results, we find that evidence consistent with an increase in the number of parties and candidates as a possible mechanism behind these findings.

Keywords: electoral rules, political competition, incumbency advantage, vote share, regression discontinuity design.

JEL Classification: H73, R58, I28

1 Introduction

For many years economists, sociologists and political scientists have developed theories and conducted field researches of how the citizens vote. One of the most important results in public choice economics is the *median voter theorem* (Black (1948)) which establishes that, under certain assumptions, a majority-rule voting system will select the outcome most preferred by the median voter. However, the conclusion could be very different under a different voting scheme; hence the importance, both in theory and in practice, of understanding the consequences of different electoral rules on the behavior and choices of voters and candidates. The first group could vote sincere or strategically depending on the context, whereas the second group could choose to merge into a party or not.

The most known prediction regarding strategic voting is *Duverger's Law*, which states that simple-majority single-ballot (SB) favors a two-party system and strategic voting whereas simple-majority dual-ballot (DB) favors multipartyism and sincere voting. Rather than focusing on the behavior of the voters Bordignon et al. (2017) have recently centered on the behavior of the candidates, emphasizing that in a model where the number of parties is endogenous, the SB system attracts fewer parties than the DB system, because SB pushes candidates to merge into parties in order to gain more votes.

Barone and De Blasio (2013) exploit the institutional setup in Italy (where SB applies to municipalities with less than 15,000 inhabitants, while DB is in place above that threshold), to study the consequences of a change from SB to DB, finding that the number of parties is larger under DB than under SB–lending empirical support to *Duverger's Law* and the model in Bordignon et al. (2017). Bordignon et al. (2016) exploit the same quasi-experimental design in Italy and conclude that DB attracts more candidates, but also reduces the influence of extremist parties, compared to SB. In the same line, Fujiwara (2011) show that a change from SB to DB produced an increase in political competition in Brazilian municipal elections, and document a displacement of the third-placed candidate under SB, which does not occur under DB.[1]

In this paper, we estimate the causal effect of a change from a single-ballot majority system to a dual-ballot majority system on two measures of political competition, the probability of reelection of the incumbent mayor (the personal incumbency advantage) and the vote share of winning candidate, through a regression discontinuity design (RDD), finding that the DB increases political competition along both dimensions. One possible explanation of these results is that SB encourage the voters to make a strategic vote, voting more for candidates with high probability of being elected increasing the strength of the incumbent (who usually is one of the top candidates) or the top candidate. Another possible explanation is that the DB increase the number of parties and candidates giving the citizens more options and reducing votes of the incumbent or the top candidate, these two effects are further explained in section 7.

The institutional setup that characterizes Brazilian municipalities is ideal to explore the

[1] Chamon et al. (2019) use this change from SB to DB to instrument political competition in their exploration of the consequences of different electoral rules on fiscal spending.

effect of a change from SB to DB, since municipalities with less than 200,000 registered voters face a SB while a DB is in place above that threshold. Under SB there is only one round with the winner being the one who obtains the most votes, while DB is a two-round system: first an election is held and if a candidate obtains more than 50% of the votes, she is elected; otherwise, a second round of voting ensues where only the two most voted candidates in the first round face off. We exploit the change in electoral rules at the 200,000 voters cutoff to identify a causal effect through a sharp RDD.

To our knowledge, we are the first to exploit the causal effect of electoral rules on the personal incumbency advantage. Klašnja and Titiunik (2013) and Brambor and Ceneviva (2011) have also documented an incumbency disadvantage through an RDD, but exploiting close electoral races under a given electoral rule. The first research studies the party incumbency advantage whereas the second studies the personal incumbency advantage.

2 Econometric strategy

2.1 Empirical Strategy

Brazil is constituted by 26 states, a federal district and has more than 5000 municipalities, the smallest level of government in the country. Each municipality is run by a *Prefeito* who is elected every 4 years. Municipal elections are regulated by federal legislation, and all municipalities have the first round of the election at the beginning of October and the second round at the end of October.

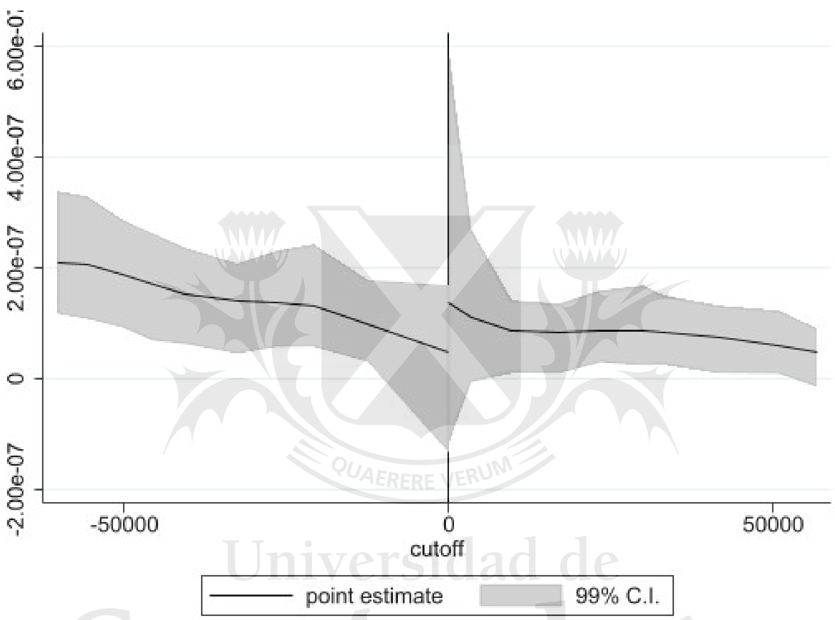
Brazilian legislation requires all adult citizens to register to vote in their municipality of residence, voter registration is compulsory. Moreover, since the reform of 1998 the Constitution states that mayoral elections should be run under SB in municipalities with less than 200,000 voters, while municipalities with 200,000 voters or more must have their elections under DB.

The Brazilian scheme of voting creates an environment suitable for applying a standard regression discontinuity design. Under mild assumptions, municipalities that receive very similar score values on the opposite sides of the cutoff are comparable to each other in all relevant aspects, except for the treatment status. In other words, the reason that they are on a particular side of the threshold can be interpreted as random assigned that should not be related to the outcome of interest. This argument is formalized by Lee (2008).

Other than the voting rule, any observed or unobserved variable that could affect voting should be the same for all municipalities that are sufficiently close to the threshold. Under the mild assumptions of the RDD, any difference in outcomes between these two groups is a causal consequence of the different electoral rules. Since every municipality with more than 200,000 voters face a DB rule, whereas every municipality with less than 200,000 voters face a SB rule the RDD is sharp.

The first threat to the validity of the quasi-experimental design that could invalidate the analysis is the possibility of strategic manipulation of the forcing variable because it would produce self-selection. It is possible that some incumbent has preference for SB over DB (or vice versa) and could try to manipulate the registration of voters to have the desired political rule. This kind of behavior would produce a discontinuity in registration rate or in the number of cities that are above or below the cutoff. Fortunately, this issue can be tested and rejected in the data. For our particular analysis we have to make two tests of continuity of the number of registered voters around the cutoff; one keeping only municipalities that have a candidate that ran for reelection and another one keeping just one observation per election (municipality-election level).[2] These two tests are presented in Figure 1 and Figure 2 respectively.[3]

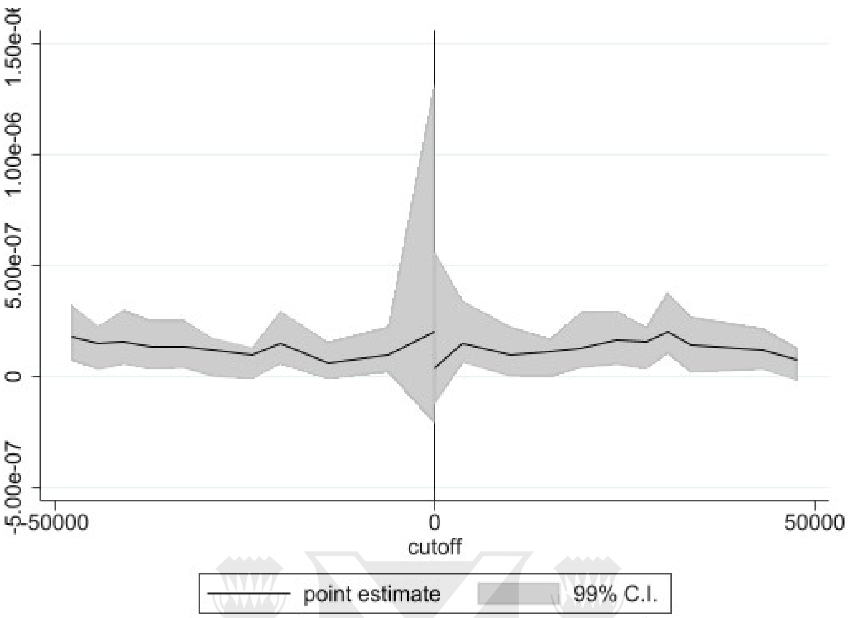
Figure 1: Estimated density of the score-only candidates suitable for reelection



[2]To perform the test we used the *rddendensity* command developed in Cattaneo et al. (2016).

[3]In the first test the selected bandwidth was 20,000 in order to have enough observations to perform the test.

Figure 2: Estimated density of voters in Brazil elections from 2000 to 2016



In the first test the value of the statistic is 1.5706 and the associated p-value is 0.1163. This means that under the continuity-based approach, we fail to reject the null hypothesis of no difference in the density of treated and control observations at the cutoff. Meanwhile in the second test the value of the statistic is -0.8624 and the associated p-value is 0.3885.

Another threat to the validity of our design would occur if a change from SB to DB affected voter turnout because it would not pass the test of predetermined covariates.[4] However, Brazilian law makes registration and voting compulsory for all citizens aged 18-70. Failing to register or vote in a previous election renders a citizen ineligible for several publicly provided services until a fine is paid. Moreover, elections are held on Sunday and voters are allocated to polls close to their residence in order to have a better turnout. This feature makes the difference in turnout under SB and DB virtually zero. The evidence reported in Fujiwara (2011) and Chamon et al. (2019) further shows that turnout is indeed unaffected by the electoral rules in Brazil.

2.2 Estimation Framework

Let ν be the number of registered voters in a municipality. The treatment effect of a change from SB to DB on outcome y is given by

$$ATE = \lim_{v \uparrow 200{,}000} \mathrm{E}[y \mid v] - \lim_{v \downarrow 200{,}000} \mathrm{E}[y \mid v]$$

Under the standard assumption of continuity of the conditional expectation of y on ν, the first term converges to the expected outcome of a municipality with 200,000 voters and SB,

[4]Further explained in section 5.

while the second term to the expected outcome of a municipality of 200,000 voters and DB. As long as the distribution of treatment effects is continuous at the threshold, ATE identifies the treatment effect of changing from SB to DB for a municipality of 200,000 voters.[5]

The estimation method used here follows the guidelines in Cattaneo et al. (2019). The limits on the right hand side are estimated non-parametrically by local linear regressions. This approach uses only observations that are between $200,000 - bw$ and $200,000 + bw$, where bw is a so-called bandwidth that determines the size of the neighborhood around the cutoff. Within the bandwidth, it is common to adopt a weighting scheme to ensure that the observaions closer to 200,000 receive more weight than those further away; the weights are determined by a kenerl function.[6]

A key decision is the kernel bandwidth bw, which controls the width of the neighborhood around the cutoff. This parameter directly affects the properties of local polynomial estimation and inference procedures. Choosing a smaller bw will reduce the misspecification error (also known as smoothing bias) of the local polynomial approximation, but will simultaneously tend to increase the variance of the estimated coefficients because fewer observations will be available for estimation. On the other hand, a larger bw will result in more smoothing bias if the unknown function differs considerably from the polynomial model used for approximation, but will reduce the variance because the number of observations in the interval $[200,000bw; 200,000 + bw]$ will be larger. For this reason, the choice of bandwidth is said to involve a bias-variance trade-off. To select the most appropriate bw we use the command *rdbwselect* (developed in Calonico et al. (2017)), which seeks to minimize the MSE of the local polynomial RD point estimator, given a choice of polynomial order and kernel function. In this particular natural experiment the optimal bw is 47,781 if we are taking as outcome the probability of being reelected, and 35,881 when running the regression using the vote share of the most voted candidate as dependent variable.

3 Data

3.1 Dataset, sources and variables

The principal dataset is at the candidate-election level, containing data from the mayoral elections of 2000, 2004, 2008, 2012 and 2016 available from the Tribunal Superior Eleitoral.[7] For the whole analysis we have excluded elections in which there is only one candidate, and the first round of elections where there was second round. The main variables are the number of votes per candidate-election, the number of voters per municipality-election and a categorical variable that takes value of 1 if the candidate won the election, 4 if the candidate lost the election, 6 if the candidate passed to the second round[8] and 10 if the candidate resigned or died. From these three variables we create the cutoff and the outcomes:

[5]Continuity offers one justification for using observations just below the cutoff to approximate the average outcome that units just above the cutoff would have had if they had received the control condition instead of the treatment

[6]In the main specifications we use triangular kernels to make the observations near the cutoff weigh more

[7]`http://www.tse.gov.br`

[8]Only for municipalities with more than 200,000 voters

- We compute the running variable as the number of voters minus 200,000, so as to have the cutoff equal to zero.
- In constructing the variable *Reelection* we consider only the candidates that are running for reelection. The variable takes a value of 1 if the candidate is reelected, and zero otherwise.[9]
- The variable *Vote Share* is the number of votes per winning candidate over the total number of votes in a given election and municipality.

Note that the first outcome is created to evaluate the effect of the political rules over the personal incumbency advantage (or disadvantage), whereas the second outcome is created to evaluate the effect over political competition.

From Brazil's 2000 Demographic Census we obtain GDP per capita, the share of economically active population and the share of urban population. From *Pesquisa de Informacoes Basicas Muncipais* we obtain the share of population without sewerage and the share of population with electricity.[10]

In analyzing the incumbency advantage, a decision must be made regarding the level at which this advantage is going to be analyzed. Fujiwara (2011) and Bordignon et al. (2016), for instance, consider the incumbent mayor's advantage (the personal incumbency advantage), whereas Titiunik (2009), Lee (2008) and Klašnja and Titiunik (2013) analyze the incumbent party's advantage. Since Brazil is a country in which voters do not show a strong loyalty to parties, candidates frequently change their political party in order to gain more votes (Brambor and Ceneviva (2011)), we favor the study of the personal incumbency advantage in this paper. Since the Brazilian electoral system produces incentives for individual politicians to seek a personal vote, we also construct our dataset at the individual candidate level to have a better measure of political power.

3.2 Descriptive statistics

This section contains the descriptive statistics for the subsample with 155,000-245,000 voters. We present information for municipalities with electorates above 200,000 (DB) and for municipalities under 200,000 (SB). Table 1 show the summary statistics of our measures of political competition. Table 2 has descriptive statistics of demographic information[11] For both Tables there is also show the t-stat of a difference in means for each reported variable.

[9]It is important to notice that we are using the probability of being reelected conditional on the fact that the incumbent ran for reelection. It would be interesting to test the uncondicional probability of being reelected in further research.

[10]Both sources are available from the Instituto Brasileiro de Geografia e Estaistica (IBGE): `http://www.ibge.gov.br/english`.

[11]In section 5 we use this same variables to run validity tests.

Table 1: Descriptive statistics of the outcomes.

	(1)	(2)	(3)
Variable	Less than 200,000 voters	More than 200,000 voters	(1)-(2)
Pr of reelection	0.600	0.730	0.130
	(0.495)	(0.450)	(0.103)
Vote share of the winning cand	0.189	0.172	-0.017
	(0.199)	(0.194)	(0.012)

* p<0.1; ** p<0.05; *** p<0.01

Notes. Means of the two outcomes. Column (1) has races with 155,000-200,000 voters and column (2) has races with 200,000-245,000 voters. For column (1)-(2) standard deviation are in parenthesis. For column (3) standard errors from a t-test of equity of means are reported in parenthesis.

Table 2: Descriptive statistics of the demographic characteristics.

	(1)	(2)	(3)
Variable	Less than 200,000 voters	More than 200,000 voters	(1)-(2)
Share of electricity	98.812	99.059	0.247
	(3.054)	(2.229)	(0.162)
Share or no-sewerage	2.233	1.976	-0.257
	(3.202)	(2.758)	(0.180)
Share of urban population	95.207	95.789	0.583*
	(6.198)	(4.904)	(0.337)
Per capita GDP	8,409.885	8,220.404	-189.480
	(6,486.250)	(5,538.588)	(365.196)
Share of economically active pop	58.755	58.664	-0.091
	(3.745)	(3.102)	(0.207)

* p<0.1; ** p<0.05; *** p<0.01

Notes. Means of the two outcomes. Column (1) has races with 155,000-200,000 voters and column (2) has races with 200,000-245,000 voters. For column (1)-(2) standard deviation are in parenthesis. For column (3) standard errors from a t-test of equity of means are reported in parenthesis.

4 Main Results

We present here two main results on the effect of a change from SB to DB: a decrease of the personal incumbency advantage and a reduction in the vote share of the winning candidate.

The first column of Table 3 presents the change in the incumbent mayor's probability of being reelected due to facing a DB electoral race rather than a SB electoral race. This probability is reduced by more than 30 percentage points under DB as compared to SB.[12] In other words, DB decreases the probability of reelection by more than 30%.

[12] It is important to notice that the mean of the variable *Reelection* in municipalities just below the cutoff is nearly 1 and just below the cutoff it decreases to 0.6 if we are looking at a window of 10,000 voters

In the other two columns we report very similar results from the same specification but with different bandwidths.[13] In the second column the effect is not significant if we use robust standard errors, because with a narrower bandwidth the number of observations drop significatively. The effect is slightly smaller if the bandwidth is bigger, as expected. In the next section we show that using a higher polynomial degree does not affect our results.

Table 3: Treatment effect on the personal incumbency advantage.

	Optimal BW	Lower BW	Bigger BW
RD_Estimate	-0.30192*	-0.40841	-0.21013**
N	9250	9250	9250
Effective observations	93	47	109
Left	55	25	67
Right	38	22	42
Bandwidth	47781	25000	55000

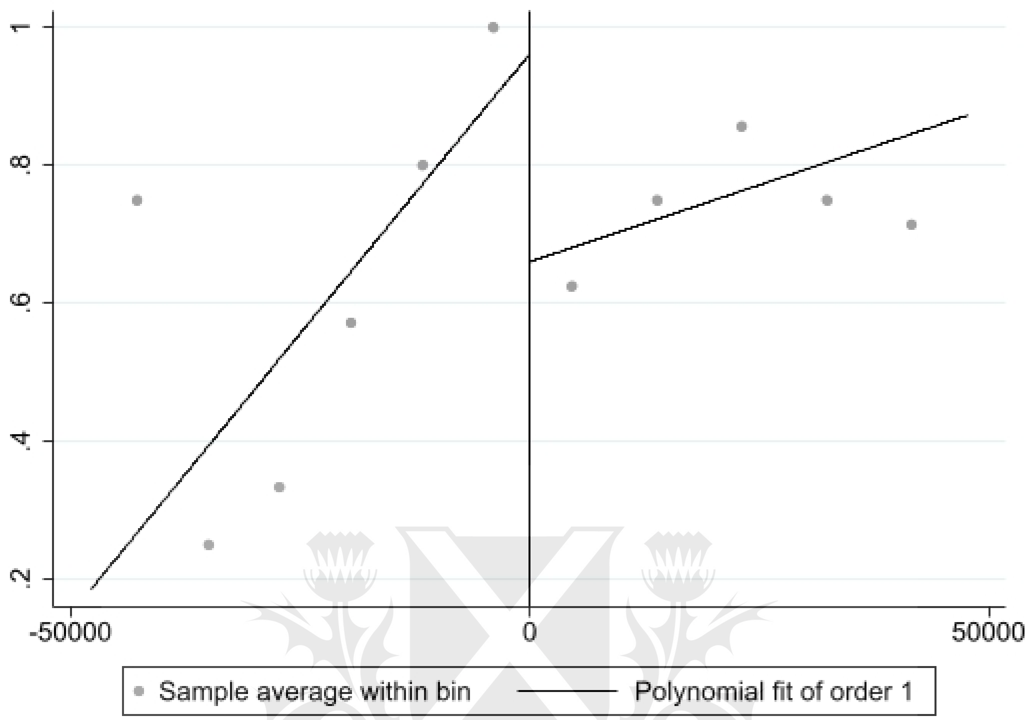
* p<0.1; ** p<0.05; *** p<0.01

Notes. Election year between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at candidate-election level. Dependent variable: Probability of being reelected. The statistical significance is determined by the robust standard errors.

To facilitate visualization, Figure 3 presents the probability of being reelected, against the forcing variable (number of registered voters minus 200,000), using the optimal bandwidth (Table 3 column1).

[13] We ran the regressions in Stata using the rdrobust command developed in Calonico et al. (2017). All specifications use a triangular kernel and linear approximations

Figure 3: Probability of reelection



Notice that the probability of being reelected is quite heterogeneous inside the window we are looking at: small on the left-hand side of the window and large to the right of the cutoff.[14] To verify that the results are not driven by our use of triangular kernels (which weigh observations near the cutoff more heavily), we re-ran the regressions with Epanechnikov and uniform kernels, and the result still holds.

Our next result concerns the effect of a change in the electoral rule on political competition, as measured by the vote share of the winning candidate. The first column of Table 4 shows that changing from SB to DB decreases this vote share by 9 percentage points (using the optimal bandwidths). The results are robust to considering different bandwidths, as shown in the remaining columns.

[14]This particular result is intriguing. One possible explanation is that the cost of changing policies is smaller in smaller districts than in larger cities, making it easier for the actual mayor to retain the votes in larger cities. Another explanation could be that voters in smaller districts are better informed of the behavior of the mayor which could make them more critical of her politics.

Table 4: Treatment effect on the vote share of the winning candidate.

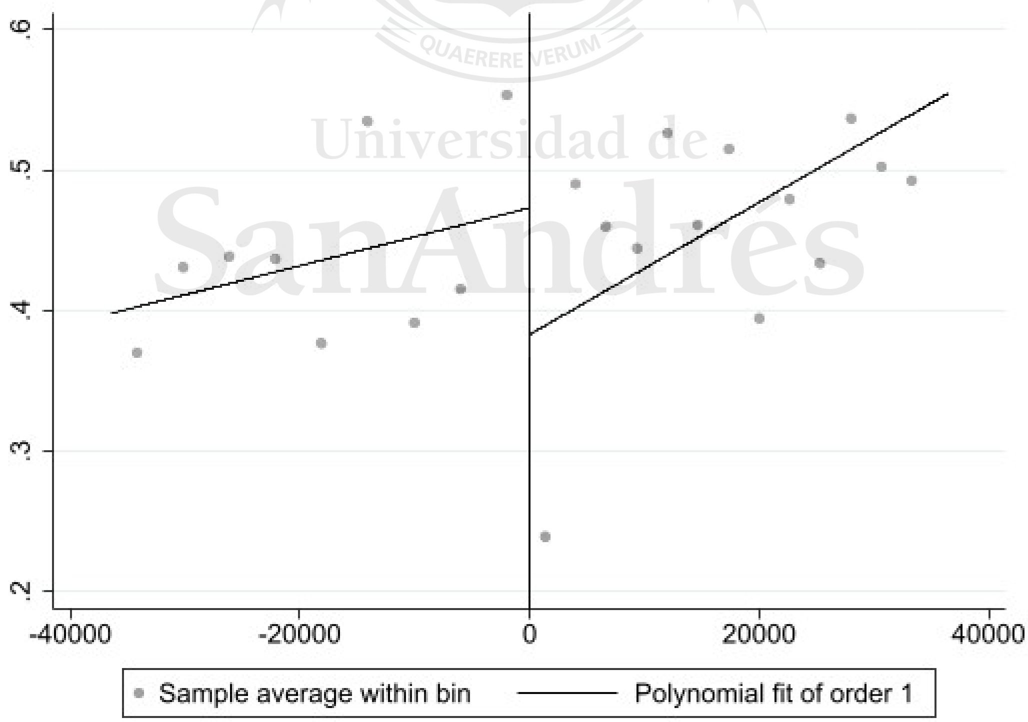
	Optimal BW	Lower BW	Bigger BW
RD_Estimate	-0.09024*	-0.13925***	-0. 07578**
N	27434	27434	27434
Effective observations	187	118	234
Left	102	61	137
Right	85	57	98
Bandwidth	36429	25000	45000

* p<0.1; ** p<0.05; *** p<0.01

Notes. Election year between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at candidate-election level. Dependent variable: Vote Share of the winning candidate. The statistical significance is determined by the robust standard errors.

Figure 4 plots the vote share against the forcing variable (number of registered voters minus 200,000), using the optimal bandwidth (Table 4, column1).

Figure 4: Vote Share of winning candidate



5 Validity of the Methodology

In this section we provide a set of empirical tests to show that our RD design is valid.

5.1 Treatment effect at artificial cutoff values

The key identifying assumption is the continuity (or lack of abrupt changes) of the regression functions for treatment and control units at the cutoff in the absence of treatment. Evidence of continuity away from the cutoff can be interpreted as potentially casting doubt on the RD design. To test this continuity in other regions, we run a regression replacing the true cutoff value by another value at which the treatment status does not really change, and conduct the estimation using this artificial cutoff point. A priori, the expectation is that no significant treatment effect will occur at placebo cutoff values.

In the next tables we present a series of regressions using different cutoffs: the first two tables display the effect of the cutoff[15] $+4000, -4000, +8000, -8000, +20000$ and -20000 on the probability of being reelected, whereas the last two tables examine the effect on the vote share of the winning candidate at each alternative cutoff. As expected, no statistically significant effect shows up in any of the regressions.

[15]Notice that the cutoff is the number of voters minus 200,000

Table 5: Multiple placebo cutoffs over the personal incumbency advantage

	Cutoff+4000	Cutoff-4000	Cutoff+8000
RD_Estimate	0.01171	-0.00883	0.27557
N	9250	9250	9250
Effective observations	140	159	153
Left	93	112	102
Right	47	47	51
Bandwidth	64874	79762	68693
	* p<0.1; ** p<0.05; *** p<0.01		

	Cutoff-8000	Cutoff+20000	Cutoff-20000
RD_Estimate	0.10559	0.24529	0.05692
N	9250	9250	9250
Effective observations	202	346	74
Left	148	277	44
Right	54	69	33
Bandwidth	102407	106819	47523
	* p<0.1; ** p<0.05; *** p<0.01		

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at candidate-election level. Dependent variable: Probability of being reelected. The statistical significance is determined by the robust standard errors.

Table 6: Multiple placebo cutoffs over the vote share of the winning candidate

	Cutoff+4000	Cutoff-4000	Cutoff-8000
RD_Estimate	0.00922	0.0081	0.00215
N	27434	27434	27434
Effective observations	360	140	414
Left	236	77	277
Right	124	63	137
Bandwidth	61467	34940	67327
	* p<0.1; ** p<0.05; *** p<0.01		

	Cutoff-8000	Cutoff+20000	Cutoff-20000
RD_Estimate	0.02867	-0.00816	0.0081
N	27434	27434	27434
Effective observations	316	446	140
Left	211	309	77
Right	105	137	63
Bandwidth	63083	63797	34940
	* p<0.1; ** p<0.05; *** p<0.01		

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at candidate-election level. Dependent variable: Vote Share of the winning candidate. The statistical significance is determined by the robust standard errors.

5.2 Predetermined Covariates

This falsification test is based on the effect of the treatment on predetermined covariates. All predetermined covariates should be analyzed in the same way as the outcome of interest, and to validate the methodology we must find that predetermined covariates are not affected by the treatment.

In the following graphs we present the effect of a change from SB to DB on several predetermined covariates (all variables are from the 2000s). None of the effects are statistically significant at the usual levels of confidence, lending further support to our experimental design. The graphs (a) contain only the municipalities where there is a mayor running for reelection, whereas the graphs (b) contain one municipality per election.

Figure 5: Share of electricity

(a) (b)

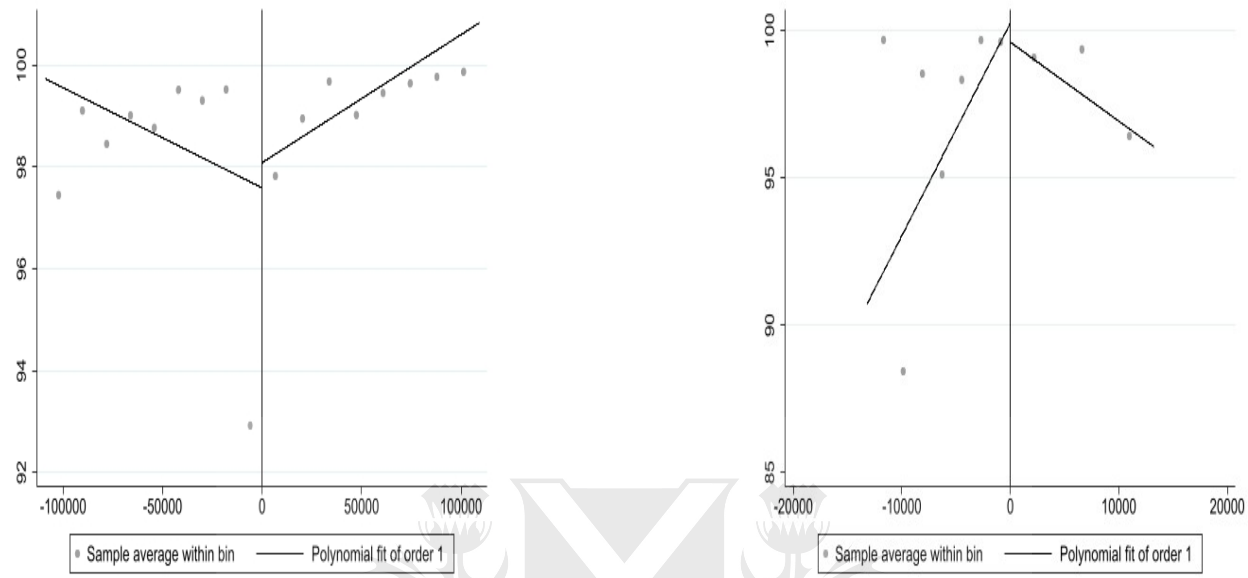


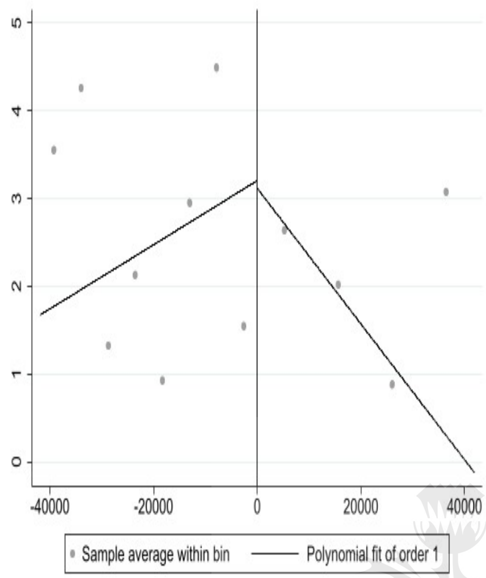
Table 7: Treatment effect on share of population with electricity

	Reelection Sample	Whole Sample
RD_Estimate	0.49407	-0.65014
N	9158	27195
Effective observations	252	60
Left	190	26
Right	62	34
Bandwidth	107488	13208
	* p<0.1; ** p<0.05; *** p<0.01	

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at municipality-election level. Dependent variable: Share of population with electricity. The statistical significance is determined by the robust standard errors.

Figure 6: Share of population without sewerage

(a)



(b)

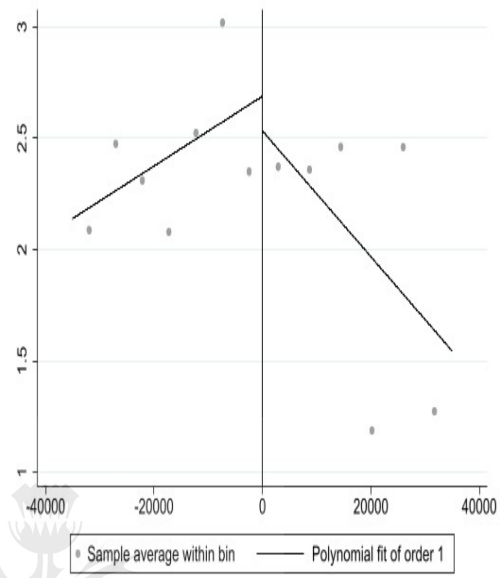


Table 8: Treatment effect on share of population without sewerage

	Reelection Sample	Whole Sample
RD_Estimate	-0.08593	-0.1558
N	9124	27089
Effective observations	80	180
Left	45	95
Right	35	85
Bandwidth	41880	34999
	* p<0.1; ** p<0.05; *** p<0.01	

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at municipality-election level. Dependent variable: Share of population without sewerage. The statistical significance is determined by the robust standard errors.

Figure 7: Per cap GDP

(a) (b)

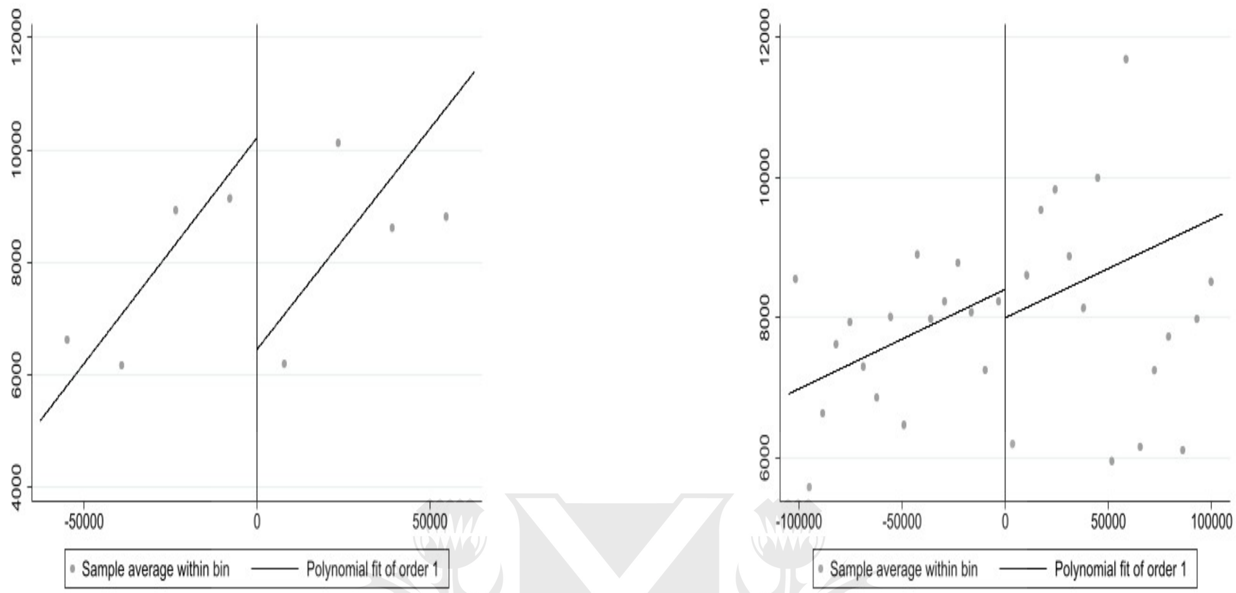


Table 9: Treatment effect on share over the per capita GDP

	Reelection Sample	Whole Sample
RD_Estimate	-3,765.5	-407.36
N	9148	27160
Effective observations	132	664
Left	87	500
Right	45	164
Bandwidth	65750	105151
	* p<0.1; ** p<0.05; *** p<0.01	

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at municipality-election level. Dependent variable: Per capita GDP. The statistical significance is determined by the robust standard errors.

Figure 8: Share of economically active population

(a) (b)

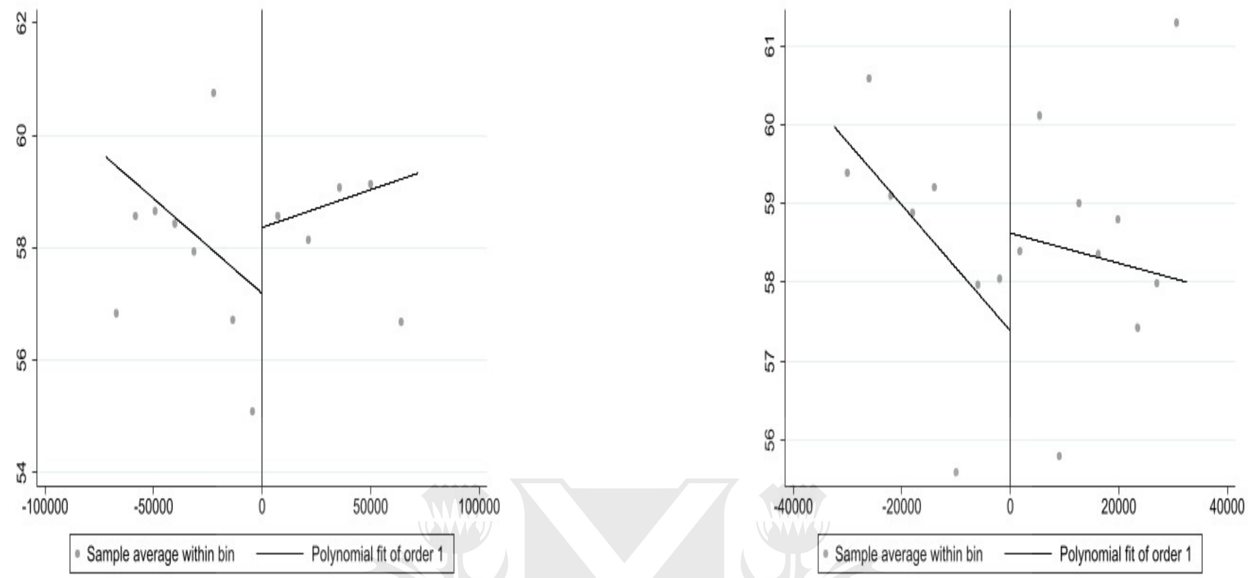


Table 10: Treatment effect over the share of economically active population

	Reelection sample	Whole Sample
RD_Estimate	1.1731	1.2367
N	9157	27195
Effective observations	149	166
Left	101	86
Right	48	80
Bandwidth	72048	32514
	* p<0.1; ** p<0.05; *** p<0.01	

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at municipality-election level. Dependent variable: Share of economically active populaiton. The statistical significance is determined by the robust standard errors.

Figure 9: Share of urban population

(a) (b)

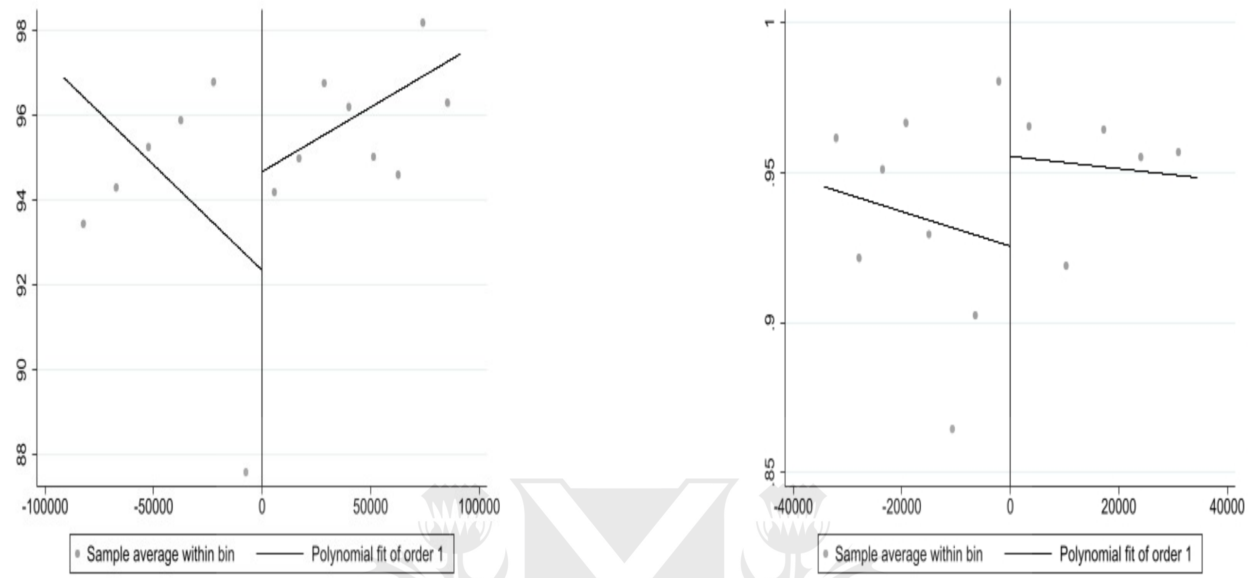


Table 11: Treatment effect over the share of urban population

	Reelection Sample	Whole Sample
RD_Estimate	2.3271	2.9991
N	9156	27190
Effective observations	190	178
Left	135	94
Right	55	84
Bandwidth	91437	34434
	* p<0.1; ** p<0.05; *** p<0.01	

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at candidate-election level. Dependent variable: Share of urban population. The statistical significance is determined by the robust standard errors.

5.3 Regression with higher polynomial degree

In this subsection we increase the order of the polynomial using a local quadratic fit instead of a local linear to reduce the approximation error in the estimation of the RD effect. We show that although the optimal bandwidths are larger the results of both outcomes are almost invariant.[16] In Table 12 and Figure 10 we present the results.

[16]It is not unusual to observe a change in the point estimate as one changes the polynomial order used in the estimation. Unless the higher-order terms in the approximation are exactly zero, incorporating those

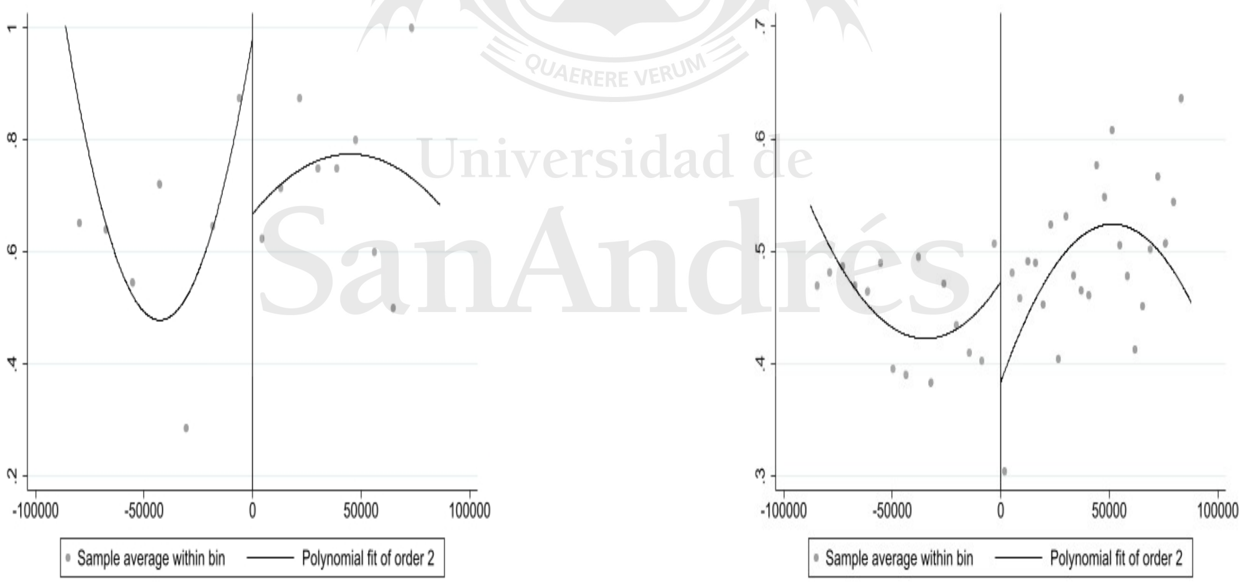
Table 12: Treatment effect on political competition outcomes with higher polynomial approximation

	Prob of reelection	Vote Share of the winner
RD_Estimate	-0.31515 *	-0.08892*
N	9250	27434
Effective observations	180	526
Left	127	375
Right	53	151
Bandwidth	86358	87754
	* p<0.1; ** p<0.05; *** p<0.01	

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at candidate-election level. Dependent variable: Probability of being reelected; Vote Share of the winning candidate. The statistical significance is determined by the robust standard errors.

Figure 10: Regression with higher polynomial degree

(a) Probability of being reelected (b) Vote share of the winner



6 Robustness check

In this section we will show that the result is robust to leaving out municipalities where the probability of a second round was negligible. To that end we drop the observations

terms in the estimation will reduce the approximation error and thus lead to changes in the estimated effect.

where cutoff is greater than zero and the vote share of the candidate that won the election is greater than 70%.

If we assume that the voters have a prior of how the elections will end, this result must be clearer than when we use all the sample since we drop the observations where there is no possibility of second round.[17] This is important for this research since we are studying the effect of a possible second round and this estimation allow us to conclude that the effect is nearly unchanged when we estimate the effect of a change from single-ballot to dual-ballot with a second round. Table 13 and Figure 11 display the results of this exercise (optimal bandwidths are used).

Table 13: Treatment effect on political competition outcomes dropping elections where a ballotage was impossible

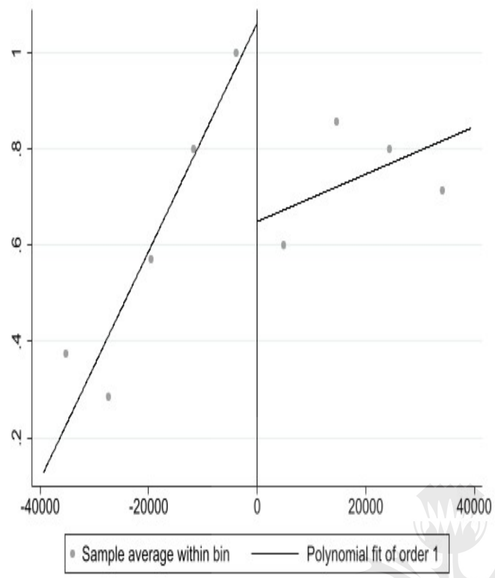
	Prob of reelection	Vote share of the winner
RD_Estimate	-0.28926 *	-0.10145**
N	9233	27402
Effective observations	89	211
Left	55	128
Right	34	83
Bandwidth	47968	41968
	* p<0.1; ** p<0.05; *** p<0.01	

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at candidate-election level. Dependent variable: Probability of being reelected; Vote Share of the winning candidate. The statistical significance is determined by the robust standard errors.

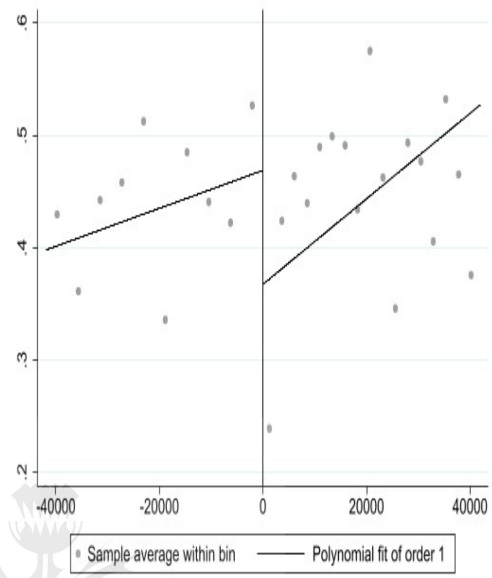
[17] Here the effect of stretegic vs sincere voting must be greater

Figure 11: Robustness check

(a) Probability of being reelected



(b) Vote share of the winner



7 Possible Mechanisms

In this section we discuss two possible, non-exclusive, mechanisms behind our results:

1. **Voters**: According to *Duverger's Law* strategic voting is more likely under SB than under DB. This could have a negative impact on the reelection chances of the incumbent or the vote share of the top candidate since it is not necessary to discard all candidates that have a low probability of being elected to be a pivotal voter.

2. **Parties**: Another prediction of *Duverger's Law* is that DB results in a larger number of political parties. Barone and De Blasio (2013) provide evidence for Italy that supports this prediction.

Bordignon et al. (2017) show theoretically that under DB there are more parties contending than under SB, because in the latter case it is more likely that candidates merge into parties. In Table 14 we provide evidence that the number of parties indeed increases when there is a change from SB to DB (column 1). In the last two columns we perform placebo tests with different cutoffs.

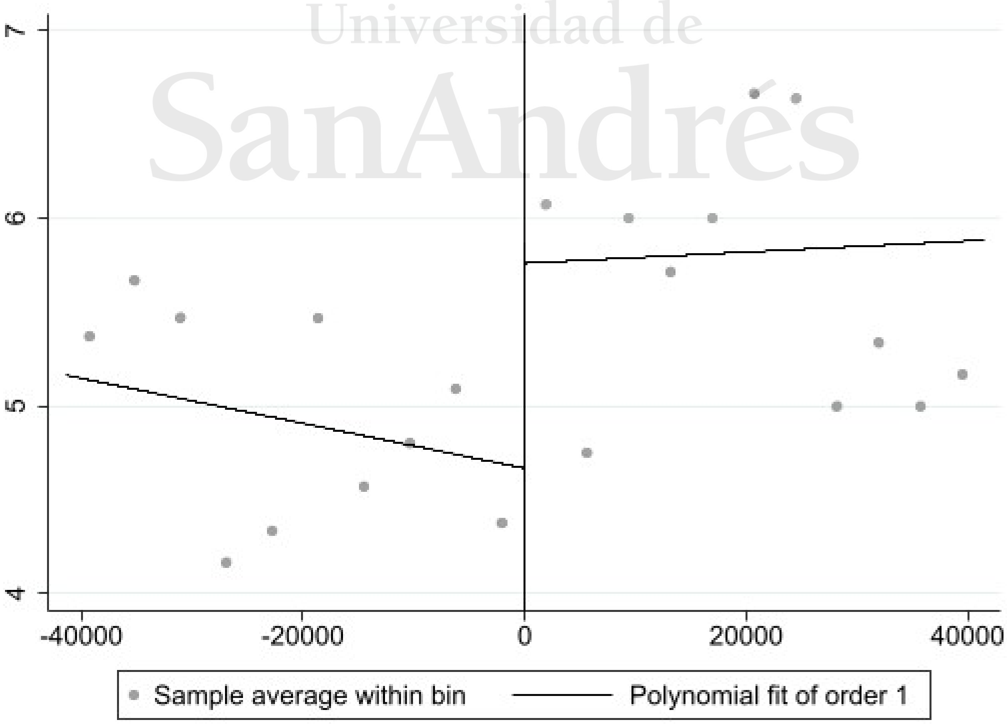
Table 14: Treatment effect on the number of parties.

	Cutoff	Cutoff+4000	Cutoff-4000
RD_Estimate	1.09178**	0.42072	0.58394
N	27434	27434	27434
Effective observations	215	410	364
Left	124	279	251
Right	91	131	113
Bandwidth	41444	69400	67543
	* p<0.1; ** p<0.05; *** p<0.01		

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at municipality-election level. Dependent variable: Number of parties. The statistical significance is determined by the robust standard errors.

Note that we run the regression at the municipality-election level in order to have the number of parties in a given election. We also examine whether this result is robust to different types of kernels .Figure 12 plots the number of parties against the forcing variable (number of registered voters minus 200,000), using the optimal bandwidth, the cutoff and a triangular kernel (Table 14 column 1).

Figure 12: Number of parties



Having more parties gives citizens more options. If we assume that citizens are distributed uniformly across the political spectrum and parties seek to differentiate themselves from each other, this produce a better match of a party to each voter's preferences.

Since Brazil is a country characterized by weak parties and strong individual candidates, we also examine whether a change from SB to DB increases the number of candidates.[18] Table 15 presents the results of this exercise at the true cutoff and two placebo cutoffs.

Table 15: Treatment effect on the number of candidates.

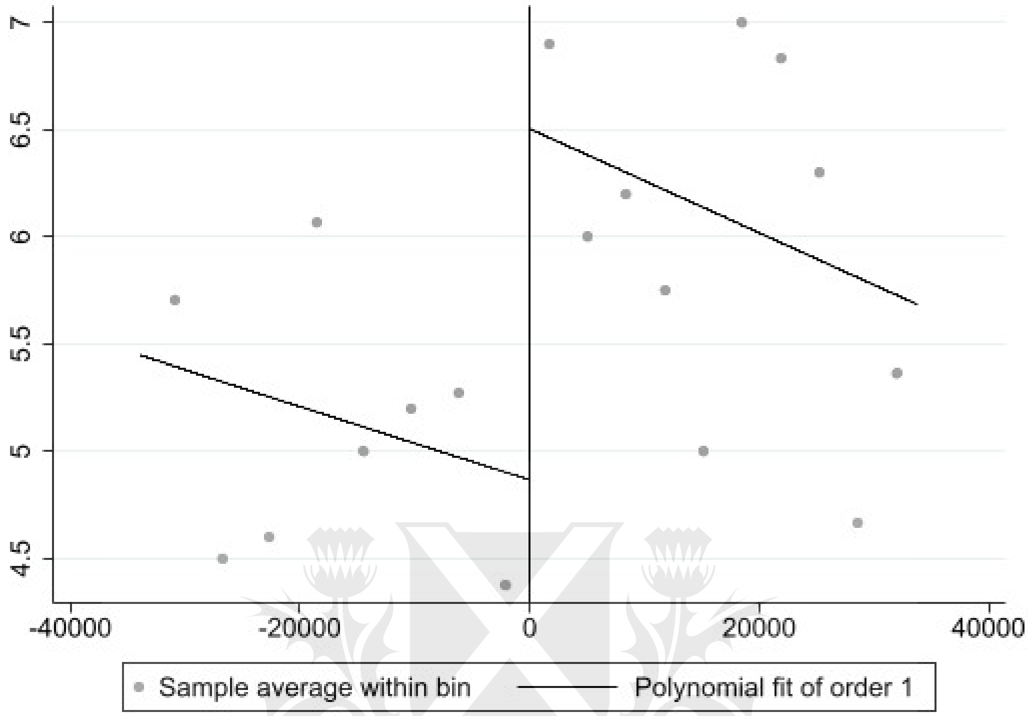
	Cutoff	Cutoff+4000	Cutoff-4000
RD_Estimate	1.63771***	0.61884	-0.10076
N	27434	27434	27434
Effective observations	173	345	451
Left	90	227	317
Right	83	118	134
Bandwidth	33878	59168	80112
	* p<0.1; ** p<0.05; *** p<0.01		

Notes. Election years between 2000 and 2016; the number of municipalities is ruled by the bandwidth. The dataset is at municipality-election level. Dependent variable: Number of candidates. The statistical significance is determined by the robust standard errors.

Figure 13 plots the number of candidates against the forcing variable, using the optimal bandwidth, the cutoff and a triangular kernel (Table 15 column 1).

[18]Chamon et al. (2019) also finds that the number of candidates increase from a change of SB to DB

Figure 13: Number of candidates



8 Concluding remarks

One of the most important factors in a country's development is who rules and what policies are implemented. This statement is also true for states and cities. We have shown evidence that the final outcome of an electoral race is affected by the electoral rules that are chosen – thereby contributing to further our understanding of the consequences of different voting schemes. Our results could be particularly relevant for public policies that aim to increase the competition between candidates in order to increase the quality of candidates.

This paper exploited a discontinuity in Brazilian electoral rules to show that DB elections increase political competition in more than one dimension. This result is in line with a large body of theoretical and empirical evidence on electoral rules and electoral competition such as Fujiwara (2011), Chamon et al. (2019), and Bordignon et al. (2016).

By taking advantage of the discontinuity in the electoral rule as a function of the electorate size, we can unequivocally identify the effect of a change from SB to DB over the personal incumbency advantage and the vote share of the winning candidate. The causal validity of this result is likely to hold given the quasi-random assignment of electoral rules generated by their discontinuous assignment across municipalities. The validity of this regression discontinuity design is supported by a number of validity tests and a robustness check.

References

Barone, G. and De Blasio, G. (2013). Electoral rules and voter turnout. *International Review of Law and Economics*, 36:25–35. 2, 22

Black, D. (1948). On the rationale of group decision-making. *Journal of political economy*, 56(1):23–34. 2

Bordignon, M., Nannicini, T., and Tabellini, G. (2016). Moderating political extremism: single round versus runoff elections under plurality rule. *American Economic Review*, 106(8):2349–70. 2, 7, 25

Bordignon, M., Nannicini, T., and Tabellini, G. (2017). Single round vs. runoff elections under plurality rule: A theoretical analysis. *European Journal of Political Economy*, 49:123–133. 2, 22

Brambor, T. and Ceneviva, R. (2011). Incumbency advantage in brazilian mayoral elections. In *APSA 2011 Annual Meeting Paper*. 3, 7

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372–404. 6, 9

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press. 6

Cattaneo, M. D., Jansson, M., and Ma, X. (2016). rddensity: Manipulation testing based on density discontinuity. *The Stata Journal (ii)*, pages 1–18. 4

Chamon, M., Firpo, S., Mello, J. M. d., and Pieri, R. (2019). Electoral rules, political competition and fiscal expenditures: regression discontinuity evidence from brazilian municipalities. *The Journal of Development Studies*, 55(1):19–38. 2, 5, 24, 25

Fujiwara (2011). A regression discontinuity test of strategic voting and duverger's law. *Quarterly Journal of Political Science*, 6(3-4):197–233. 2, 5, 7, 25

Klašnja, M. and Titiunik, R. (2013). Incumbency disadvantage in weak party systems: Evidence from brazil. *Unpublished manuscript, University of Michigan*. 3, 7

Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, 142(2):675–697. 3, 7

Titiunik, R. (2009). Incumbency advantage in brazil: Evidence from municipal mayor elections. *University of California-Berkeley, mimeo*. 7