

Supplemental Online Appendix to “The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability”

Marko Klašnja*
Post-Doctoral Fellow
Princeton University

Rocío Titiunik†
Assistant Professor
University of Michigan

INTENDED FOR ONLINE PUBLICATION ONLY

February 6, 2016

*Georgetown University, School of Foreign Service and Department of Government, ICC 302-N, 3700 O St NW, Washington, DC 20057. Email: <marko.klasnja@georgetown.edu>, Web: <http://markoklasnja.com>.

†Corresponding author. Department of Political Science, 505 South State St., 5700 Haven Hall, University of Michigan, Ann Arbor, MI 48109. Email: <titiunik@umich.edu>, Web: <http://www.umich.edu/~titiunik>.

Contents

S1	Overview	2
S2	Definition of Treated and Control Groups in Different Analysis	2
S3	Additional Details About Formal Model	5
S4	Formalization of the RD design	6
S5	Validity of the RD Design	8
S6	RD Effects Disaggregated by Year	15
S7	RD Effects on Candidacy $t + 1$, Conditional Victory $t + 1$ and Margin of Victory $t + 1$	17
S8	Bound Analysis for Conditional Victory $t + 1$	21
S9	Positive Incumbency Advantage Cannot Explain Negative Result	24
S10	Additional Career Path Analysis	25
	S10.1 Overall	25
	S10.2 PT Versus Other Parties	29
S11	Additional Analysis for Incumbent vs. Open Seat Samples	31
S12	Public Good Provision Indicators for PT	33
S13	Mayoral Victories: Incumbent versus Non-incumbent Candidates	34
S14	Histogram of Close Races in Brazil and United States	35
S15	Exploring the Correlation Between Negative Effects of Incumbency and Party Weakness Across Countries	36

S1 Overview

This document is the supplemental appendix to the manuscript “The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability”, and is intended for online publication only.

S2 Definition of Treated and Control Groups in Different Analysis

Table S1: Description of Treatment and Control Groups in Incumbent Party Analysis

Treatment Group
<ul style="list-style-type: none"> • A party wins at $t - 1$, so it is the incumbent party at t; • The same party runs again at t and (barely) wins, irrespective of whether its candidate is incumbent or a new candidate; • We analyze outcomes for the party at $t + 1$, when it is still an incumbent party (because it barely won election t) and has either an incumbent candidate running for reelection or a non-incumbent candidate.
Control Group
<ul style="list-style-type: none"> • A party wins at $t - 1$, so it is the incumbent party at t; • The same party runs again at t and (barely) loses, irrespective of whether its candidate is incumbent or a new candidate; • We analyze outcomes for that party at $t + 1$, when it is no longer the incumbent party (because it barely lost the t election), but some other first-term incumbent party either has an incumbent candidate who runs for reelection, has a new candidate, or does not have a candidate and it is an open race.
RD effect
<ul style="list-style-type: none"> • Party outcome at $t + 1$ when party is an incumbent running with either an incumbent candidate or a non-incumbent candidate vs. • Party outcome at $t + 1$ when the party is in opposition against either an incumbent or non-incumbent candidate of the party who won at t, or in open seat.

Table S2: Description of Treatment and Control Groups in Individual Party Analysis

Treatment Group
<ul style="list-style-type: none">• A party (barely) wins at t, irrespective of whether its candidate is incumbent or a new candidate;• We analyze outcomes for that party at $t+1$, when it is an incumbent party and has either an incumbent candidate running for reelection or a non-incumbent candidate.

Control Group
<ul style="list-style-type: none">• A party (barely) loses at t, irrespective of whether its candidate is incumbent or a new candidate;• We analyze outcomes for that party at $t + 1$, when it is not an incumbent party, but some other first-term incumbent party either has an incumbent candidate who runs for reelection, or has a new candidate, or does not have a candidate and there is no incumbent party in the race.

RD effect
<ul style="list-style-type: none">• Party outcome at $t + 1$ when party is an incumbent party running with either an incumbent candidate or a non-incumbent candidate vs.• Party outcome at $t + 1$ when the party is in opposition against either an incumbent or non-incumbent candidate of the party who won at t, or there is no incumbent party in the race.

S3 Additional Details About Formal Model

Proof of Proposition 1 in the main text:

Proof. In the third period, there are no reelection incentives. Given the player utilities, neither type of politician exerts effort, and neither type of party exacts punishment. The voter is thus indifferent with respect to the possible payoff, as she always receives 0. Therefore, she focuses on $t = \{1, 2\}$.

Given the voter's and party's strategies, the good politician receives $r_1 - g_1 + r_2 - g_2 = 2(r - 1)$ from providing the public good ("working") in both periods. Working in the first period, shirking in the second period and being punished by the party brings $r_1 - g_1 + r_2 - p_2 = 2(r - 1)$. Shirking in the first period brings $r_1 - p_1$. Since $r_t > g_t$, the good politician prefers providing the public good in the first period to shirking, all else equal. In turn, the party does not need to exact punishment in the first period, i.e. $p_1 = 0$. If facing the prospect of punishment due to shirking, the good politician therefore gets at least as much benefit from working in both periods as shirking in the second period. Therefore, the good politician will work in the first period. Combined with the observation that the bad politician never provides $g_t = 1$, the voter and the party perfectly infer the politician's type in the first period. This allows the voter and the party to fully condition the second-period strategy solely on the observation of the second-period outcome and second-period punishment.

Since by assumption $c(1) = L \leq r$ for any $\kappa = L$, the party facing a low cost will punish the shirking politician. On the other hand, it is immediate that the party will choose $p_2 = 0$ if $c(1) = H > r$. The parties perfectly separate in the second period, and so the voter chooses $\sigma(g_2 = 0, p_2 = 0) = 0$, thus punishing the high-cost party. This establishes the claim in the proposition. \square

We assume in the model that voters do not directly observe the type of the politician or the strength of the party. These assumptions are important for the result in Proposition 1.¹ We believe that these assumptions are not unrealistic; informational asymmetry is commonly invoked in other standard agency models (e.g. Besley 2007; Persson and Tabellini 2002). Moreover, the main purpose of our model is to identify meaningful implications that can be tested empirically, rather than build a comprehensive and robust theory of electoral dynamics with term limits and weak parties, which we believe merits separate effort. Nonetheless, here we discuss how changing the assumptions affects our results and what alternative assumptions can be made to get the same result as in Proposition 1.

Suppose that all politicians are "good" and that a party's type is known to the voter. Then, there is simple perfect "(non)compliance:" when $\kappa = H$ is sufficiently low, then all parties punish, all politicians provide the public good, and voters always reelect. If the high cost is too high, then no party punishes, no politician behaves, and the voter never reelects. Empirically, this type of result helps us distinguish electoral consequences for strong and weak parties, but it does not help us draw implications of the interaction of party strength and term limits.

¹Another assumption on the information structure is that the party does not observe the politician's type. While this is in line with our conceptual framework of a politician being an agent of both the party and the voter, our results in the paper come through if we relax this assumption and allow the party to observe the type of the politician.

Suppose instead that party's strength is unobserved by the voter, but that all politicians are good, implying that all politicians can be induced to provide the public good either by the party or the voter. In this case, when faced with an election in t_1 , voters know that they can throw the incumbent out and that if they do so they will be guaranteed a politician they can control in t_2 , because that politician will face reelection. Unless there are some politicians they cannot control and there is some uncertainty about the politician's type, they may want to pursue this strategy in t_1 , particularly if the probability of a weak party is high.² This again rules out our main result, because we get an equilibrium with no lame duck politicians, which clearly deviates from empirical evidence.

An alternative to assuming politicians of different type and imperfect information about them is to assume that the voter can ex ante credibly commit to a voting rule, for example:

$$\sigma(g_t) = \begin{cases} 0 & \text{if } g_t = 0 \\ 1 & \text{if } g_t = 1, \end{cases}$$

In this case, the voter is indifferent between getting $g_1 = 1; g_2 = 0; g_3 = 0$ if she reelects the politician who provides the public good in the first period (whatever the party strength) and $g_1 = 0; g_2 = 1; g_3 = 0$ if she induces the strong party to reveal its type in the first period, thereby giving the same result as in Proposition 1.

S4 Formalization of the RD design

Let municipality i at election t have J political parties that dispute the municipal mayoral elections. For $j = 1, \dots, J$, let $V_{it,j}$ be the vote share obtained by party j in municipality i in election t and $V_{it,(1)}, \dots, V_{it,(J)}$ be the corresponding order statistics. The margin of victory for party k is defined as the vote share obtained by party k minus the vote share obtained by party k 's strongest opponent, where the latter is defined as the party that obtains the highest vote share if party k loses the election and the party that obtains the *second* highest vote share if party k wins. Formally, party k 's margin of victory is given by:

$$M_{it,k} \equiv \begin{cases} V_{it,k} - V_{it,(J-1)} & \text{if } V_{it,k} = V_{it,(J)} \\ V_{it,k} - V_{it,(J)} & \text{otherwise.} \end{cases} \quad (1)$$

It follows that the rule that determines the incumbency status of party k at election $t+1$ in municipality i , denoted by $I_{it+1,k}$ is:

$$I_{it+1,k} = \begin{cases} 1 & \text{if } M_{it,k} \geq 0 \\ 0 & \text{if } M_{it,k} < 0. \end{cases} \quad (2)$$

Let $Y_{it+1,k}^1$ denote the outcome of interest for party k in municipality i at election $t+1$ when $I_{it+1,k} = 1$ and $Y_{it+1,k}^0$ denote the outcome of interest for party k when $I_{it+1,k} = 0$. The effect of interest is $\tau_k \equiv \mathbb{E}(Y_{it+1,k}^1 - Y_{it+1,k}^0)$. Of course, for a given election in a

²When the probability that the party is weak (call it γ) is high, the voter knows that using the strategy of reelecting when seeing $g_1 = 1$ implies there is a high chance that she will get $g_2 = 0$. By design, she will get $g_3 = 0$. Therefore, her expected utility is close to 1. But she then may decide to use the strategy of not reelecting when seeing $g_1 = 1, p_1 = 0$ in order to induce the low-cost party to reveal its type in the first period. If it does so, the voter will know that she will get $g_2 = 1$ for sure, and would be willing to reelect even if $g_1 = 0$ but $p_1 = 1$. In this case, our result is not an equilibrium for all values of γ .

given municipality, a party cannot be the incumbent and not the incumbent simultaneously, and hence one only observes $Y_{it+1,k} = I_{it+1,k}Y_{it+1,k}^1 + (1 - I_{it+1,k})Y_{it+1,k}^0$. Without further assumptions, τ_k cannot be recovered.

Assuming that $\mathbb{E}(Y_{it+1,k}^1|M)$ and $\mathbb{E}(Y_{it+1,k}^0|M)$ are continuous at $M_{it,k} = 0$ (Hahn, Todd, and van der Klaauw 2001), the expected causal effect of incumbency status on the outcome of interest can be recovered from observed outcomes at the discontinuity point. Formally,

$$\tau_k^{RD} \equiv \mathbb{E}(Y_{it+1,k}^1 - Y_{it+1,k}^0|M = 0) = \lim_{M \downarrow 0} \mathbb{E}(Y_{it+1,k}|M) - \lim_{M \uparrow 0} \mathbb{E}(Y_{it+1,k}|M)$$

Therefore, the discontinuity in the rule that determines which party wins office provides an opportunity to observe the average difference in potential outcomes by comparing points on either side of the $M_{it,k} = 0$ threshold. The crucial assumption is the continuity of the expected potential outcomes at the threshold. This assumption is inherently unobservable but, as we show in Section S5 below, several empirical tests strongly support this assumption in our data.

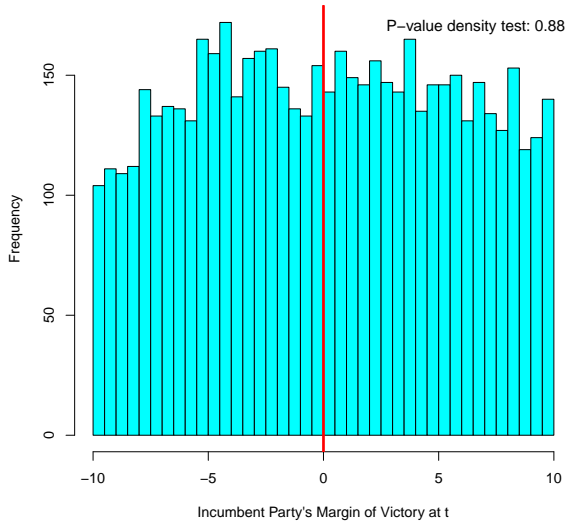
S5 Validity of the RD Design

We assess the validity of the RD design in our application. In order to interpret the outcome change that occurs at the vote margin cutoff as the effect of the party winning the mayoral office, any factors correlated with both the outcome of interest and electoral victory at t must be continuous at the cutoff. This is an identification assumption and, as such, is untestable. But if this assumption is true, it is reasonable to expect that certain empirical implications will hold, and it is now standard to present empirical evidence to validate this assumption.

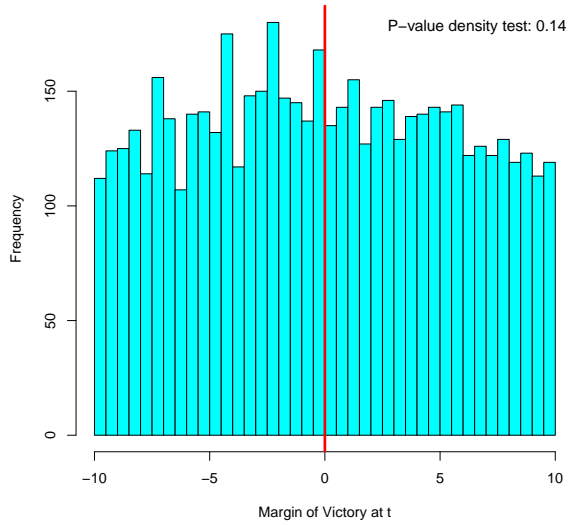
The first piece of evidence we show is that the density of the running variable, the party's vote margin at t , is not discontinuous at the cutoff. If parties had the ability to influence precisely whether they lose or win, we would likely observe very few parties that barely lose, and many more parties that barely win. Since manipulation of electoral results at t would likely be correlated with future electoral performance, a discontinuity in the density of the vote margin at t right around the cutoff might raise doubts about the design. But we do not observe any such discontinuity. Figures S1(a), S1(b), S1(c) and S1(d) show histograms of the margin of victory at election t for the incumbent party, the PMDB, the PSDB and the DEM party. Each figure also reports the p -value of the null hypothesis that the density of the running variable is continuous at the cutoff using the local polynomial density estimator developed by Cattaneo, Jansson, and Ma (2015a) (see Cattaneo, Jansson, and Ma 2015b, for details about Stata implementation). The figures show that there is nothing peculiar about the parties' vote margin at t around the cutoff, a conclusion that we corroborate formally as the density test fails to reject the null hypothesis in all cases (p-values range from 0.14 to 0.88).³

³The p-value corresponding to the density test of the PP and PT parties (the two parties not included in Figure S1) are, respectively, 0.60 and 0.81.

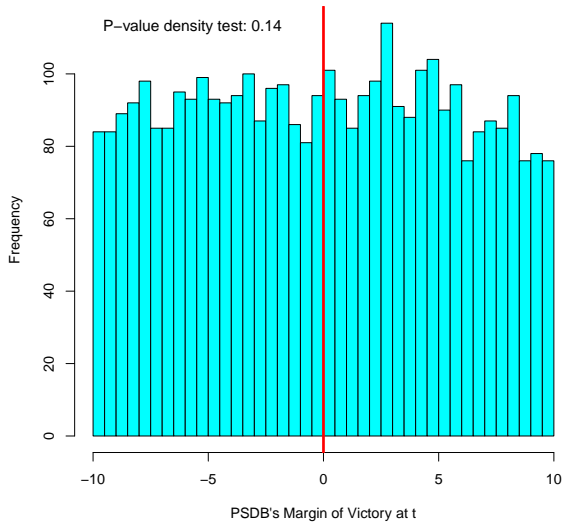
Figure S1: Histogram of margin of victory for various parties



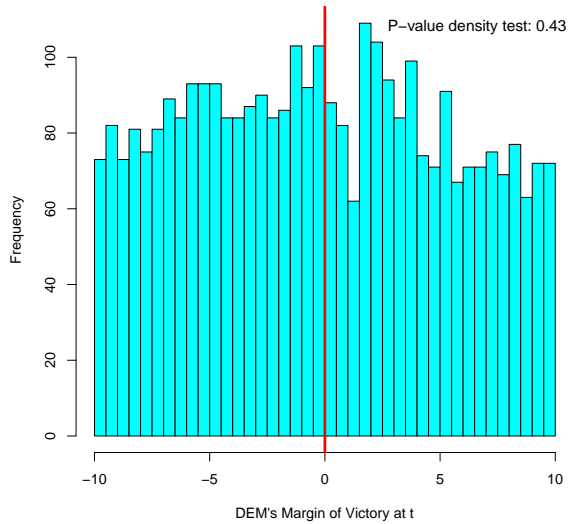
(a) Incumbent Party, 2000-2012



(b) PMDB, 1996-2012



(c) PSDB, 1996-2012



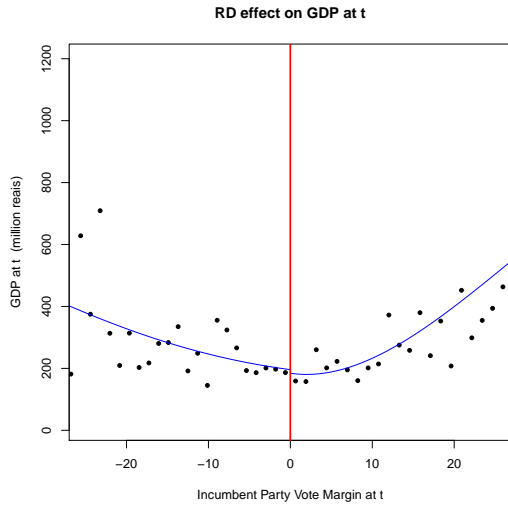
(d) DEM, 1996-2012

In addition, we estimate “placebo” RD effects on *predetermined* covariates—covariates that are determined before the treatment is assigned at t . Since treatment is assigned after these covariates are realized and measured, we expect the effect of parties’ barely winning at t on these covariates to be indistinguishable from zero. A significant effect would be a strong indication that unobserved confounders are spuriously causing the effects on the outcomes of interest. We present some results graphically, and also include full details of the analysis in Tables S3 and S4.

Figures S2(a), S2(b) and S2(c) show the effect of the incumbent party barely winning on the municipality's GDP at t , the municipality's population at t , and the number of effective parties in the mayoral election at t , respectively, while Figures S2(d), S2(e) and S2(f) show the effect of barely winning on the victory in the previous election for the PMDB, PSDB and DEM parties, respectively. As the figures show, the RD effect is indistinguishable from zero in all cases, as expected in a valid RD design. The formal RD estimation results are provided in Table S3 for all the covariates reported in the figures and also for additional covariates, for the effect of the incumbent party winning. In addition, Table S4 reports RD effects for every individual party—in particular, it reports the effect of each individual party winning on the party's lagged victory, a very important predetermined covariate.

Finally, we estimate the same effects as in Table 2 in the main paper, but in each case including covariates in the local regression estimation. The covariates included are GDP, population, number of effective parties, dummy for municipality located in the north, dummy for municipality located in south, dummy for municipality located in northeast, total municipality revenues and total municipality expenditures.

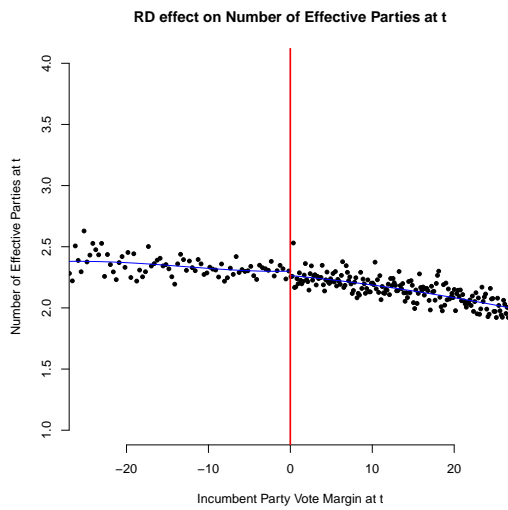
Figure S2: RD effects on various predetermined covariates



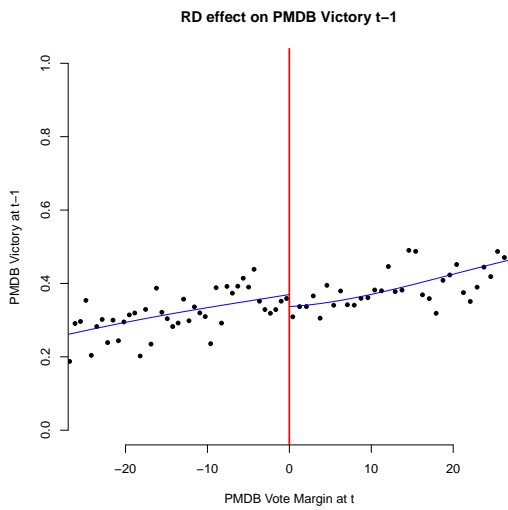
(a) Incumbent Party, GDP t



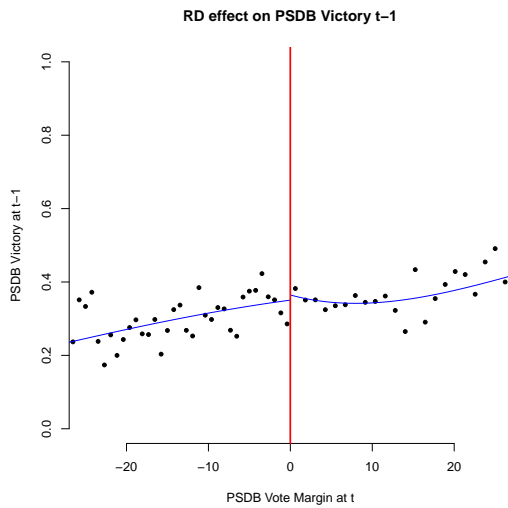
(b) Incumbent Party, Population t



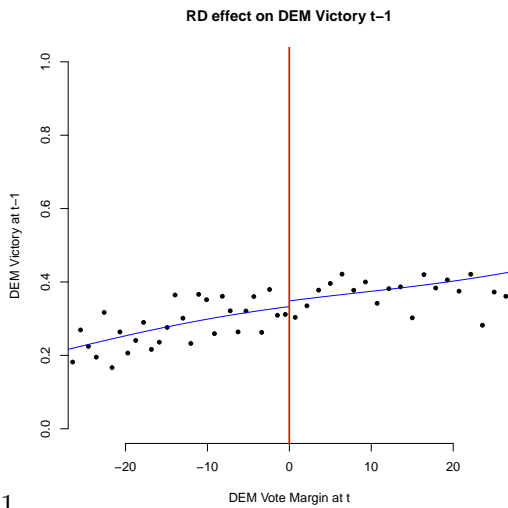
(c) Incumbent Party, Effective Parties t



(d) PMDB, Victory t-1



(e) PSDB, Victory t-1



(f) DEM, Victory t-1

Table S3: RD Effect of Victory on Predetermined Covariates—Incumbent Party

Party	Outcome	Estimate	CI	pval	h	Ntr	Nco
No. effective parties	All seats	-0.04	[-0.1,0.02]	0.17	12.78	3485	3374
	Incumbent	-0.05	[-0.13,0.01]	0.11	12.74	2174	1897
	Open Seat	-0.01	[-0.1,0.08]	0.86	15.15	1302	1497
GDP per capita	All seats	-149.03	[-1211.61,1016.49]	0.86	13.48	3593	3484
	Incumbent	-141.63	[-1562.93,1191.51]	0.79	16.95	2696	2284
	Open Seat	64.25	[-1621.95,2029.37]	0.83	15.27	1307	1501
Population	All seats	756.02	[-2557.52,5102.17]	0.51	10.74	3022	2948
	Incumbent	-886.54	[-5545.19,4773.53]	0.88	10.76	1887	1688
	Open Seat	3527.11	[-1585.33,10158.72]	0.15	9.75	976	1069
Winner age	All seats	1.30	[0.33,2.61]	0.01	15.00	3478	3432
	Incumbent	1.69	[0.42,3.15]	0.01	18.50	2591	2174
	Open Seat	1.74	[-0.08,4.31]	0.06	10.83	930	1059
Winner educ	All seats	-0.03	[-0.21,0.16]	0.83	18.73	4035	3889
	Incumbent	0.02	[-0.19,0.26]	0.74	20.77	2789	2283
	Open Seat	-0.05	[-0.33,0.23]	0.71	23.99	1436	1732
Winner male	All seats	0.00	[-0.02,0.03]	0.72	20.23	4219	4044
	Incumbent	0.01	[-0.03,0.05]	0.54	18.54	2595	2175
	Open Seat	-0.01	[-0.05,0.05]	0.90	17.81	1274	1509
Total expenditures	All seats	-2339779.86	[-6633890.15,3003351.15]	0.46	6.48	1816	1796
	Incumbent	-4900398.70	[-11565585.7,1756952.1]	0.15	6.01	1007	979
	Open Seat	3708757.94	[-2633914.95,11940171.65]	0.21	8.00	795	857
Total revenue	All seats	-2199736.84	[-6477261.51,3242688.59]	0.51	7.48	2087	2046
	Incumbent	-5167476.15	[-12438638.92,1611321.55]	0.13	6.27	1047	1017
	Open Seat	3403149.14	[-3104914.81,11778503.05]	0.25	7.95	792	852
Centerwest	All seats	-0.01	[-0.03,0.01]	0.43	19.03	4554	4322
	Incumbent	-0.01	[-0.04,0.03]	0.83	17.52	2787	2332
	Open Seat	-0.02	[-0.07,0.02]	0.36	16.70	1377	1592
Northeast	All seats	0.01	[-0.03,0.06]	0.62	16.86	4220	4051
	Incumbent	0.00	[-0.06,0.06]	0.97	16.58	2658	2263
	Open Seat	0.02	[-0.06,0.1]	0.64	16.94	1390	1604
North	All seats	-0.02	[-0.06,0.01]	0.11	12.27	3366	3284
	Incumbent	0.01	[-0.03,0.04]	0.78	17.43	2776	2325
	Open Seat	-0.03	[-0.09,0.01]	0.10	12.94	1194	1353
Southeast	All seats	-0.01	[-0.06,0.03]	0.48	16.85	4220	4050
	Incumbent	-0.05	[-0.12,0.01]	0.07	15.78	2561	2195
	Open Seat	0.03	[-0.04,0.1]	0.38	17.46	1409	1632

Note: Running variable is party's vote margin at t . Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 4-8 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

Table S4: RD Effect of Victory on Unconditional Victory for Individual Parties

Party	Outcome	Estimate	CI	pval	h	Ntr	Nco
Victory t-1 PSDB	All seats	0.01	[-0.05,0.09]	0.51	15.73	1955	2022
	Incumbent	-0.02	[-0.15,0.12]	0.82	16.42	631	517
	Open Seat	0.05	[-0.01,0.13]	0.11	13.48	1199	1355
Victory t-1 PMDB	All seats	-0.04	[-0.1,0.01]	0.13	18.56	3152	3457
	Incumbent	-0.07	[-0.17,0.03]	0.15	17.35	866	706
	Open Seat	-0.02	[-0.07,0.04]	0.59	14.96	2015	2319
Victory t-1 DEM	All seats	0.02	[-0.06,0.09]	0.68	18.16	1733	1961
	Incumbent	-0.05	[-0.19,0.07]	0.36	18.41	512	469
	Open Seat	0.03	[-0.05,0.09]	0.52	16.43	1157	1396
Victory t-1 PP	All seats	0.03	[-0.05,0.13]	0.34	12.42	1246	1199
	Incumbent	-0.08	[-0.2,0.04]	0.21	15.20	397	298
	Open Seat	0.02	[-0.05,0.11]	0.47	14.08	955	1032
Victory t-1 PT	All seats	0.01	[-0.06,0.07]	0.87	25.96	1429	1856
	Incumbent	-0.04	[-0.14,0.03]	0.21	13.44	230	176
	Open Seat	0.06	[0.01,0.13]	0.02	14.95	839	1002

Note: Running variable is party's vote margin at t . Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 4-8 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

Table S5: RD effect of Incumbency at t on Victory at $t + 1$ (Unconditional on Running) for Various Parties with Covariates—Brazil Mayoral Elections, 1996-2012

OUTCOME: UNCONDITIONAL VICTORY $t + 1$						
Party	Estimate	95% CI	pval	h	Ntr	Nco
INCUMBENT	-0.14	[-0.201,-0.041]	0.0029	13.55	2645	2378
PMDB	-0.13	[-0.193,-0.001]	0.0484	14.63	1978	2121
PSDB	-0.03	[-0.093,0.107]	0.8904	18.77	1645	1649
DEM	-0.10	[-0.251,-0.019]	0.0226	12.15	1099	1190
PP	-0.20	[-0.388,-0.116]	0.0003	12.43	937	889

Note: Running variable is party's vote margin at t , outcome is dummy =1 if party wins the following election at $t + 1$, =0 otherwise. Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and including covariates, with MSE-optimal bandwidth chosen according to CCT. Columns 3-7 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

S6 RD Effects Disaggregated by Year

Table S6: Yearly RD effect of Incumbency at t on Victory at $t + 1$ (Unconditional on Running) for Various Parties—Brazil Mayoral Elections, 1996-2012

OUTCOME: UNCONDITIONAL VICTORY $t + 1$							
Year	Party	Estimate	CI	pval	h	Ntr	Nco
1996	PMDB	-0.10	[-0.209,-0.008]	0.03	16.70	832	923
1996	PSDB	-0.04	[-0.189,0.084]	0.45	16.31	575	572
1996	DEM	-0.06	[-0.198,0.051]	0.25	16.07	596	690
1996	PP	-0.05	[-0.172,0.089]	0.53	19.76	449	551
1996	PT	-0.14	[-0.441,0.12]	0.26	18.39	85	145
2000	INCUMBENT	-0.19	[-0.281,-0.122]	0.00	18.33	1120	1104
2000	PMDB	-0.14	[-0.239,-0.011]	0.03	15.49	760	806
2000	PSDB	-0.12	[-0.244,-0.02]	0.02	21.70	653	666
2000	DEM	-0.23	[-0.376,-0.135]	0.00	12.38	502	534
2000	PP	-0.26	[-0.435,-0.14]	0.00	12.92	345	337
2000	PT	-0.18	[-0.498,0.099]	0.19	18.05	120	186
2004	INCUMBENT	-0.09	[-0.209,-0.013]	0.03	13.31	881	890
2004	PMDB	-0.06	[-0.187,0.044]	0.22	15.61	693	800
2004	PSDB	0.04	[-0.087,0.173]	0.52	17.36	588	595
2004	DEM	0.00	[-0.125,0.106]	0.87	15.94	520	570
2004	PP	-0.27	[-0.48,-0.127]	0.00	10.09	284	279
2004	PT	0.00	[-0.198,0.153]	0.80	16.02	278	337
2008	INCUMBENT	-0.16	[-0.259,-0.073]	0.00	16.53	1115	856
2008	PMDB	-0.18	[-0.319,-0.066]	0.00	13.70	688	683
2008	PSDB	-0.03	[-0.163,0.108]	0.69	15.91	466	470
2008	DEM	-0.10	[-0.243,0.01]	0.07	16.98	310	360
2008	PP	-0.15	[-0.312,-0.008]	0.04	18.97	395	360
2008	PT	-0.15	[-0.311,0.019]	0.08	21.11	405	396

Note: Running variable is party's vote margin at t , outcome is dummy =1 if party wins the following election at $t + 1$, =0 otherwise. Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 4-8 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

Table S7: Yearly RD effect of Incumbency at t on Victory at $t + 1$ (Conditional on Running) for Various Parties—Brazil Mayoral Elections, 1996-2012

OUTCOME: CONDITIONAL VICTORY $t + 1$							
Year	Party	Estimate	CI	pval	h	Ntr	Nco
1996	PMDB	-0.15	[-0.283,-0.017]	0.03	15.21	610	561
1996	PSDB	-0.15	[-0.341,0.008]	0.06	16.33	414	303
1996	DEM	-0.12	[-0.291,0.032]	0.12	14.93	409	373
1996	PP	-0.14	[-0.325,0.073]	0.22	18.30	277	238
1996	PT	-0.13	[-0.481,0.223]	0.47	16.01	63	93
2000	INCUMBENT	-0.32	[-0.471,-0.21]	0.00	16.13	575	569
2000	PMDB	-0.22	[-0.351,-0.088]	0.00	21.13	618	597
2000	PSDB	-0.14	[-0.327,0.094]	0.28	14.38	291	274
2000	DEM	-0.28	[-0.467,-0.139]	0.00	16.66	357	355
2000	PP	-0.32	[-0.555,-0.137]	0.00	13.69	204	186
2000	PT	-0.22	[-0.558,0.077]	0.14	18.62	106	150
2004	INCUMBENT	-0.12	[-0.264,-0.017]	0.03	16.73	638	562
2004	PMDB	-0.11	[-0.271,0.006]	0.06	16.13	524	493
2004	PSDB	-0.05	[-0.246,0.112]	0.46	19.74	414	307
2004	DEM	0.07	[-0.126,0.287]	0.44	15.26	248	194
2004	PP	-0.31	[-0.578,-0.115]	0.00	10.64	186	143
2004	PT	-0.07	[-0.298,0.109]	0.36	18.63	259	227
2008	INCUMBENT	-0.20	[-0.343,-0.083]	0.00	19.92	716	504
2008	PMDB	-0.33	[-0.502,-0.186]	0.00	13.17	483	388
2008	PSDB	-0.08	[-0.269,0.13]	0.50	17.06	304	252
2008	DEM	-0.38	[-0.687,-0.123]	0.00	18.07	139	94
2008	PP	-0.45	[-0.735,-0.241]	0.00	13.50	189	137
2008	PT	-0.16	[-0.35,0.102]	0.28	13.63	248	192

Note: Running variable is party's vote margin at t , outcome is dummy =1 if party wins the following election at $t + 1$, =0 if it runs and loses. Sample includes only municipalities where party contests the $t + 1$ election. Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 4-8 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

S7 RD Effects on Candidacy $t + 1$, Conditional Victory $t + 1$ and Margin of Victory $t + 1$

In this section, we consider the effects of barely winning on additional outcomes not reported in the main body of the paper. First, we analyze the effect of barely winning at t on whether the party is a candidate at $t + 1$, which we treat as an outcome in its own right. Second, we analyze the effect of barely winning at t on the party's conditional victory at $t + 1$ —the party's victory at $t + 1$ given that the party t contests the $t + 1$ election. Finally, we report the effects of barely winning on margin of victory at $t + 1$.

Table S8 presents the results for candidacy at $t + 1$, a dummy variable that is equal to one if the party contests the mayoral election at $t + 1$ and equal to zero otherwise. The first row reports the results from the incumbent party analysis for the 2000-2012 period; the other rows report the individual party analysis in the full 1996-2012 period for the four largest parties described above: PMDB, PSDB, DEM and PP. The results for the incumbent party indicate that when the incumbent party barely wins the t election it is 6 percentage points less likely to contest the following election than when it barely loses, and this effect is different from zero at 5% level (the robust 95% confidence interval ranges between -0.12 and -0.001). We show this result graphically in Figure (S3a). This negative effect on candidacy at $t + 1$, however, is not observed for any of the individual parties analyzed.

Table S8: RD effect of Incumbency at t on Candidacy at $t + 1$ for Various Parties

Brazil Mayoral Elections, 1996-2012

OUTCOME: CANDIDACY $t + 1$						
Party	Estimate	95 CI	pval	h	Ntr	Nco
INCUMBENT	-0.06	[-0.119,-0.001]	0.05	17.55	3295	3012
INDIVIDUAL PARTIES						
PMDB	0.04	[-0.012,0.084]	0.14	20.38	3498	3868
PSDB	0.06	[-0.006,0.124]	0.08	20.27	2464	2522
DEM	-0.01	[-0.091,0.045]	0.51	15.36	1965	2184
PP	0.00	[-0.087,0.073]	0.87	16.55	1579	1615

Note: Running variable is party's vote margin at t , outcome is dummy =1 if party contests the following election at $t + 1$, =0 otherwise. Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 3-7 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

Table S9 shows the results for margin of victory at $t + 1$. Since the margin of victory is

undefined for races where a party does not contest the election, these results condition on the party contesting the $t + 1$ election.

Table S9: RD effect of Incumbency at t on Margin of Victory at $t + 1$ (Conditional on Running) for Various Parties

Brazil Mayoral Elections, 1996-2012

OUTCOME: VOTE MARGIN AT $t + 1$						
Party	Estimate	95 CI	pval	h	Ntr	Nco
INCUMBENT	-7.39	[-10.865,-3.974]	0.00	12.67	1496	1361
INDIVIDUAL PARTIES						
PMDB	-7.92	[-11.114,-5.212]	0.00	12.90	1890	1740
PSDB	0.05	[-3.801,4.761]	0.83	16.03	1361	1115
DEM	-4.73	[-9.472,-0.717]	0.02	13.36	1029	911
PP	-7.73	[-12.3,-4.088]	0.00	13.01	800	675
PT	-7.21	[-12.137,-1.671]	0.01	15.59	653	610

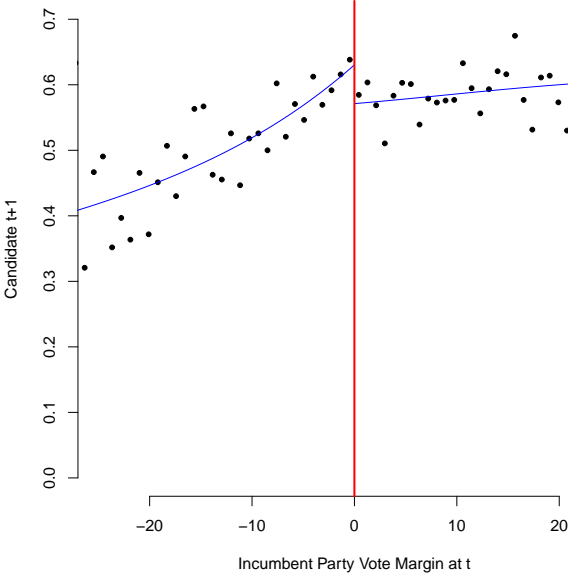
Note: Running variable is party's vote margin at t , outcome is margin of victory at in the following election at $t + 1$. Sample includes only municipalities where party contests the $t + 1$ election. Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 3-7 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

Table S10 shows the results for electoral victory at $t + 1$, conditional on the party contesting the $t + 1$ election. The outcome analyzed in this table is a dichotomous variable that is equal to one if the party contested and won the $t + 1$ election, and equal to zero if the party contested but lost the election.

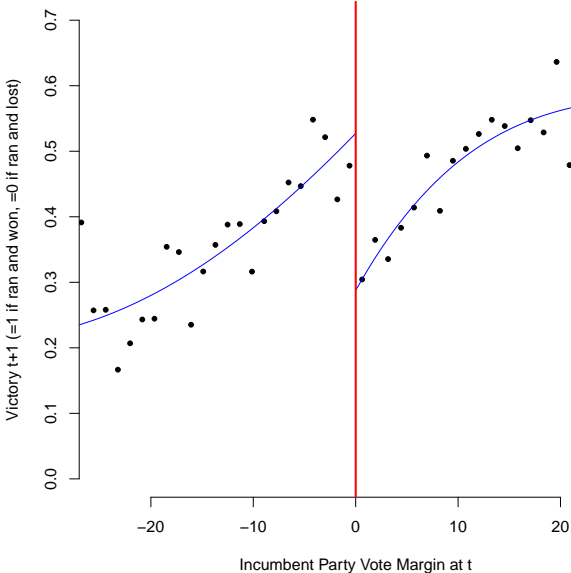
The results are strongly negative for both the incumbent party and the individual party analysis, in sharp contrast with the positive results that are commonly found in the American literature (see, e.g., Ansolabehere and Snyder Jr 2002; Cattaneo, Frandsen, and Titiunik 2015; Erikson and Titiunik 2015). The effect in the first row indicates that in those municipalities where the incumbent party is barely reelected at t it is 21 percentage points less likely to win at $t + 1$ than in those municipalities where it barely lost at t (among those municipalities where the incumbent party contests the $t + 1$ election). The illustration of this effect in Figure (S3b) shows that the incumbent party wins in about 50% of the municipalities where it barely lost at t and runs at $t + 1$, where it only wins in roughly 30% of the municipalities where it barely won at t and runs at $t + 1$. The effects are strongly statistically significant. These negative results are seen in the individual party analysis as

well, with point estimates that range between -0.11 and -0.27, all strongly distinguishable from zero.

Figure S3: RD effects on for incumbent party



(a) Candidate t+1



(b) Victory t+1 (given running)

These results are stronger (more negative) than the unconditional effects reported in the

Table S10: RD effect of Incumbency at t on Victory at $t + 1$ (Conditional on Running) for Various Parties—Brazil Mayoral Elections, 1996-2012

OUTCOME: CONDITIONAL VICTORY $t + 1$						
Party	Estimate	95% CI	pval	h	Ntr	Nco
INCUMBENT	-0.21	[-0.303,-0.14]	0.00	14.33	1664	1470
PMDB	-0.19	[-0.272,-0.127]	0.00	14.68	2088	1911
PSDB	-0.11	[-0.206,-0.029]	0.01	19.33	1571	1242
DEM	-0.15	[-0.259,-0.059]	0.00	14.51	1098	960
PP	-0.28	[-0.41,-0.184]	0.00	13.75	845	690

Note: Running variable is party's vote margin at t , outcome is dummy =1 if party wins the following election at $t + 1$, =0 if it runs and loses. Sample includes only municipalities where party contests the $t + 1$ election. Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 3-7 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

paper. However, given the significant effect of incumbency on Candidacy at $t + 1$ reported in Table S8 and the fact that parties choose not to contest a large proportion of mayoral elections, looking at the effects of incumbency on future victory only for the subset of municipalities where the party contests the $t + 1$ election could introduce biases. There are two issues that should be considered. The first is whether the difference in the proportion of contested elections at $t + 1$ between the treatment and control groups is driving the negative results. Consider an example with 100 municipalities in each of the treatment and control groups where, at $t + 1$, (i) the party contests all 100 elections in the treatment group but only 50 elections in the control group, and (ii) the party wins 25 elections at $t + 1$ in each group. Analyzing only the municipalities where the party runs at $t + 1$ yields a negative effect of -0.25, since the party wins 25% (25/100) of races in the treatment group and 50% (25/50) in the control group. However, an unconditional analysis that compares whether the party wins regardless of whether it ran, yields an effect of 0, as the proportion of electoral victories is 25/100 in both groups.

The results for unconditional $t + 1$ victory reported in the main body of the paper show that a scenario such as the one just described is not driving our results. Table 2 in the main paper shows that, with one exception, the negative effects reported in Table S10 are somewhat reduced but are still large, negative and strongly statistically significant in the unconditional analysis.

S8 Bound Analysis for Conditional Victory $t + 1$

An alternative strategy to deal with uncontested races is to treat all municipalities where the party does not contest the $t + 1$ election as missing data, and calculate bounds for the effect of interest under the different values that the missing data could have taken. We note that the most plausible assumption is that parties avoid running in municipalities where they expect to do poorly at $t + 1$, a situation that in a single sample would tend to overestimate the advantages to incumbency.

However, our situation is complicated by the fact that we are comparing two samples, municipalities where the party barely won (treated group) versus municipalities where the party barely lost (control group), so we need to consider the impact of endogenously different decision to run in each group. Our task is much simplified by the fact that we find negative results. The negative effects of incumbency on conditional victory at $t + 1$ reported in Table S10 could arise if the parties' selectively avoided all municipalities in the control group where they expected a poor electoral performance at $t + 1$ but did not avoid such municipalities in the treatment group, a situation we cannot entirely rule out.

We now use bounds (see, e.g. Manski 2007) to show that, under plausible scenarios regarding the missing data, we can rule out non-negative effects on conditional $t + 1$ victory for the incumbent party analysis.

Let $Y_{it+1,k}$ denote the outcome of interest for party k in municipality i at election $t + 1$, and $I_{it,k}$ be equal to one if party k wins the t election in municipality i and zero otherwise. Let $Z_{it+1,k}$ be equal to one if party k contests the $t + 1$ election and zero otherwise.

$$\begin{aligned} \lim_{m \downarrow 0} E(Y_{it+1,k} | M_{it,k} = m) = \\ \lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot \lim_{m \uparrow 0} \Pr(Z_{it+1,k} = 1 | M_{it,k} = m) + \\ \lim_{m \downarrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 0) \cdot \lim_{m \downarrow 0} \Pr(Z_{it+1,k} = 0 | M_{it,k} = m) \end{aligned}$$

All the terms can be estimated except for $E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 0)$. Assuming the support of Y is bounded by γ_l and γ_u , the identification region for $\lim_{m \downarrow 0} E(Y_{it+1,k} | M_{it,k} = m)$ is:

$$\begin{aligned} H^{Tr} = \left[\lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot \lim_{m \uparrow 0} \Pr(Z_{it+1,k} = 1 | M_{it,k} = m) + \right. \\ \left. \gamma_l \cdot \lim_{m \downarrow 0} \Pr(Z_{it+1,k} = 0 | M_{it,k} = m), \right. \\ \left. \lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot \lim_{m \uparrow 0} \Pr(Z_{it+1,k} = 1 | M_{it,k} = m) + \right. \\ \left. \gamma_u \cdot \lim_{m \downarrow 0} \Pr(Z_{it+1,k} = 0 | M_{it,k} = m) \right] \end{aligned}$$

Defining $p^{Tr} = \lim_{m \uparrow 0} \Pr(Z_{it+1,k} = 1 | M_{it,k} = m)$ this simplifies to:

$$H^{Tr} = \left[\lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot p^{Tr} + \gamma_l \cdot (1 - p^{Tr}), \right. \\ \left. \lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot p^{Tr} + \gamma_u \cdot (1 - p^{Tr}) \right]$$

And H^{Co} can be defined analogously for the control group, with $p^{Co} = \lim_{m \downarrow 0} \Pr(Z_{it+1,k} = 1 | M_{it,k} = m)$.

Letting $H^{Tr} = [L^{Tr}, U^{Tr}]$ and $H^{Co} = [L^{Co}, U^{Co}]$, the identification region for the average treatment effect at the cutoff is:

$$H = [L^{Tr} - U^{Co}, U^{Tr} - L^{Co}]$$

In our case, Y is a binary outcome that indicates whether party k won election $t + 1$, so $\gamma_l = 0$ and $\gamma_u = 1$, which leads to the extreme value identification regions:

$$\check{H}^{Tr} = \left[\lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot p^{Tr}, \right. \\ \left. \lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot p^{Tr} + (1 - p^{Tr}) \right]$$

and:

$$\check{H}^{Co} = \left[\lim_{m \downarrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot p^{Co}, \right. \\ \left. \lim_{m \downarrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot p^{Co} + (1 - p^{Co}) \right]$$

As discussed above, however, assuming that parties would have won every election that they did not contest is implausible and, in addition, not even possible since at most one party can win each election and there are many parties competing in every election.

Thus, to construct the bounds we still set $\gamma_l = 0$, but use the alternative assumption that in those municipalities where the party did not contest the election, its electoral performance would have been no better than it was in those municipalities where it did run, $\gamma_u = \lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1)$, which leads to the alternative identification regions:

$$\tilde{H}^{Tr} = \left[\lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot p^{Tr}, \lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \right] \equiv [\tilde{L}^{Tr}, \tilde{U}^{Tr}] \\ \tilde{H}^{Co} = \left[\lim_{m \downarrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \cdot p^{Co}, \lim_{m \downarrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1) \right] \equiv [\tilde{L}^{Co}, \tilde{U}^{Co}]$$

So the bounds we estimate and report are:

$$\tilde{H} = [\tilde{L}^{Tr} - \tilde{U}^{Co}, \tilde{U}^{Tr} - \tilde{L}^{Co}]$$

Table S11: Bounds for Victory t+1 – Full Sample

Variable	NTr	NCo	Prz_Tr	Prz_Co	mu1	mu0	Ident Region	CI Ident Region
Inc	1664	1470	0.57	0.63	0.30	0.51	[-0.339,-0.021]	[-0.422,0.069]
PSDB	1571	1242	0.64	0.57	0.44	0.55	[-0.273,0.125]	[-0.358,0.209]
PMDB	2088	1911	0.74	0.70	0.39	0.58	[-0.293,-0.015]	[-0.373,0.067]
DEM	1098	960	0.55	0.57	0.45	0.60	[-0.351,0.109]	[-0.461,0.216]
PT	735	718	0.78	0.75	0.49	0.63	[-0.249,0.019]	[-0.386,0.16]
PP	845	690	0.60	0.59	0.37	0.65	[-0.428,-0.013]	[-0.548,0.117]

We estimate \tilde{L}^{Tr} , \tilde{U}^{Co} , \tilde{U}^{Tr} and \tilde{L}^{Co} from our data, and calculate confidence intervals using bootstrapping. Since the effect we estimate is negative, we are interested in whether $\tilde{U}^{Tr} - \tilde{L}^{Co}$ is less than zero. If $\tilde{U}^{Tr} - \tilde{L}^{Co} < 0$, the conclusions of our analysis remain unchanged. In contrast, if $\tilde{U}^{Tr} - \tilde{L}^{Co} \geq 0$, a nonnegative effect of incumbency cannot be ruled out with our assumption that $\gamma_u = \lim_{m \uparrow 0} E(Y_{it+1,k} | M_{it,k} = m, Z_{it+1,k} = 1)$ (note that $\gamma_l = 0$ is not an assumption).

S9 Positive Incumbency Advantage Cannot Explain Negative Result

Imagine there is a positive personal incumbency advantage equal to γ which is the same for incumbent candidates of all parties and all time periods, and also that each party i receives baseline vote B_i in open seat races that is constant over time. When there is an incumbent candidate running for Party i , the vote that Party i receives in open seats at election $t + 1$ is $V_{it+1} = B_i$ and the vote that it receives when Party i is running is $V_{it+1} = B_i + \gamma$.

Now define the RD estimand as $V_{it+1}^{Tr} - V_{it+1}^{Co}$, where Tr and Co indicate, respectively, the municipalities where Party i barely won and lost the previous election (election t). Consider the table below, and note that a positive advantage can never result in a negative RD effect in the Incumbent Sample. When Party i loses, some other opposition party wins; we call this opposition party Party j . We assume this party is an opposition party (i.e., the votes that j gets, it takes away from i).

Table S12: Possible scenarios and sign of RD effect at $t + 1$ under positive personal incumbency advantage

	Treatment Group	Control Group	RD effect ($V_{it+1}^{Tr} - V_{it+1}^{Co}$)
(1)	Incumbent candidate runs	Incumbent candidate runs	$(B_i + \gamma) - (B_i - \gamma) = 2\gamma > 0$
(2)	Incumbent candidate runs	Open Seat	$(B_i + \gamma) - B_i = \gamma > 0$
(3)	Open Seat	Incumbent candidate runs	$B_i - (B_i - \gamma) = \gamma > 0$
(4)	Open Seat	Open Seat	$B_i - B_i = 0$

Note: Treatment group defined as municipalities where Party i won election t and control group as municipalities where Party i lost to some opposition party j at election t .

In Incumbent Sample, we only have scenarios (3) and (4), since there are no incumbents running at $t + 1$ in the treatment group. We can see that the effect is either 0 or γ , never negative. In the Open Seat Sample, we have all four scenarios, so the effect can be 0, γ or 2γ . On average, the RD effect will be larger in the Open Seat sample, due to the inclusion of scenarios (1) and (2) (specially (1)). Thus, a positive personal incumbency advantage may explain why the effect in the Incumbent Seat Sample is *smaller* than the effect in the Open Seat sample, but it cannot explain why it is negative.

Thus, if the loss of the mayor’s personalistic support were the true cause of the difference between the samples that we report in the main paper, the negative effects we see would be caused by an unknown factor that affects both samples and causes an effect of equal absolute value to the effect observed in the Incumbent sample, but with a positive sign—this is how large the negative effect would have to be to turn a positive personal incumbency advantage into the negative results of the magnitude we observe. In the incumbent party analysis, this unknown factor would need to account for a negative effect of 21 percentage points, equally affecting both samples. The fact that we have no theory for what this unknown factor may be, together with the fact that below we corroborate the second empirical implication of our model, suggests that this alternative explanation is implausible.

S10 Additional Career Path Analysis

S10.1 Overall

Table S13: Career Path of Mayors Elected in 1996 (Full Sample)

1998	Yes	No
Runs	26	5350
	Yes	No
Wins	2	24
Runs with same party	9	17
Runs and wins with same party	2	
2000	Yes	No
Runs	3646	1730
	Yes	No
Wins	2097	1549
Runs with same party	2575	1071
Runs and wins with same party	1504	
2002	Yes	No
Runs	275	5101
	Yes	No
Wins	43	232
Runs with same party	116	159
Runs and wins with same party	19	
2004	Yes	No
Runs	1282	4094
	Yes	No
Wins	377	905
Runs with same party	630	652
Runs and wins with same party	198	
2006	Yes	No
Runs	390	4986
	Yes	No
Wins	123	267
Runs with same party	151	239
Runs and wins with same party	63	
2008	Yes	No
Runs	2105	3271
	Yes	No
Wins	605	1500
Runs with same party	912	1193
Runs and wins with same party	292	
2010	Yes	No
Runs	221	5155
	Yes	No
Wins	85	136
Runs with same party	85	136
Runs and wins with same party	40	
2012	Yes	No
Runs	1079	4297
	Yes	No
Wins	459	620
Runs with same party	26	399
Runs and wins with same party	188	680

Note: All cells report counts, i.e. the number of mayors in each category. Results for all mayors elected in

Table S14: Career Path of Mayors Elected in 2004

	Full Sample		Elected to 2nd term		Elected to 1st term	
2006	Yes	No	Yes	No	Yes	No
Runs	36	5484	23	1332	13	4152
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	8	28	7	16	1	12
Runs with same party	15	21	9	14	6	7
Runs and wins with same party	5		4		1	
2008	Yes	No	Yes	No	Yes	No
Runs	3251	2269	26	1329	3225	940
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	2141	1110	13	13	2172	1053
Runs with same party	2253	998	13	13	2240	985
Runs and wins with same party	1510		8		1502	
2010	Yes	No	Yes	No	Yes	No
Runs	334	5186	188	1167	146	4019
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	79	255	49	139	30	116
Runs with same party	163	171	95	93	68	78
Runs and wins with same party	46		26		20	
2012	Yes	No	Yes	No	Yes	No
Runs	1244	4276	621	734	623	3542
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	481	763	241	380	240	383
Runs with same party	683	561	359	262	324	299
Runs and wins with same party	274		145		129	

Note: All cells report counts, i.e. the number of mayors in every category. First two columns (labeled *Full Sample*) report results for all mayors who were elected in 2004, while the sets of columns labeled *Elected to 2nd term* and *Elected to 1st term* subset these results by reelection status. Columns labeled *Elected to 2nd term* report results for the subset of mayors elected in 2004 who in 2004 were reelected to their second consecutive term, while columns labeled *Elected to 1st term* report results for the subset of mayors who was elected in 2004 for their first consecutive term.

Table S15: Career Path of Mayors Elected in 2008

	Full Sample		Elected to 2nd term		Elected to 1st term	
2010	Yes	No	Yes	No	Yes	No
Runs	55		28		27	
	Yes	No	Yes	No	Yes	No
Wins	12	43	10	18	2	25
Runs with same party	28	27	20	8	8	19
Runs and wins with same party	12		10		2	
2012	Yes	No	Yes	No	Yes	No
Runs	2523		22		2501	
	Yes	No	Yes	No	Yes	No
Wins	1393	1130	11	11	1382	1119
Runs with same party	2034	489	15	7	2019	482
Runs and wins with same party	1132		7		1125	

Note: All cells report counts, i.e. the number of mayors in every category. First two columns (labeled *Full Sample*) report results for all mayors who were elected in 2008, while the sets of columns labeled *Elected to 2nd term* and *Elected to 1st term* subset these results by reelection status. Columns labeled *Elected to 2nd term* report results for the subset of mayors elected in 2008 who in 2008 were reelected to their second consecutive term, while columns labeled *Elected to 1st term* report results for the subset of mayors who was elected in 2008 for their first consecutive term.

S10.2 PT Versus Other Parties

Table S16: Career Path of Mayors Reelected in 2004 to Second Consecutive Term: PT vs Other Parties

	All		PT		Other Parties	
	Yes	No	Yes	No	Yes	No
2006						
Runs	23	1332	2	70	21	1262
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	7	16	1	1	6	15
Runs with same party	9	14	1	1	8	13
Runs and wins with same party	4		1		3	
2008						
Runs	26	1329	1	71	25	1258
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	13	13	1	0	13	12
Runs with same party	13	13	0	1	13	12
Runs and wins with same party	8		0		8	
2010						
Runs	188	1167	33	39	155	1128
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	49	139	12	21	37	118
Runs with same party	95	93	25	8	70	85
Runs and wins with same party	26		11		15	
2012						
Runs	621	734	33	39	588	695
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	241	380	8	25	233	355
Runs with same party	359	262	28	5	331	257
Runs and wins with same party	145		7		138	

Note: All cells report counts, i.e. the number of mayors in every category. First two columns (labeled *All*) report results for all mayors who were reelected in 2004 for their second consecutive term, while the sets of columns labeled *PT* and *Other Parties* subset these results by type of party. Columns labeled *PT* report results for PT mayors who in 2004 were reelected to their second consecutive term, while columns labeled *Other Parties* report results for mayors from all other parties who were reelected in 2004 for their second consecutive term.

Table S17: Career Path of Mayors Reelected in 2008 to Second Consecutive Term: PT vs Other Parties

	All		PT		Other Parties	
2010	Yes	No	Yes	No	Yes	No
Runs	28	2098	7	196	21	1902
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	10	18	4	3	6	15
Runs with same party	20	8	6	1	14	7
Runs and wins with same party	10		4		6	
2012	Yes	No	Yes	No	Yes	No
Runs	22	2104	1	202	21	1902
	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>No</u>
Wins	11	11	1	0	10	11
Runs with same party	15	7	1	0	14	7
Runs and wins with same party	7		1		6	

Note: All cells report counts, i.e. the number of mayors in every category. First two columns (labeled *All*) report results for all mayors who were reelected in 2008 for their second consecutive term, while the sets of columns labeled *PT* and *Other Parties* subset these results by type of party. Columns labeled *PT* report results for PT mayors who in 2008 were reelected to their second consecutive term, while columns labeled *Other Parties* report results for mayors from all other parties who were reelected in 2008 for their second consecutive term.

S11 Additional Analysis for Incumbent vs. Open Seat Samples

Table S18: RD effect on Candidacy at $t+1$ —Open Seat vs. Incumbent sample

Party	Subsample	Estimate	95% CI	pval	h	Ntr	Nco
INCUMBENT	Incumbent	-0.11	[-0.193,-0.046]	0.00	19.35	2280	1815
	Open Seat	0.03	[-0.07,0.104]	0.70	18.15	1044	1159
PMDB	Incumbent	-0.07	[-0.196,0.076]	0.39	17.12	700	556
	Open Seat	0.06	[0.001,0.111]	0.05	17.63	2505	2973
PSDB	Incumbent	-0.10	[-0.24,0.039]	0.16	18.99	598	471
	Open Seat	0.11	[0.033,0.175]	0.00	20.03	1831	2020
DEM	Incumbent	-0.05	[-0.201,0.134]	0.69	16.22	418	394
	Open Seat	-0.01	[-0.095,0.056]	0.61	14.82	1525	1764
PP	Incumbent	-0.11	[-0.315,0.125]	0.40	12.22	282	202
	Open Seat	0.04	[-0.05,0.135]	0.37	15.64	1195	1309
PT	Incumbent	-0.12	[-0.407,0.153]	0.37	13.00	149	96
	Open Seat	0.07	[-0.021,0.141]	0.15	21.18	736	1113

Note: Running variable is party's vote margin at t , outcome is dummy =1 if party contests the following election at $t + 1$, =0 otherwise. Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 4-8 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth.

Table S19: RD effect on Victory at t+1—Open Seat vs. Incumbent sample

Party	Subsample	Estimate	95% CI	pval	h	Ntr	Nco	Difference
INCUMBENT	Incumbent	-0.33	[-0.437,-0.233]	0.00	14.72	959	824	-0.255
	Open Seat	-0.07	[-0.211,0.047]	0.21	15.68	697	634	[-0.418, -0.089]
PMDB	Incumbent	-0.33	[-0.515,-0.149]	0.00	16.12	392	301	-0.158
	Open Seat	-0.17	[-0.253,-0.107]	0.00	16.40	1831	1741	[-0.349, 0.045]
PSDB	Incumbent	-0.27	[-0.439,-0.096]	0.00	19.31	300	247	-0.181
	Open Seat	-0.09	[-0.189,0.001]	0.05	20.04	1290	1014	[-0.369, 0.0230]
DEM	Incumbent	-0.21	[-0.454,-0.007]	0.04	19.05	205	195	-0.062
	Open Seat	-0.15	[-0.271,-0.053]	0.00	14.49	931	788	[-0.317,0.181]
PP	Incumbent	-0.56	[-0.842,-0.331]	0.00	13.13	135	105	-0.342
	Open Seat	-0.22	[-0.352,-0.112]	0.00	14.98	759	605	[-0.637,-0.073]
PT	Incumbent	-0.07	[-0.407,0.295]	0.75	15.85	127	69	0.082
	Open Seat	-0.16	[-0.293,-0.029]	0.02	17.84	580	614	[-0.270, 0.480]

Note: Running variable is party's vote margin at t , outcome is dummy =1 if party contests the following election at $t + 1$, =0 if it runs and loses. Sample includes only municipalities where party contests the $t + 1$ election. Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 4-8 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth. Last column reports difference in point estimates between Incumbent and Open Seat sample and corresponding 95% confidence interval.

S12 Public Good Provision Indicators for PT

Table S20: RD Effects on Public Good Provision Indicators for PT

PUBLIC FINANCE DATASET								
Outcome	Subsample	Estimate	CI	pval	h	Ntr	Nco	Difference
Share HESA Expend.	All seats	-0.48	[-2.47,1.14]	0.47	16.39	814	972	-1.77 [-7.49,3.86]
	Incumbent	-1.90	[-7.44,3.24]	0.44	14.09	160	98	
	Open Seat	-0.14	[-2.21,1.64]	0.77	17.25	657	909	
MUNICIPAL SURVEY DATASET								
Outcome	Subsample	Estimate	95% CI	pval	h	Ntr	Nco	Difference
Adm. Employment Growth	All seats	-0.13	[-8.08,6.76]	0.86	19.64	678	789	-9.29 [-33.38,9.36]
	Incumbent	-8.17	[-30.02,8.55]	0.28	13.89	133	80	
	Open Seat	1.12	[-7.92,10.47]	0.79	16.30	471	588	
Social Assistance Expend.	All seats	-9271.50	[-24107.04,2336.28]	0.11	13.06	361	382	-37510.48 [-101740.66,23361.86]
	Incumbent	-40734.09	[-105253.35,18378.77]	0.17	11.50	65	39	
	Open Seat	-3223.61	[-13802.49,5306.72]	0.38	13.52	287	353	
Housing Program	All seats	0.09	[0.02,0.17]	0.01	23.33	883	1138	-0.08 [-0.31,0.15]
	Incumbent	0.02	[-0.18,0.25]	0.76	17.12	178	113	
	Open Seat	0.11	[0.03,0.2]	0.01	24.41	674	1060	
Housing Materials Program	All seats	-0.01	[-0.14,0.09]	0.68	15.42	703	768	-0.32 [-0.72,0.03]
	Incumbent	-0.26	[-0.65,0.05]	0.09	10.37	112	84	
	Open Seat	0.05	[-0.09,0.17]	0.53	16.08	547	689	

Note: Running variable is party's vote margin at t . Estimate is average treatment effect at cutoff estimated with local linear regression with triangular kernel and MSE-optimal bandwidth chosen according to CCT implementation. Columns 4-8 report, respectively, 95% robust confidence intervals, robust p-value, main optimal bandwidth, treated observations within bandwidth, and control observations within the bandwidth. Last column reports difference in point estimates between Incumbent and Open Seat sample and corresponding 95% confidence interval.

S13 Mayoral Victories: Incumbent versus Non-incumbent Candidates

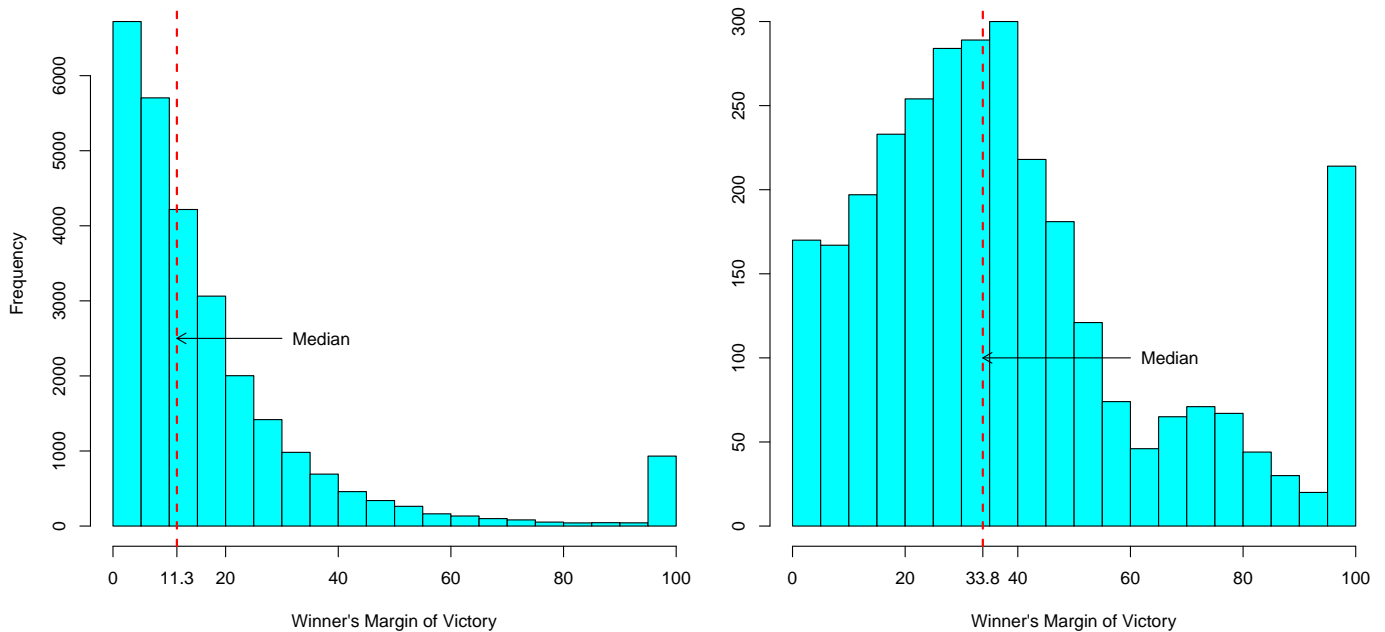
Table S21: Mayoral Victories of Incumbent Party by Type of Seat, 2000-2012

	Non-incumbent candidate runs	Incumbent Mayor Runs	Total
Incumbent Party Loses	2,492	3,201	5,693
Incumbent Party Wins	2,042	4,954	6,996
Total	4,534	8,155	12,689

S14 Histogram of Close Races in Brazil and United States

Figure S4a shows the winner party's margin of victory in all mayoral elections in Brazil between 1996 and 2012, and shows that the majority of mayoral elections are highly competitive. In this period, the median margin of victory by the winning party is just 11.3 percentage points, and roughly a quarter of the races are decided by five percentage points or less. This is in stark contrast to noncompetitive settings such as the U.S. House elections, shown in Figure S4(b) for a similar time period, where the median margin of victory by the winning party is 33.8 percentage points, and where only about five percent of races are decided by less than five percentage points.

Figure S4: Histogram of Margin of Victory for Winner Party: Brazil vs. United States



(a) Brazilian Municipalities, 1996-2012

(b) U.S. House Districts, 1996-2010

S15 Exploring the Correlation Between Negative Effects of Incumbency and Party Weakness Across Countries

We developed a proxy measure for party weakness based on the data from the Democratic Accountability and Linkages Project (DALP, Kitschelt 2013) in order to conduct a tentative cross-country analysis of the correlation between negative effects of incumbency and party weakness. Our measure incorporates the extent to which parties maintain continued presence at the municipal level, resort to clientelistic strategies, and are programmatically structured.⁴ We performed individual party RD analyses analogous to those reported in rows 2-4 of Table 2 for Brazil, studying the effect of barely winning office at t on the probability of unconditional victory at $t+1$ for 21 different political parties from the six countries analyzed—Brazil, Chile, Colombia, Costa Rica, Mexico and Peru. Then, for each party, we created a “negative effect” indicator variable equal to one if the RD effect was statistically different from zero (at 5%) and negative, and zero otherwise.

In order to explore the correlation between negative party effects, term limits and party weakness, we regressed the negative effect indicator on the party weakness measure and an indicator equal to one if term limits restrict the reelection of incumbent mayors. We stress that this analysis is highly tentative: the sample size is limited, the party weakness measure we use does not perfectly capture the nature of the ties between individual candidates and party organizations central to our framework, and we are simply looking at associations. Thus, we see this analysis as suggestive and exploratory, not as definitive. The results are reported in Table S22. First, consistent with our findings in Figure 3 in the main paper, the first column (second row) shows that there is a positive correlation between the presence of term limits and a significant negative effect of incumbency. Second, there is tentative evidence that party weakness, as measured by our proxy, is also positively associated with an incumbent party suffering subsequent electoral losses (the first row of column 1). Moreover, in line with our theoretical framework, this positive correlation between party weakness and costs of incumbency is present even when countries with term limits are analyzed separately (column 2), suggesting an interaction between party weakness and term limits.⁵

In sum, our preliminary analysis from five additional Latin American countries shows that the electoral losses associated with winning access to office are not unique to Brazil, and tentatively suggests that these effects are related to the existence of term limits and the extent of weakness of political parties. This preliminary analysis together with the generality of our theoretical framework suggests that our framework could be applied more widely to study the evolution of party systems in developing democracies.

⁴In particular, each party’s score is the average of the local presence (variables a1-a3), clientelism (b1-b3), and the inverse of the CoSalPo score of programmatic structuration. See DALP codebook for details (Kitschelt 2014).

⁵The results are substantively unchanged if we fit a non-linear probability model such as probit rather than an OLS model.

Table S22: Party-Level Negative Effects, Term Limits, and Party Weakness

	All Countries	Countries With Term Limits
Party weakness score	1.24*** (0.41)	1.24** (0.50)
Term limits	0.37* (0.19)	
Parties	21	12

Note: Countries included are Brazil, Chile, Colombia, Costa Rica, Mexico and Peru.

References

- Ansola-behere, Stephen, and James M Snyder Jr. 2002. "The incumbency advantage in US elections: An analysis of state and federal offices, 1942-2000." *Election Law Journal* 1 (3): 315–338.
- Besley, Timothy. 2007. *Principled Agents? The Political Economy of Good Government*. Oxford, UK: Oxford University Press.
- Cattaneo, Matias D., Brigham Frandsen, and Rocío Titiunik. 2015. "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate." *Journal of Causal Inference* 3 (1): 1–24.
- Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2015a. "Simple Local Regression Distribution Estimators with an Application to Manipulation Testing." working paper, University of Michigan.
- Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2015b. "rddensity: Manipulation Testing in Stata." working paper, University of Michigan.
- Erikson, Robert, and Rocío Titiunik. 2015. "Using Regression Discontinuity to Uncover the Personal Incumbency Advantage." *Quarterly Journal of Political Science* 10 (1): 101–119.
- Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69 (1): 201–209.
- Kitschelt, Herbert. 2013. *Democratic Accountability and Linkages Project*. Duke University. <http://sites.duke.edu/democracylinkage/>: Durham, NC.
- Kitschelt, Herbert. 2014. *Democratic Accountability and Linkages Project. 2008-9 Dataset. Dataset Codebook*. Duke University. http://sites.duke.edu/democracylinkage/files/2014/12/DALP_Codebook_2014-04_01.pdf: Durham, NC.
- Manski, Charles F. 2007. *Identification for prediction and decision*. Cambridge, MA: Harvard University Press.
- Persson, Torsten, and Guido Enrico Tabellini. 2002. *Political Economics: Explaining Economic Policy*. MIT press.