Online Appendix: Parties, Brokers and Voter Mobilization: How Turnout Buying Depends Upon the Party's Capacity to Monitor Brokers

Horacio Larreguy* John Marshall† Pablo Querubín‡

1 Proofs

Proof of Proposition 1: The argument in the text derives voter behavior $v^*(d, e_A^*, e_B^*; \sigma)$, which pins down e_i^* for parties using backward induction. The individual rationality (IR) constraint in the program of equation (4) of the main text is clearly satisfied if the incentive compatibility (IC) constraint is satisfied. Note that the IC constraint comes from parties optimally choosing to punish, by withdrawing the wage, when they observe effort $e_i < \hat{e}_i$. The parties will clearly then let the IC bind at any interior optimum. Substituting for $\hat{w}_i$ using the binding IC constraint yields the following first-order condition in $\hat{e}_i$:

$$\psi\alpha d[1+\alpha d(1-\hat{e}_A)b(d)] \leq \frac{\gamma_A d\hat{e}_A}{p}, \quad (1)$$

$$\psi\alpha d[1-\alpha d(1-\hat{e}_B)b(d)] \leq \frac{\gamma_B d\hat{e}_B}{p}, \quad (2)$$

for parties A and B respectively. The second-order conditions for a unique equilibrium are satisfied when $\gamma_A + p\psi\alpha^2 db(d) > 0$ for party A and $\gamma_B - p\psi\alpha^2 db(d) > 0$ for party B. These necessary conditions are given at the beginning of the proposition.

When $1+\alpha db(d) \leq 0$, A's first order condition is weakly negative for any value of $\hat{e}_A$, and thus $\hat{e}_A^* = 0$. Similarly, when $1-\alpha db(d) \leq 0$, B's first order condition is weakly negative for any value of $\hat{e}_B$, and thus $\hat{e}_B^* = 0$.

*Assistant Professor, Department of Government, Harvard University, 1737 Cambridge St., Cambridge, MA 02138 (hlarreguy@fas.harvard.edu).

†Ph.D. candidate, Department of Government, Harvard University, 1737 Cambridge St., Cambridge, MA 02138 (jlmarsh@fas.harvard.edu).

‡Assistant Professor, Department of Politics, New York University, 19 W 4th Street, New York, NY 10012 (pablo.querubin@nyu.edu).

However, when $1+\alpha db(d)>0$ and $1-\alpha db(d)>0$ respectively hold, parties A and B respectively engage in turnout buying. Solving solving first order conditions respectively yields:

$$\hat{e}_A^* = \frac{p\psi\alpha[1+\alpha db(d)]}{\gamma_A + p\psi\alpha^2 db(d)}, \tag{3}$$

$$\hat{e}_B^* = \frac{p\psi\alpha[1-\alpha db(d)]}{\gamma_B - p\psi\alpha^2 db(d)}, \tag{4}$$

where it follows that $p\psi\alpha<\gamma_A$ and $p\psi\alpha<\gamma_B$ must hold for an interior solution where $\hat{e}_A^*<1$ and $\hat{e}_B^*<1$, respectively. Otherwise, $\hat{e}_A^*=\hat{e}_B^*=1$.

The binding IC constraint then determines the optimal wage $\hat{w}_i^*$. By virtue of satisfying the IC constraint, the broker optimally chooses $e_i^*=\hat{e}_i^*$. ■

Proof of Proposition 2: Given the closed form solutions in Proposition 1, the comparative statics are straight-forward to identify for any interior solution (as defined by the conditions in Proposition 1).

The first result follows from differentiating e_i^* with respect to p to yield:

$$\frac{\partial e_A^*}{\partial p} = \frac{\psi\alpha\gamma_A[1+\alpha db(d)]}{[\gamma_A + p\psi\alpha^2 db(d)]^2} \geq 0, \tag{5}$$

$$\frac{\partial e_B^*}{\partial p} = \frac{\psi\alpha\gamma_B[1-\alpha db(d)]}{[\gamma_B - p\psi\alpha^2 db(d)]^2} \geq 0. \tag{6}$$

Both differentials are positive at any interior solution because $e_i^*\in(0,1)$ and thus the numerators of the solutions for each e_i^* must also be positive; the denominators are positive in any equilibrium.[1]

We now turn to the second result. We first differentiate $\frac{\partial e_A^*}{\partial p}$ by d to yield:

$$\frac{\partial^2 e_A^*}{\partial p \partial d} = \frac{\psi\gamma_A\alpha^2[b(d)+db'(d)]\Big[\gamma_A - p\psi\alpha[2+\alpha db(d)]\Big]}{[\gamma_A+\psi\alpha^2 pdb(d)]^3}. \tag{7}$$

From our existence condition (given in Proposition 1), the denominator is positive. $\psi\gamma_A\alpha^2[b(d)+db'(d)]$ is always positive, given the assumption that $b(d)+db'(d)>0$ (i.e. $\varepsilon(d)>-1$). The cross-partial thus depends on the term in large brackets in the numerator. As $d\downarrow 0$, the condition $\gamma_A>2p\psi\alpha$ ensures that this term is positive, and thus equation (7) is positive for $d\downarrow 0$. Since the term in large brackets term is monotonically decreasing in d, given that $b'(d)>0$, and is negative as $d\to\infty$, there must then exist a cut-point $\overline{d}_A>0$ such that $\gamma_A - p\psi\alpha[2+\alpha\overline{d}_A b(\overline{d}_A)]=0$ and equation (7) is negative for $d>\overline{d}_A$.

The result similarly follows for party B from

$$\frac{\partial^2 e_B^*}{\partial p \partial d} = -\frac{\psi\gamma_B\alpha^2[b(d)+db'(d)]\Big[\gamma_B - p\psi\alpha[2-\alpha db(d)]\Big]}{[\gamma_B-\psi\alpha^2 pdb(d)]^3} < 0. \tag{8}$$

[1]Note that since Π^i is increasing in e_i^* (or $\hat{e}_i^*$), these results for e_i^* (or $\hat{e}_i^*$) equally apply to Π^i.

As above, our existence conditions ensures that the denominator is positive. Given $\varepsilon(d) > -1$, the cross-partial thus again depends on the term in large brackets in the numerator. As $d \downarrow 0$, the condition $\gamma_B > 2p\psi\alpha$ ensures that this term is positive, and thus $\frac{\partial^2 e_B^*}{\partial p \partial d} < 0$. Since the term in large brackets is monotonically increasing in d, $\frac{\partial^2 e_B^*}{\partial p \partial d} < 0$ always holds.[2] ■

2 Effects of congestion on party vote share

As noted in the main text, an important concern is that our results are instead driven by a reduction in congestion costs. To identify the implications of this concern, we re-examine the implications of the model. In particular, we assume that no turnout buying occurs ($e_i = 0$) and voters instead face the cost $c(d, f) = \alpha d + f$, where $f > 0$ is the cost associated with congestion. To identify the implications of congestion, we simply differentiate the party vote share (as a proportion of registered voters)—Π^A and Π^B—by f, before examining the cross-partial effect with distance d.

Unsurprisingly, the vote share for each party decreases in f:

$$\frac{\partial \Pi^A}{\partial f} = -\psi[1 + b(d)(\alpha d + f)] < 0, \quad (9)$$

$$\frac{\partial \Pi^B}{\partial f} = -\psi[1 - b(d)(\alpha d + f)] < 0. \quad (10)$$

The signs follow from $b(d) \in [-\psi, \psi]$ and $c(d, f) \in [0, \frac{1}{\psi}]$ (in the main text). A reduction in congestion thus increases the vote share of both parties, and particularly the party experiencing the bias (e.g. $b(d) > 0$ for a bias toward A and $b(d) < 0$ for a bias toward B).

Differentiating again to examine the heterogeneous effects of congestion by distance yields:

$$\frac{\partial^2 \Pi^A}{\partial f \partial d} = -\psi[\alpha[b(d) + db'(d)] + b'(d)f] < 0, \quad (11)$$

$$\frac{\partial^2 \Pi^B}{\partial f \partial d} = \psi[\alpha[b(d) + db'(d)] + b'(d)f] > 0, \quad (12)$$

where the signs follow from the condition $b(d) + db'(d) > 0$. These cross-partial effects demonstrate that the increased number of votes for the PRI—or rural party A—due to reduce congestion costs should be monotonically increasing in distance, while the increased number of votes for the PAN and PRD are expected to be monotonically decreasing in distance.

[2] Differentiating Π^i to examine when these results for e_i^* carry over to the vote shares of parties A and B: $\frac{\partial^2 \Pi^i}{\partial p \partial d} = \frac{\partial^2 e_i^*}{\partial p \partial d} \frac{\gamma_i d e_i^*}{p} + \frac{\partial e_i^*}{\partial p} \frac{\gamma_i e_i^*}{p}$, where we exploit the envelope condition that $\partial e_i^* / \partial d = 0$ and substitute using the first-order conditions. Given $\frac{\gamma_i e_i^*}{p} \geq 0$, the sign of $\frac{\partial^2 e_i^*}{\partial p \partial d}$ determines the first term of $\frac{\partial^2 \Pi^i}{\partial p \partial d}$. We call this the "first-order effect". Whenever this exceeds the second term, the first-order effect dominates, and thus follows the sign of $\frac{\partial^2 \Pi^i}{\partial p \partial d}$.

3 Variable definitions and summary statistics

Our variables are defined below. Summary statistics are provided in Table 1 for our main variables; summary statistics for our balancing variables are available in our replication code. Most of the data was obtained from the IFE using freedom of information requests. Codebooks defining our Census and other balancing variables are available upon request.

Turnout. Proportion of voters at a given polling station (within an electoral precinct) that turned out at the legislative election. This includes all votes, not just valid votes. Source: IFE.

PAN/PRD/PRI vote share. PAN/PRD/PRI legislative vote share, as a proportion of the registered electorate, at a given polling station. Source: IFE.

Registered voters in electoral precinct. Number of voters registered to vote in a given electoral precinct. Source: IFE.

Registered voters deviation. Difference in the number of registered voters in a given electoral precinct from the nearest multiple of 750.

Split. Indicator coded one if registered voters deviation is greater than zero. (as defined in the main text).

Registered voters in polling station. Number of voters registered to vote at a given polling station. Source: IFE.

Year. National legislative election year; 2000, 2003, 2006, 2009 or 2012.

Distance. We used the set of IFE localities, which differ from National Institute of Statistics and Geography (INEGI) localities because the INEGI groups individuals in bigger localities, and calculated the Euclidean distance (in kilometers) from each locality to the locality of the polling stations in the section. We assigned zero distance to the voters who voted in the locality where the polling station was located since we do not know the spatial distribution of the voters within the locality. Computing the voter-weighted distance of each locality from the polling stations in the section required three types of data that we obtained from the IFE through various freedom of information requests. These three types of data are: a) data on the number of registered voters in each IFE locality (available for 2006, 2009, and 2012), b) data on the coordinates of each IFE locality (available for 2006, 2009, and 2012), and c) the coordinates of each polling station (available for 2006, 2009, and 2012). While the three types of data provide an IFE locality code, these do not always matched across the different data sets. However, using a fuzzy name matching algorithm in Stata (reclink), together with extensive matching by hand, we were able to match the localities that represented at least 95% of the registered voters for each electoral precinct for 99.5% of the precincts. Finally, to calculate the average distance of voters to the polling station, we summed all distances weighting by the locality registered population divided by the total registered population in the electoral precinct. Due to the time required to complete the name matching procedure that had to be done by hand, we only executed the full matching procedure, and thus computed the average distance measure, for almost all the electoral precincts for the year 2012, which we use for all years. However, restricting attention to those electoral precincts where the fuzzy name matching algorithm matched the localities that represented at least 95% of the registered voters, we observe an extremely high correlation in the average distance measure across years.

PRD state governor. Indicator coded one if the state governor is from the PRD at the time of the election.

Area (log). The natural logarithm of the electoral precinct area in kilometers.

Voter density (log). The natural logarithm of the registered precinct electorate divided by total area in kilometers.

Share economically active. Percentage of electoral precinct population that is economically active. Source: 2010 Census.

Share employed. Percentage of electoral precinct population that is employed. Source: 2010 Census.

Share medical insurance. Percentage of electoral precinct population that has medical insurance. Source: 2010 Census.

Share illiterate. Percentage of electoral precinct population above 15 that is illiterate. Source: 2010 Census.

Incomplete primary school. Percentage of electoral precinct population above 15 with incomplete primary schooling. Source: 2010 Census.

Complete primary school. Percentage of electoral precinct population above 15 with complete primary schooling. Source: 2010 Census.

Incomplete secondary school. Percentage of electoral precinct population above 15 with incomplete secondary schooling. Source: 2010 Census.

Complete secondary school. Percentage of electoral precinct population above 15 with complete secondary schooling. Source: 2010 Census.

Share owns house. Percentage of electoral precinct owning a house. Source: 2010 Census.

Share basic amenities. Percentage of households in the electoral precinct with all electricity, piped water, toilet and drainage. Source: 2010 Census.

Share with radio. Percentage of households in the electoral precinct with radio. Source: 2010 Census.

Share with TV. Percentage of households in the electoral precinct with a television. Source: 2010 Census.

Share with fridge. Percentage of households in the electoral precinct with a refrigerator. Source: 2010 Census.

Share washing machine. Percentage of households in the electoral precinct with a washing machine. Source: 2010 Census.

Share with car. Percentage of households in the electoral precinct with a car or truck. Source: 2010 Census.

Share with telephone. Percentage of households in the electoral precinct with a landline telephone. Source: 2010 Census.

Share cell phone. Percentage of households in the electoral precinct with a cellphone. Source: 2010 Census.

Share with internet. Percentage of households in the electoral precinct with internet access. Source: 2010 Census.

Table 1: Summary statistics: discontinuity and full (national) samples

	Discontinuity sample (20 voter bandwidth)					**Full sample**				
	Obs.	Mean	Std. dev.	Min.	Max.	Obs.	Mean	Std. dev.	Min.	Max.
Turnout	27,697	0.54	0.14	0.06	1	517,255	0.53	0.14	0.00	1
PRI vote share	27,697	0.19	0.08	0	0.83	517,255	0.19	0.08	0	1.00
PAN vote share	27,697	0.17	0.11	0	0.86	517,255	0.17	0.11	0	0.86
PRD vote share	27,697	0.10	0.09	0	0.99	517,255	0.10	0.09	0	0.99
Registered voters in electoral precinct	27,697	1585.45	1638.26	731	24010	517,255	1618.22	1524.21	375	24010
Registered voters deviation	27,697	2.62	11.21	-19	20	517,255	1.75	215.70	-375	375
Split	27,697	0.61	0.49	0	1	517,255	0.53	0.50	0	1
Registered voters in polling station	27,697	570.56	154.08	375	750	517,255	575.04	100.58	375	750
Year	27,697	2006.44	4.21	2000	2012	517,255	2006.40	4.20	2000	2012
Distance	27,417	0.22	0.65	0	18.03	512,774	0.22	0.67	0	24.33

4 Continuity around the discontinuity

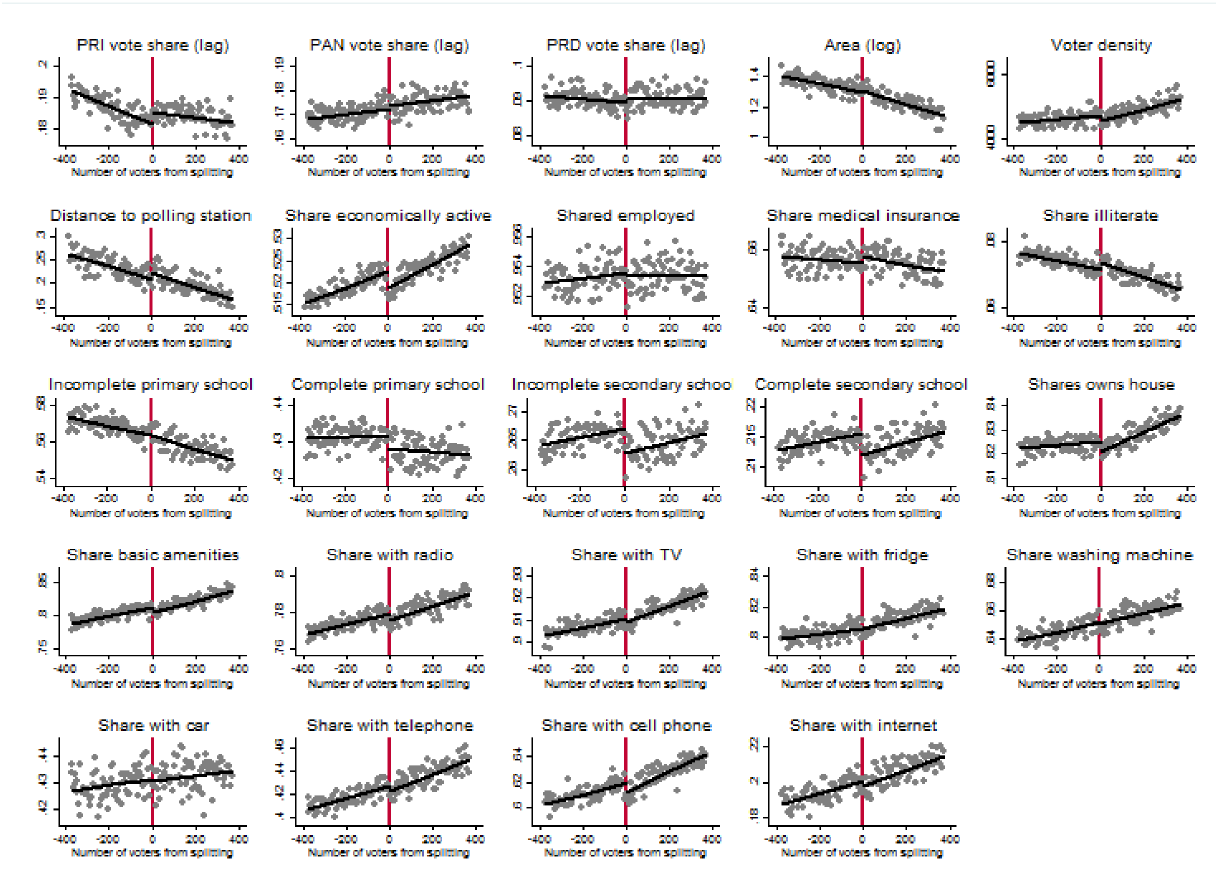

Figure 1: Scatter plots in the running variable for each balancing variable

5 Distribution of polling station splits

Our main analysis examines electoral precincts within 20 registered voters of receiving an additional polling station. As the Figure in the main text suggests, these precincts are evenly split across the country. Table 2 shows this claim more clearly, demonstrating that the proportion of precincts from each state in our discontinuity sample almost exactly reflects the proportion of all available electoral precincts. Table 3 shows that the distribution across election years is also similar.

6 Efficacy of turnout-buying

Table 4 instruments for turnout using our polling station discontinuity in order to estimate the effectiveness of turnout buying on a party's total vote share (as a percentage of registered voters).

Table 2: Distribution of polling stations across states: discontinuity and full (national) samples

State	Discontinuity sample (20 voter bandwidth), % of total			Full sample, % of total
	All	Split	Unsplit	
Aguascalientes	1.16	1.17	1.15	1.06
Baja California	3.13	3.15	3.09	2.97
Baja California Sur	0.52	0.51	0.53	0.55
Campeche	0.73	0.80	0.61	0.79
Chiapas	2.62	2.85	2.27	2.63
Chihuahua	0.49	0.52	0.44	0.61
Coahuila	4.00	3.71	4.46	3.58
Colima	4.30	4.63	3.79	3.99
Durango	1.94	2.09	1.71	1.89
Guanajuato	5.45	5.68	5.09	5.31
Guerrero	3.01	2.97	3.07	3.79
Hidalgo	2.50	2.34	2.76	2.65
Jalisco	7.16	7.05	7.35	6.96
Mexico	13.22	12.73	13.99	13.26
Michoacan	4.57	4.58	4.56	4.46
Morelos	1.83	1.86	1.78	1.76
Nayarit	1.20	1.08	1.38	1.17
Nuevo Leon	4.77	4.44	5.29	4.34
Oaxaca	3.63	3.68	3.56	3.71
Puebla	5.25	5.44	4.96	4.97
Queretaro	1.47	1.50	1.42	1.51
Quintana Roo	1.35	1.28	1.45	1.04
San Luis Potosi	2.30	2.27	2.35	2.66
Sinaloa	1.83	1.86	1.79	3.76
Sonora	2.66	2.70	2.59	2.55
Tabasco	2.01	1.94	2.13	2.07
Tamaulipas	3.43	3.68	3.05	3.27
Tlaxcala	0.98	0.89	1.11	1.11
Veracruz	8.61	8.67	8.53	7.62
Yucatan	1.92	1.74	2.20	1.88
Zacatecas	1.94	2.19	1.54	2.08

Table 3: Distribution of polling stations across election years: discontinuity and full (national) samples

Election year	Discontinuity sample (20 voter bandwidth), % of total			Full sample, % of total
	All	Split	Unsplit	
2000	17.06	17.30	16.67	17.59
2003	18.32	18.38	18.22	18.66
2006	20.58	20.13	21.29	20.23
2009	21.15	21.08	21.26	21.50
2012	22.90	23.11	22.56	22.01

The results, as cited in the main text, indicate that turnout buying is relatively effective: a percentage point increase in turnout translates into a 0.62 percentage point increase in votes for the PRI, and a 0.48 percentage point increase in votes for the PAN. As noted in the main text, this is likely to be an under-estimate if multiple parties simultaneously buy turnout in some precincts.

Table 4: IV estimates of the effect of turnout on vote share

	PRI vote share (1)	PAN vote share (2)	PRD vote share (3)
Turnout	0.5339***	0.4791***	0.0324
	(0.1496)	(0.1548)	(0.0980)
Observations	27,697	27,697	27,697
First stage F statistic	9.4	9.4	9.4
Observations	27,770	27,770	27,770

Notes: All specifications include district-year fixed effects, and are estimated with 2SLS. All results are for a 20 voter bandwidth. Standard errors are clustered by state. * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

7 Effect of an additional polling station in PRD strongholds

Using interactions, Table 5 shows that there is no evidence of PRD turnout buying in states with a PRD governor. Unreported robustness checks show that in Michoacán and Guerrero, states where the PRD inherited the PRI's local apparatus, the interaction coefficient is negative.

Table 5: Effect of an additional polling station in PRD strongholds

	PRD vote share
Split	-0.0002
	(0.0010)
Split × PRD governor	0.0032
	(0.0025)
Observations	27,697

Notes: All specifications include district-year fixed effects, and are estimated with OLS. All results are for a 20 voter bandwidth. Standard errors are clustered by state. * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

8 Robustness checks

Tables 6 and 8 report the additional robustness checks cited in the main text. Table 6 shows the local linear regression estimates where linear trends in the running variable are included either side of the discontinuity. Although the estimates are somewhat noisier, the effect sizes generally rise (although the larger coefficient for the PAN ceases to be statistically significant). This increase is not especially surprising given that the main text shows that the trend either side of the discontinuity is declining; although the slope around the discontinuity is relatively shallow, in general such trends imply that comparing means will underestimate the effect at the discontinuity.

Table 7 presents our results when controlling linearly for our balance test variables. The estimates show that our findings are highly robust to such controls, which is not surprising given that the discontinuity design works well. The estimates do change slightly because we lose around a quarter of the sample.

Table 8 shows the PRI interaction results with distance when controlling for our balancing variables. This table also includes the district-year-discontinuity fixed effects, cited in the main text, in specification (24); these capture a wide variety of possible concerns such as interactions with race-specific characteristics and the party of the state governor. The results indicate that our theoretical claims are robust: the non-linear PRI interaction with distance is highly robust to the inclusion of Census variables. The distance interactions also weaken in ways consistent with our model when we control for interactions with lagged vote share and (log) area. First, since lagged vote share is (like distance to the polling station) also a good proxy for the number of potential voters in a precinct, it is not surprising to find that both distance and lagged vote share are capturing similar variation. This explains the smaller coefficients in the first three specifications, and in fact indicates that distance is doing a job at capturing potential voters. Nevertheless, our results indicate that the interaction with distance generally remains statistically significant, which is encouraging since lagged vote share is a good proxy for potential voters but does not as effectively capture the costs of mobilizing brokers. Second, area is highly correlated ($\rho > 0.5$) with distance to the

Table 6: Local linear regression estimates

	Turnout	PRI vote share	PAN vote share	PRD vote share
	(1)	(2)	(3)	(4)
Panel A: Average effects				
Split	0.0102***	0.0052**	0.0049	0.0003
	(0.0024)	(0.0025)	(0.0032)	(0.0015)
Observations	27,697	27,697	27,697	27,697
Panel B: Heterogeneous effects				
Split	0.0108***	0.0042*	0.0053*	0.0007
	(0.0023)	(0.0024)	(0.0031)	(0.0015)
Distance	-0.0067	0.0058	-0.0089***	-0.0017
	(0.0049)	(0.0038)	(0.0031)	(0.0021)
Distance squared	-0.0003	-0.0006	0.0003	0.0000
	(0.0006)	(0.0008)	(0.0006)	(0.0003)
Split $\times$ Distance	0.0014	0.0110***	-0.0035	-0.0002
	(0.0052)	(0.0038)	(0.0029)	(0.0031)
Split $\times$ Distance squared	-0.0009	-0.0030**	0.0010*	-0.0005
	(0.0016)	(0.0013)	(0.0006)	(0.0007)
Observations	27,417	27,417	27,417	27,417

Notes: All specifications include district-year fixed effects, and are estimated with OLS. All results are for a 20 voter bandwidth. Block-bootstrapped standard errors are clustered by district. Locality-weighted distance to the polling station was unavailable for a small number of electoral precincts. * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

Table 7: Controlling for balance variables

	Turnout (1)	PRI vote share (2)	PAN vote share (3)	PRD vote share (4)
Panel A: Average effects				
Split	0.0070***	0.0031***	0.0031***	0.0005
	(0.0010)	(0.0008)	(0.0010)	(0.0007)
Observations	20,788	20,788	20,788	20,788
Panel B: Heterogeneous effects				
Split	0.0073***	0.0029***	0.0030***	0.0008
	(0.0012)	(0.0007)	(0.0009)	(0.0009)
Distance	-0.0016	-0.0027	-0.0002	-0.0001
	(0.0037)	(0.0023)	(0.0025)	(0.0017)
Distance squared	-0.0000	0.0002	0.0001	-0.0000
	(0.0007)	(0.0003)	(0.0004)	(0.0003)
Split × Distance	-0.0013	0.0035	0.0003	-0.0004
	(0.0045)	(0.0027)	(0.0025)	(0.0032)
Split × Distance squared	-0.0001	-0.0015**	0.0003	-0.0005
	(0.0014)	(0.0006)	(0.0006)	(0.0009)
Observations	20,788	20,788	20,788	20,788

Notes: All specifications include district-year fixed effects, and are estimated with OLS. All results are for a 20 voter bandwidth. Block-bootstrapped standard errors are clustered by district. Locality-weighted distance to the polling station was unavailable for a small number of electoral precincts. * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

Table 8: Heterogeneous effects of splitting polling stations on PRI vote share, condition upon distance and controlling for alternative interactions

	(1) PRI vote share (lag)	(2) PAN vote share (lag)	(3) PRD vote share (lag)	(4) Area (log)	(5) Voter density	(6) Share economically active	(7) Share employed	(8) Share medical insurance
Split × Distance	0.0033	0.0057*	0.0064*	0.0067	0.0103**	0.0096***	0.0112***	0.0113***
	(0.0025)	(0.0034)	(0.0034)	(0.0050)	(0.0047)	(0.0033)	(0.0040)	(0.0038)
Split × Distance sq.	-0.0013**	-0.0019*	-0.0019*	-0.0022*	-0.0028**	-0.0027**	-0.0031**	0.0030**
	(0.0005)	(0.0010)	(0.0011)	(0.0013)	(0.0015)	(0.0010)	(0.0013)	(0.0012)
Observations	22,450	22,450	22,450	24,355	24,355	27,404	27,404	27,404
	(9) Share illiterate	(10) Incomplete primary school	(11) Complete primary school	(12) Incomplete secondary school	(13) Complete secondary school	(14) Share owns house	(15) Share basic amenities	(16) Share with radio
Split × Distance	0.0116***	0.0105***	0.0107***	0.0105***	0.0103***	0.0106***	0.0086**	0.0096***
	(0.0037)	(0.0035)	(0.0038)	(0.0037)	(0.0035)	(0.0036)	(0.0037)	(0.0039)
Split × Distance sq.	-0.0028**	-0.0027**	-0.0030**	-0.0030**	-0.0030**	-0.0030**	-0.0025**	-0.0028**
	(0.0012)	(0.0011)	(0.0013)	(0.0013)	(0.0012)	(0.0013)	(0.0011)	(0.0014)
Observations	27,404	27,404	27,404	27,404	27,404	27,409	27,409	27,409
	(17) Share with TV	(18) Share with fridge	(19) Share washing machine	(20) Share with car	(21) Share with telephone	(22) Share with cell phone	(23) Share with internet	(24) District-year-discontinuity fixed effects
Split × Distance	0.0104***	0.0100***	0.0093***	0.0104***	0.0082**	0.0085**	0.0094**	0.115**
	(0.0038)	(0.0038)	(0.0036)	(0.0035)	(0.0035)	(0.0034)	(0.0037)	(0.0048)
Split × Distance sq.	-0.0028**	-0.0028**	-0.0026**	-0.0028**	-0.0025**	-0.0024**	-0.0026**	-0.0031*
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0010)	(0.0012)	(0.0016)
Observations	27,409	27,409	27,409	27,409	27,409	27,409	27,409	26,697

Notes: Each specification is estimated separately using OLS and including district-year fixed effects, in addition to quadratic interactions of the variable listed at the top of each column with the split indicator (with the exception of column (24) which includes district-year-discontinuity fixed effects). All results are for a 20 voter bandwidth.Block-bootstrapped standard errors are clustered by district. * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

Table 9: Controlling for district-year-discontinuity fixed effects

	Turnout (1)	PRI vote share (2)	PAN vote share (3)	PRD vote share (4)
Distance	-0.0071	0.0065*	-0.0102***	-0.0017
	(0.0052)	(0.0037)	(0.0038)	(0.0024)
Distance squared	-0.0003	-0.0006	0.0004	0.0001
	(0.0007)	(0.0009)	(0.0006)	(0.0004)
Split × Distance	0.0034	0.0115**	-0.0016	-0.0004
	(0.0061)	(0.0048)	(0.0041)	(0.0041)
Split × Distance squared	-0.0013	-0.0031**	0.0008	-0.0006
	(0.0017)	(0.0016)	(0.0008)	(0.0008)
Observations	26,701	26,701	26,701	26,701

Notes: All specifications include district-year-discontinuity fixed effects (and thus the lower-order split term is omitted), and are estimated with OLS. All results are for a 20 voter bandwidth. Block-bootstrapped standard errors are clustered by district. Locality-weighted distance to the polling station was unavailable for a small number of electoral precincts. * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

polling station. It is thus unsurprising to find a slight decrease in coefficient magnitude and loss of precision for these interactions. Finally, Table 9 shows that the results are robust to the inclusion of district-year-discontinuity fixed effects.