What is the Effect of Turnout Buying? Theory and Evidence from Mexico *

Horacio A. Larreguy[†] John Marshall[‡] Pablo Querubín[§]

June 2014

Vote and turnout buying are widespread in developing democracies, but little is known about the conditions under which such practices are effective. Rather than focusing on the party-voter relationship, we argue that the efficacy of turnout buying depends upon the moral hazard problem underlying the party-broker relationship. Parties can extract more effort from their brokers where they can better monitor their performance. Exploiting random variation in the number of polling stations—and thus information about broker performance—in Mexican electoral precincts, we show that greater monitoring capacity increases turnout and the number of votes for the PAN and especially the PRI. Consistent with our model, the PRI vote share depends non-linearly upon the distance of voters to the polling station: it first increases because rural voters generally favor the PRI but face prohibitive costs of voting, before declining as the cost of incentivizing brokers increases.

*We thank Gustavo Rivera Loret de Mola for useful comments. All remaining errors are our own.
[†]Department of Government, Harvard University, hlarreguy@fas.harvard.edu.
[‡]Department of Government, Harvard University, jmarsh@fas.harvard.edu.
[§]Department of Politics, New York University, pablo.querubin@gmail.com.

1 Introduction

The exchange of goods for voters turning out or voting for a particular political party is widely reported across the developing world.[1] Given that political parties struggle to effectively monitor their political brokers on the ground and voters once in the polling booth, the efficacy of such strategies remains uncertain. In Mexico and many other contexts, newspaper reports and political surveys suggest that neither vote nor turnout buying is especially effective or prevalent.[2] However, although voters likely under-report selling their votes (e.g. Gonzalez-Ocantos et al. 2012; Nichter and Palmer-Rubin forthcoming) and election monitors consistently identify substantial malpractice,[3] researchers have struggled to systematically quantify the effectiveness of such practices. In this article, we exploit variation in the number of polling stations in Mexican electoral precincts to identify the effect of increasing the capacity of political parties to monitor their brokers on turnout buying.

Despite emerging from decades of one-party rule by the Institutional Revolutionary Party (PRI), Mexican elections are still characterized by clientelism and vote buying. Although deeply embedded clientelistic ties provide voters with ongoing employment and access to government programs remain prevalent and effective (Cornelius 2004; Diaz-Cayeros, Estévez and Magaloni forthcoming; Magaloni 2006), this paper focuses on voter mobilization outside of these struc-

[1]For recent surveys of this literature, see e.g. Hicken (2011), Kitschelt and Wilkinson (2007), Stokes et al. (2013), Vicente and Wantchekon (2009).

[2]See newspaper reports such as "Dádivas, sin incidencia en elección", *Excélsior*, 27th August 2012, and journalistic election accounts like Gasca (2013), which use qualitative and survey data to examine the number of voters reporting receipt of a gift and the number that then voted for the party they received it from. Similarly, the Mexican Panel Study typically only reports that around 5% of voters received a gift in exchange for their vote (e.g. Nichter and Palmer-Rubin forthcoming). However, in a detailed election report based on interviews and qualitative assessments, Ugalde and Rivera Loret de Mola (2013) argue that little is known about how voter mobilization strategies actually work.

[3]See Alianza Cívica, Boletín de Prensa, July 3rd 2012, who estimate that 14% of voters were brought to the polling station while 21% did not cast their vote secretly.

tures just before and especially on election day. Since 2000, Mexico's main political parties have engaged extensively in turnout buying strategies, offering gifts in exchange for turning out and and illegally hiring taxis to drive voters to polling stations.[4] Political brokers hired by political parties play the essential intermediary role in this process, mobilizing voters on election day in exchange for cash and bonuses (or sanctions) based on local electoral performance (Ugalde and Rivera Loret de Mola 2013).

While previous research has focused on the monitoring problem between parties and voters under a secret ballot (e.g. Nichter 2008; Stokes 2005), limited attention has been paid to the relationship between parties and political brokers. This neglect is surprising given that brokers typically implement voter mobilization strategies on election day and possess detailed knowledge about local voters that political parties lack. Stokes et al. (2013) regard the interaction with brokers principally as a selection problem for political parties seeking to employ the best-connected brokers. However, because brokers are generally hired contractors rather than actors with incentives closely tied to political parties, this misses a critical moral hazard concern: political parties hire well-informed local brokers to procure votes, but brokers face strong incentives to shirk given that parties cannot easily monitor their actions (Larreguy 2013).

Guided by this moral hazard problem, we formalize a theoretical model identifying the conditions under which brokers mobilize turnout. In our model, parties hire political brokers to identify and mobilize their pool of "potential voters"—favorable voters that would not otherwise turn out. Given brokers will shirk if they believe that they can go undetected and still receive payment, parties can demand more votes in exchange for higher payment in locations where their monitoring capacity is greater. The level of turnout buying also depends upon local characteristics determining the pool of potential voters. In particular, we show that the effects of monitoring capacity depend non-linearly on the distance to the polling station: parties that are relatively popular among ru-

[4]See summaries such as Nichter and Palmer-Rubin (forthcoming) and Ugalde and Rivera Loret de Mola (2013), and many reports including Alianza Cívica, Boletín de Prensa, July 3rd 2012 and those in footnote 10.

ral voters (who face higher costs to turning out) have most to gain from the turnout buying that increased monitoring capacity permits, and thus contract more effort from brokers in more distant locations until the cost of hiring brokers to attract the most distant voters becomes too high. Among parties that do well in rural areas, we therefore expect to observe an "inverted-U" relationship between distance and turnout buying. If parties were buying the votes of non-supporters, as vote buying models predict (e.g. Stokes 2005), we should instead expect a negative interaction with distance.

We take our theoretical model to the data by leveraging two sources of variation. First, differences in monitoring capacity arise from an electoral rule that requires a new polling station be created for every 750 registered voters in an electoral precinct. An additional polling station provides valuable information to political parties who use polling station-level electoral performance to reward and punish their brokers. Second, to ensure that we are picking up differences in monitoring at the discontinuity, we capture variation in political preferences and the cost of voting by calculating the average distance that voters must travel to their precinct's polling booth. On average, voters living further from their polling station are less likely to turn out and more likely to support the PRI.

Using a regression discontinuity design, comparing polling stations in electoral precincts just above and just below the threshold for creating a new polling station, we find strong evidence of turnout buying consistent with our theoretical model. An additional polling station increases electoral turnout by nearly one percentage point, significantly benefiting the right-wing National Action Party (PAN) and especially the PRI. The vote share of the Party of the Democratic Revolution (PRD), which has traditionally campaigned against clientelistic practices, is unaffected.

Following the predictions of our theoretical model, we also examine how the effect of an additional polling station varies with distance. Consistent with our theory, the change in PRI vote share at the discontinuity increases with distance until the average distance to the polling station becomes prohibitively expensive. At its maximum—where, on average, voters live 1.75km from the polling

station—the PRI gains the vote of more than one percentage of registered voters. This interaction is not present for the PAN, who stand to gain less from mobilizing rural voters. Since some non-monitoring explanations could also explain the difference at the discontinuity, demonstrating that the effect varies non-linearly with distance—for the PRI, but not other political parties—provides considerable support for our model.

Although the causal effects we identify are relatively small in magnitude, it is important to remember that our empirical strategy only quantifies the impact of a single dimension over which monitoring capacity varies. Furthermore, since parties always receive at least one signal of broker performance, our design captures large effects of relatively small differences in monitoring capacity. Ultimately, our evidence primarily serves to illuminate the monitoring mechanisms facilitating effective turnout buying by political brokers, in addition to identifying a tip of the iceberg of clientelistic practices.

Our theoretical results contribute to the nascent literature focusing on the intermediary role of political brokers. This literature departs from extant work assuming parties do not require brokers or that brokers serve as perfect agents.[5] Whereas Stokes et al. (2013) regard hiring brokers as an adverse selection problem, our model emphasizes the moral hazard problem arising from imperfect monitoring capacity. While Larreguy (2013) examines a similar signal extraction problem for longer-term clientelistic relationships, our model shows how heterogeneity in voter preferences and costs of voting causes parties to face differential incentives when mobilizing voters outside clientelistic structures.

Empirically, our results speak to previous studies identifying the extent of vote and turnout buying in Africa, Asia and Latin America (see Gonzalez-Ocantos et al. 2012; Vicente and Wantchekon 2009). We extend such research using quasi-experimental variation to isolate evidence of turnout

[5]There is now a growing formal literature examining the monitoring mechanisms employed by brokers, rather than parties, viś-a-viś voters (e.g. Gingerich and Medina 2013; Rueda 2013; Smith and de Mesquita 2012); Robinson and Verdier (2013) instead consider the reverse credibility problem. Our study, however, focuses on party monitoring of brokers.

buying. However, by demonstrating the importance of monitoring technologies for turnout buying, our findings also significantly advance this literature by evaluating the effectiveness of turnout buying. Our results differ from Larreguy (2013), who identifies the effects of monitoring with respect to state-level clientelism. Unlike observational studies of broker behavior in Argentina (Stokes et al. 2013; Szwarcberg 2012*a*), we exploit exogenous variation to identify the conditions under which electoral manipulation occurs.

The remainder of the paper is structured as follows. Section 2 provides an overview of turnout buying and the role of brokers in Mexico. Section 3 presents our theoretical model. Section 4 describes our data and explains our identification strategy. Section 5 provides our results. Section 6 concludes.

2 Electoral manipulation in Mexico

Mexico has experienced a long history of electoral malpractice. During its 71-year stranglehold on power extending back to 1929, the PRI was widely acknowledged to have engaged in clientelistic transfers, vote buying and electoral fraud (e.g. Cornelius 2004; Magaloni 2006). After allegations of widespread vote-rigging in the 1988 elections, and the rise of stronger challengers to the PRI's dominance, election monitoring—principally through the creation of the independent IFE—has become more effective at preventing the most flagrant electoral violations (Cornelius 2004).

However, according to a slew of qualitative evidence contained in newspaper articles, surveys and election reports, Mexico's main political parties—the PAN, PRD and especially PRI—continue to pressure voters using more subtle tactics (see also Nichter and Palmer-Rubin forthcoming). This has occurred in spite of the PRI's ultimately victorious Presidential candidate, Enrique Peña Nieto, promising to break from the electoral manipulation typically associated with the PRI in 2012. We focus on what Gans-Morse, Mazzuca and Nichter (2013) term "electoral clientelism", or the distribution of benefits by political parties during election campaigns.

2.1 Electoral clientelism

Reports of voters receiving gifts, including money, food, clothing and gift cards, from political parties at election time are extensive. Although gifts that are not conditional on voting for a particular party are legal under Mexican law, vote buying—where gifts are exchanged for voting a particular way—is still regarded as a regular phenomenon. One of the most egregious examples from 2012 was the widely reported allegation that the PRI distributed millions of gift cards for the supermarket Soriana. Voters were told that these cards would become active upon the PRI winning the 2012 election. Based on their election monitoring, Alianza Cívica estimate that the cost of a vote is 100-800 pesos (8-60 U.S. dollars).[6]

Directly asking voters whether they received gifts from political parties suggests that only around 5% had. Furthermore, several newspaper articles have concluded that such mobilization strategies are relatively ineffective based on survey responses, and therefore do not pose a major challenge to democratic outcomes.[7] However, these figures almost certainly suffer considerable social desirability bias (Gonzalez-Ocantos et al. 2012). In 2012, Alianza Cívica reported that 28% of voters had personally received a gift, or knew someone that had.[8] Similarly, a list experiment conducted before Mexico's 2012 election found 22% received a gift (Nichter and Palmer-Rubin forthcoming). Surveys consistently indicate that voters are more likely to be contacted by the PRI and then the PAN and PRD (Stokes et al. 2013).

However, not all gifts and incentives are provided in exchange for voters switching their vote intention. Given the difficulty of parties and brokers monitoring voter behavior once inside the polling booth, voters may renege on their promises with impunity (Stokes 2005).[9] When voters

[6] Alianza Cívica, Boletín de Prensa, July 3rd 2012.

[7] Gasca (2013) suggests less than one third of voters vote for the party they received an incentive from. See also e.g. "Dádivas, sin incidencia en elección", *Excélsior*, 27th August 2012.

[8] Alianza Cívica, Boletín de Prensa, July 3rd 2012.

[9] Although voters can be observed in the booth by children or provided with mobile phones to photograph their marked ballot (Ugalde and Rivera Loret de Mola 2013), Alianza Cívica reports that only 21% of votes are not conducted in secret.

cannot be effectively monitored, Nichter (2008) argues that parties skirt the commitment problem by instead mobilizing voters who they expect to support the party but would not otherwise turn out to vote. Consistent with this, Nichter and Palmer-Rubin (forthcoming) found that gifts were most frequently targeted at weak PRI supporters.

One of the most widespread turnout buying practices, *acarreo*, involves transporting voters to the polling station. *Acarreo* is illegal under Article 403 of the Mexican Federal Penal Code. Nevertheless, newspaper accounts from across the country reported extensive use of *acarreo* in 2012 by hired coaches and especially groups of taxi drivers.[10] Alianza Cívica report that the proportion of voters brought to polling stations increased in both 2009 and 2012 to reach 14%.[11] Transportation of this sort appears to have been particularly prevalent in areas where the polling station is not easily accessible to voters. Although the PAN and PRD have also been accused of engaging in *acarreo*, it has predominantly been associated with the PRI. In fact, one report suggests that the PRI attempted to disguise such activities by labeling its taxis with PRD stickers.[12] Another popular practice, known as *operación tamal*, entails gathering a large group of voters together for breakfast before transporting them to the polling station in exchange for additional gifts.[13]

[10]For example, see: "Gana Pri En Huauchinango En Medio De Señalamientos De Compra De Votos, Acarreo De Gente E Intimidaciones," Diario Reforma, July 3rd 2000; "Compra de votos, falta de boletas en casillas especiales y acarreo, las quejas recurrentes," SinEmbargo.mx, July 1st 2012; "Evidente acarreo de votantes en elecciones del PRD," ABC Tlaxcala, April 8th 2013; "Gana Pri En Huauchinango En Medio De Sealamientos De Compra De Votos, Acarreo De Gente E Intimidaciones," El Imparcial de la Sierra Norte, July 3rd 2013; 'Acusan Al PRI De 'Acarreo'," *El Siglo de Torreón*, July 8th 2013; "Vecinos denuncian presunto acarreo en Miguel Hidalgo," *El Universal*, September 1st 2013.

[11]Alianza Cívica, Boletín de Prensa, July 3rd 2012. Levitsky (2014) finds that brokers perform a similar role in Argentina.

[12]"Muchos ojos, pero pocos votos, en la zona conurbada y rural de Acapulco," *La Jornada*, July 6th 2009.

[13]Such practices could also incorporate vote buying as well. This is what Gans-Morse, Mazzuca and Nichter (2013) call "double persuasion". Our empirical analysis, however, provides good reasons to believe that we are identifying turnout buying rather than vote buying.

2.2 The role of brokers

Given the scale and extensive information requirements of these turnout buying operations, parties hire non-party local operatives to implement these strategies on the ground. Political brokers typically operate at a very local level (within electoral precincts), and possess detailed knowledge of the vote and turnout intentions of the local population that state and municipal officials lack.[14] These brokers—who provide transport, round up groups of potential voters, monitor voting at polling stations, and distribute gifts—are available to the highest bidder (Ugalde and Rivera Loret de Mola 2013).[15] In general, brokers are paid throughout the campaign and receive a bonus—in terms of either cash or political favors—for strong electoral performance (Ugalde and Rivera Loret de Mola 2013). For example, taxi drivers can be paid up to 2,000 (150 U.S. dollars) pesos for a day's work repeatedly ferrying voters to polling stations in their electoral precinct.

However, parties struggle to monitor the efforts of their brokers. The challenge for parties is to differentiate the effects of broker activity from other factors determining local vote outcomes. Since brokers realize that parties cannot easily identify departures from agreed efforts to mobilize voters, brokers have an incentive to shirk. Parties have sought to mitigate this moral hazard problem by demanding lists of voters whom the broker intends to bring to the polling station. These lists can be cross-checked using the "bingo system", whereby party representatives at the polling station on election day with access to the list of citizens that voted compare the two lists (Gasca 2013). Figure 1 depicts an example of such a list embossed with the PRI logo, where the broker would fill in the address, electoral precinct and voter ID of voters they promise to bring to the polls. It is also

[14] In Argentina and Peru, respectively, Stokes et al. (2013) and Finan and Schechter (2012) provide compelling survey evidence indicating that brokers possess sufficient information to target voters that they expect to reciprocate or favor a given party. More qualitative work also supports that importance of reciprocity (Auyero 2000), and points to the importance of broker centrality in local networks (Levitsky 2014; Szwarcberg 2012*a*).

[15] Szwarcberg (2012*b*) points to a similar logic in Argentina, where local political brokers are aspiring politicians who learn about the preferences and propensity of turn out of local voters. Brokers in Argentina appear to differ in that they are more interested in rising in the party hierarchy.

important to note that brokers are themselves designated to electoral precincts.

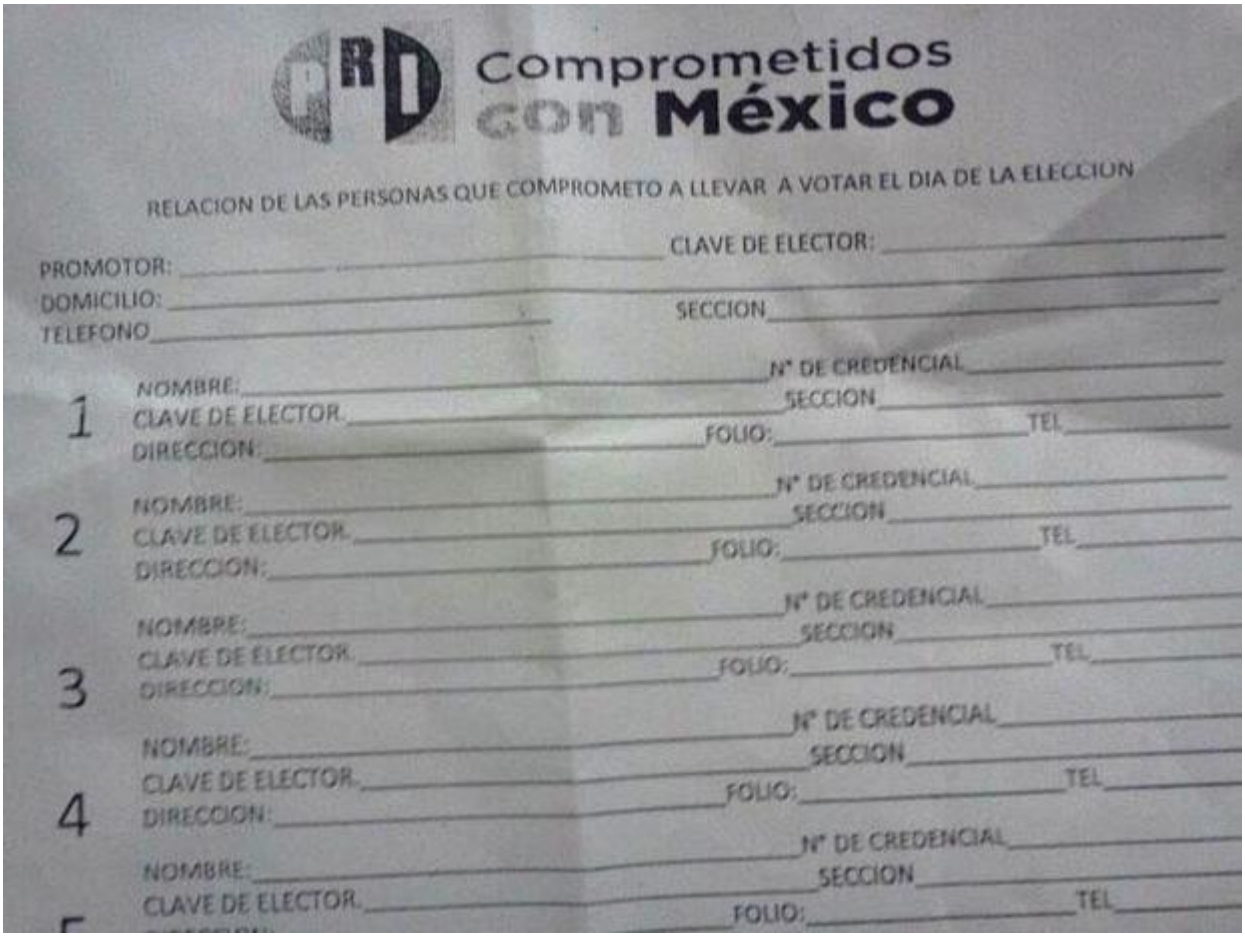

Figure 1: List of promised voters for the PRI to be completed by a political broker in a given electoral precinct

Notes: The top of the sheet (first three rows) indicates the name of the broker, address, telephone number and electoral precinct. Below this are the details of voters, including their name, electoral card number, electoral precinct, address and phone number.

Where brokers fail to meet electoral expectations, payments or rewards can be withheld. Based on interviews with Mexican politicians, Ugalde and Rivera Loret de Mola (2013) explain that parties evaluate the effectiveness of their voter mobilization apparatus at the precinct or polling station level, rather than at the voter level.

The role of political brokers has received limited attention. Most descriptive accounts of broker

behavior simply assume that brokers serve as perfect agents of their party principals. Breaking this trend, Stokes et al. (2013) focus primarily on the selection of brokers, arguing that brokers signal their competence by amassing large networks of voters. Given parties cannot identify voters in a broker's network that would have already voted for the party, and brokers have an incentive to amass these relatively cheap voters, this process may be fairly inefficient. Such models ignore the moral hazard problem described above by assuming that both parties and the brokers they hire face equal incentives to win the election (see also Szwarcberg 2012*a*).[16] However, Mexican brokers are typically paid both upfront and through post-election performance bonuses. While adverse selection is no doubt an important part of the story when parties first hire brokers, a bigger challenge is moral hazard because party and broker incentives are rarely perfectly aligned in practice and political parties cannot perfectly monitor the effort of their voters.

3 Theory

This section first formalizes our model of the relationship between parties, brokers and voters. We then show how the model's predictions apply in the Mexican context, and thus generate testable hypotheses for our empirical analysis.

3.1 Model

Our model examines turnout buying by political parties using brokers at the electoral precinct level. The key feature of the model is the moral hazard problem faced by political parties: they hire political brokers with the local knowledge required to mobilize favorable local voters, but cannot perfectly observe the effort exerted by political brokers in this task. Where parties are better able to monitor brokers, they can generate more electoral support through turnout buying. The second

[16]For Stokes et al. (2013), this arises because brokers only receive rents conditional on their party winning office. Qualitatively, Szwarcberg (2012*a*) observes that many brokers in Argentina are aspiring politicians.

main feature of the model is that voters' political preferences and costs of turning out, as well as the cost of compensating brokers, vary with their distance from the polling station. Consequently, parties face differential incentives to use brokers to mobilize voters facing high costs of turning out.

3.1.1 Setup

Consider a country containing n electoral precincts. At each electoral precinct there exists a continuum of voters, whose mass we normalize to unity. Electoral precincts differ in the distance $d > 0$ that voters must travel to their polling station. For simplicity, we assume all voters travel the same distance.[17] Electoral precincts also differ in the probability $p \in [0,1]$ that political parties can exactly monitor the behavior of the broker which they hire. Without loss of generality, we consider an electoral precinct defined by distance d in our analysis.

We consider two political parties $i = L,R$ competing for votes in each electoral precinct, roughly denoting a left and right wing party. Parties care about maximizing their vote share Π_i in the precinct,[18] and can hire political brokers at wage w_i to help them. Parties cannot themselves identify which voters to mobilize, and so they choose an effort-wage contract (e_i, w_i) to induce brokers to mobilize voters by exerting effort $e_b \geq 0$. If the party observes the broker's effort and the broker complied with the agreed effort level, such that $e_b \geq e_i$, then the agent receives wage w_i; if the broker is found to not have complied with the agreed effort level then he receives no payment.[19] If the party cannot observe the effort level, the broker receives wage w_i.

[17] We obtain very similar results if there is a distribution of voters, $f(d)$, in each electoral precinct, where fewer voters live further from the polling station ($f'(d) < 0$) and there are no voters at the extremes ($\lim_{d\to\infty} f(d) = 0$).

[18] If parties instead maximized their probability of winning districts or a legislative majority, the implications of our model are unchanged. Accordingly, parties maximize precinct vote share for simplicity.

[19] Brokers therefore have limited liability in that parties cannot punish brokers beyond refusing to pay their wage after observing low effort. We assume that parties have resolved the commitment problem that they would prefer to pay the broker nothing once voting has occurred. It is easy to

Political brokers enjoy an information advantage over political parties: brokers can identify the voters in their electoral precinct that need additional incentives to turn out, but would vote for party i if they do turn out. In particular, a broker chooses to exert effort e_b to increase the probability $h(e_b) \in [0,1)$ that a given voter will turn out. The broker's turnout buying technology $h(e_b)$ satisfies $h'(e_b) > 0$ and $h''(e_b) < 0$. The effectiveness of the broker thus exhibits diminishing marginal returns to effort. However, exerting turnout buying effort—which could constitute calling in favors, hiring coaches and drivers, or providing material incentives to voters—entails a cost $C(d,e_b)$ to the broker, where $C(d,0) = C(0,e) = 0$, $C_e(d,e_b) > 0$, $C_{ee}(d,e_b) > 0$, $C_{de}(d,e_b) > 0$ and $C_{dde}(d,e_b) > 0$ (subscripts denote partial derivatives). In words, exerting no effort is costless to brokers, while the cost of exerting additional effort is convex and exponentially increasing in the distance of voters from the polling station. If brokers are not hired by political parties, we assume they receive zero utility. Conditional upon engaging in the contract, a strategy for a broker hired by party i is to choose their effort level $e_b(i)$.

Finally, voters in each electoral precinct differ in the valence shock σ toward party R that they receive.[20] This valence shock is an expressive benefit (see e.g. Brennan and Hamlin 1998, 2000), such that it is only received by voters when they turn out and vote for their preferred candidate.[21] The shock is distributed over support $\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$, where $\psi > 0$ captures the shock's salience, according to the following density function that is drawn independently of d at each polling station:

$$g(\sigma;d) = \psi[1 - b(d)\sigma]. \tag{1}$$

rationalize this by considering interactions with brokers across elections. Stokes (2005) shows how this can occur between brokers and voters.

[20] Voters' policy utility is not included in the model because that is not the focus of this analysis. We could easily introduce policy utility $u(i,v)$ for voter type v from the platform of party i. However, allowing policy utilities to vary across voters does not affect the insights of the model, so we effectively assume $u(L,v) = u(R,v)$ for all voters.

[21] Since an individual's marginal effect on the probability of winning is zero with a continuum of voters, we use expressive voting to ensure non-negligible turnout (see Palfrey and Rosenthal 1985 for low turnout in large elections).

This density function captures the idea that the bias of the valence shock depends upon the distance to the polling station through $b(d) \in [-2\psi, 2\psi]$, where b is a monotonic function. Voters also face a cost to turning out to vote, $c(d) \in [0, \frac{1}{\psi}]$, such that $c(0) = 0$ and $c'(d) > 0$. A strategy for a voter is the decision to vote for a given party or not turn out: $v(\sigma, d) \in \{L, R, \emptyset\}$.

The expected bias toward party R, in an electoral precinct of type d, is $\mathbb{E}[\sigma|d] = -\frac{b(d)}{12\psi^2}$. Accordingly, party L benefits on average in electoral precincts of type d if $b(d) > 0$ because the density function $g(\sigma; d)$ reduces the likelihood of voters receiving a pro-R valence shock. In competitive precincts, where $b(d) = 0$, the vote is split in expectation. The term $b(d)$ can therefore be interpreted as bias in favor of party L. Furthermore, we assume $b'(d) > 0$ such that party L gains relatively more support among each voter type d from the valence shock vis-à-vis party R as the distance to the polling station increases.[22]

The timing of the game is as follows:

1. Parties $i = L, R$ offer brokers a contract (e_i, w_i) to induce voters to turn out.

2. The valence shock σ is realized for all voters, but is only observed by voters and brokers.

3. Voting occurs according to $v(\sigma, d)$, brokers exert effort e_b to mobilize voters that would not otherwise turn out, and e_b is observed by parties with probability p.

4. The election outcome and broker payment occur, and payoffs are realized.

We now proceed to identify the contracts that define the subgame perfect Nash equilibrium (SPNE) of this game.

3.1.2 Equilibrium

A central component of the contracting problem is the number of voters political parties can expect brokers to mobilize to turn out. Expressive voters have a dominant strategy to always vote for party

[22] This is without loss of generality to the extent that we could equally have chosen $b'(d) < 0$.

R if $\sigma \geq c(d)$, and always vote for party L if $-\sigma \geq c(d)$, without requiring broker inducements to turn out. If $c(d) > |\sigma|$, a voter will not turn out. Figure 2 illustrates this graphically for a given distance d, and shows a higher density of L voters when $b(d) > 0$. Absent the use of brokers, the political parties receive the following vote shares:

$$\pi_L \equiv \frac{1}{2} + \frac{b(d)}{8\psi} - P_L(d), \tag{2}$$

$$\pi_R \equiv \frac{1}{2} - \frac{b(d)}{8\psi} - P_R(d), \tag{3}$$

where $P_L(d) \equiv \psi c(d)\left[1 + \frac{1}{2}b(d)c(d)\right]$ and $P_R(d) \equiv \psi c(d)\left[1 - \frac{1}{2}b(d)c(d)\right]$. The $\frac{b(d)}{8\psi}$ term captures the bias noted above, while $P_i(d)$ represents the number of "potential voters"—voters who vote for the party if they reach the polling booth, but do not vote because the cost of turning out is too high.[23] The proportion of voters that would not turn out is given by $1 - (\pi_R + \pi_L) = 2\psi c(d)$, which is increasing in the cost $c(d)$ of voting.

Since individual vote choices cannot be bought, political parties care about mobilizing their potential voters. By perfectly observing the valence shock, brokers can precisely identify such voters. From the perspective of party i at the point of contracting, their expected number of potential voters is $P_i(d)$. Given $c(d) > 0$, it is clear that the number of potential voters depends upon $b(d)$. Unsurprisingly, L has more potential voters if the bias is in their favor ($b(d) > 0$), and L's number of potential voters increases with distance from the polling station because $b'(d) > 0$. The effect of distance on R's number of potential voters is ambiguous, but is only positive when the bias toward L is large and the distance is not too large.[24] Consequently, party L generally has a stronger incentive to engage in turnout buying, and especially in rural areas, because it has more

[23] Our distinction between certain and potential voters is similar to that drawn by Nichter (2008), who considers which voters a party should target when voters vary in their partisanship and their costs of voting (see also Dunning and Stokes 2008; Gans-Morse, Mazzuca and Nichter 2013). Note that our model could easily be extended to permit vote buying such that the set of potential voters could also include potential supporters of the opposing party.

[24] We can see this by observing: $P_R'(d) = \psi[c'(d)[1 - b(d)c(d)] - \frac{1}{2}b'(d)c(d)^2]$.

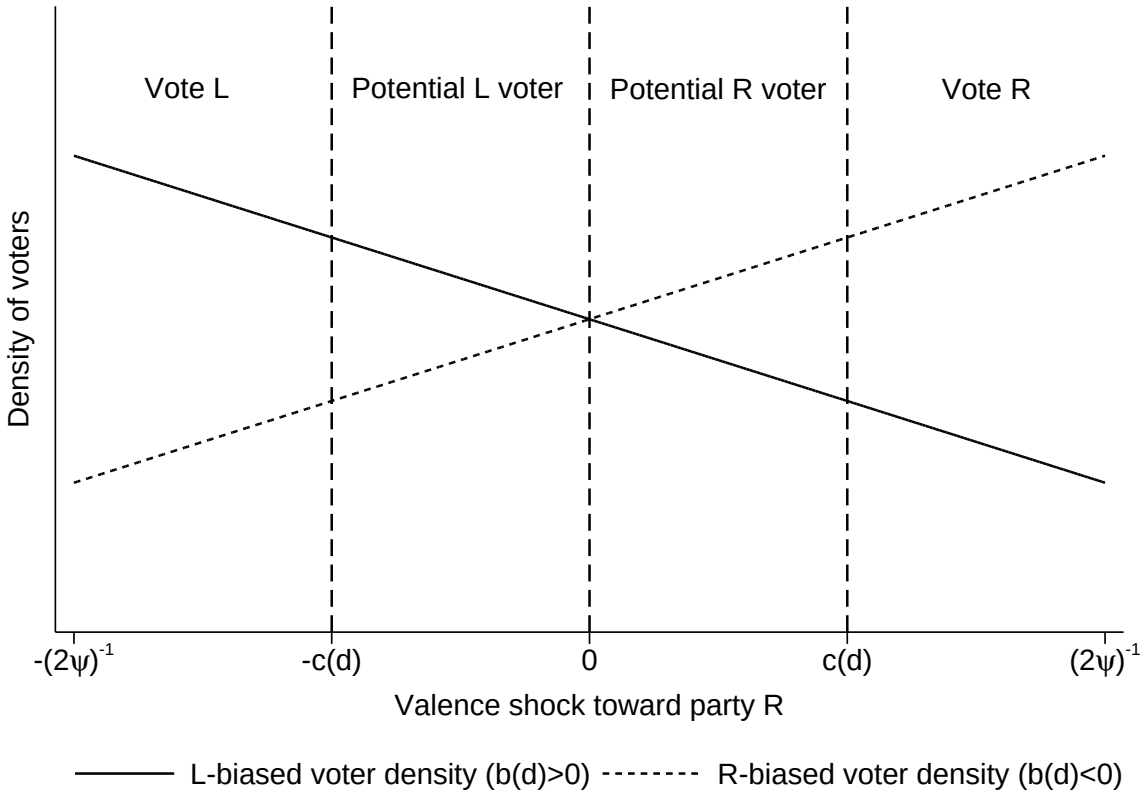

Figure 2: Vote choices at a given polling station

voters to gain.

After calculating the expected number of potential voters in a given electoral precinct, political parties offer brokers a contract. Parties design a contract (e_i, w_i) to both incentivize brokers to exert the desired level of effort (IC constraint), and to be willing to undertake the contract in the first place (IR constraint). The IC constraint induces the broker to choose to exert effort $e_b = e_i$ at cost $C(d, e_i)$, rather than choose $e_b = 0$ and receiving w_i if they are not caught (with probability $1 - p$)

and receiving zero when caught (with probability p). Party i thus solves the following program:

$$\begin{aligned} \max_{e_i, w_i} \quad & \pi_i + h(e_i)P_i(d) - w_i \qquad \text{subject to} \\ & (IC): \quad w_i - C(d, e_b(i)) \geq (1-p)w_i, \\ & (IR): \quad w_i - C(d, e_b(i)) \geq 0. \end{aligned} \tag{4}$$

Solving this problem leads directly to our equilibrium result:

Proposition 1 *The SPNE* $[(e_L^*, w_L^*), (e_R^*, w_R^*), e_b^*(i), v^*(\sigma, d)]$ *is defined by:*

$$\begin{aligned} ph'(e_i^*)P_i(d) - C_e(d, e_i^*) &= 0 \quad or \quad e_i^* = 0, \\ w_i^* &= \frac{C(d, e_i^*)}{p}, \\ e_b^*(i) &= e_i^*, \\ v^*(\sigma, d) &= \begin{cases} L & if\ \sigma \leq -c(d) \\ \emptyset & if\ \sigma \in (-c(d), c(d)) \\ R & if\ \sigma \geq c(d) \end{cases}. \end{aligned}$$

Equilibrium turnout is $T^* \equiv 1 - 2\psi\overline{c} + \sum_{i \in \{L,R\}} h(e_i^*)P_i(d)$*, of which* $\sum_{i \in \{L,R\}} h(e_i^*)P_i(d)$ *reflects turnout buying.*

Proof: see Appendix.

In equilibrium, parties offer brokers a contract to just induce optimal effort. The optimal amount of effort reflects two competing forces: the effectiveness of brokers at procuring voters (which requires that voters are available in the first place, otherwise $e_i^* = 0$), and the (probability of being monitored adjusted) cost of effort for which the broker must be compensated.

3.2 Comparative statics

The following proposition identifies the central testable predictions of the model:

Proposition 2 *In the SPNE in Proposition 1, the following comparative statics hold:*

1. T^*, Π_R^* *and* Π_L^* *are increasing in* p.
2. *Let* $B(d,e_i^*) \equiv ph'(e_i^*)P_i'(d) - C_{de}(d,e_i^*)$, *and assume* $B_d(d,e_i^*) < 0$ *and* $\lim_{d\to\infty} B(d,e_i^*) < 0$. *Then:*
   (a) *If* $B(d,e_i^*) < 0$ *for all* d, Π_i^* *is monotonically decreasing in* d.
   (b) *If* $B(d,e_i^*) > 0$ *for some* $d > 0$, *there exists a unique* $\overline{d}_i > 0$ *such that* Π_i^* *is increasing in* d *for* $d \in (0,\overline{d}_i]$ *and decreasing in* d *for* $d \geq \overline{d}_i$.
3. *If* $P_i'(d) > 0$, *the increase in* Π_i^* *due to* p *is increasing in* d.
4. Π_L^* *increases with* d *faster than* Π_R^* *increases in* d. *If* $b(d) > 0$, *then* $\Pi_L^* > \Pi_R^*$; *furthermore, the effects of* p *and* d *on* Π_L^* *are always greater than the effects of* p *and* d *on* Π_R^*.

The first prediction has an obvious interpretation: increased monitoring capacity increases turnout buying for both parties, and thus increases overall turnout in an electoral precinct. Intuitively, parties that can better monitor their brokers can more effectively threaten brokers with receiving a low wage, and can thus bargain for higher effort at a lower wage.

The more novel predictions of the model concern the effect of distance to the polling station on turnout buying. The second result in Proposition 2 shows that—at least for party L, for whom $B(d,e_i^*) > 0$ because $P_L'(d) > 0$—turnout buying is increasing in distance for small distances, but decreases in distance for sufficiently large distances. This result holds under the condition that, as the distance to the polling station increases, the marginal cost of compensating broker effort grows faster than the marginal benefit of exerting effort to mobilize potential voters, and for a

sufficiently large distance overpowers the marginal benefit. Since there is a diminishing effect of effort on mobilization (h is concave) and increasing costs to the broker ($C_{dde} > 0$), this restriction is theoretically appealing. The intuition is that one or both parties stand to gain votes by mobilizing some of the voters that would not have otherwise voted, but as the distance becomes sufficiently large the cost of inducing a broker to mobilize such voters becomes too high even for parties popular among rural voters. The third result demonstrates that monitoring always accentuates the benefits of mobilizing turnout for party L, because of its greater rural support, at any distance.

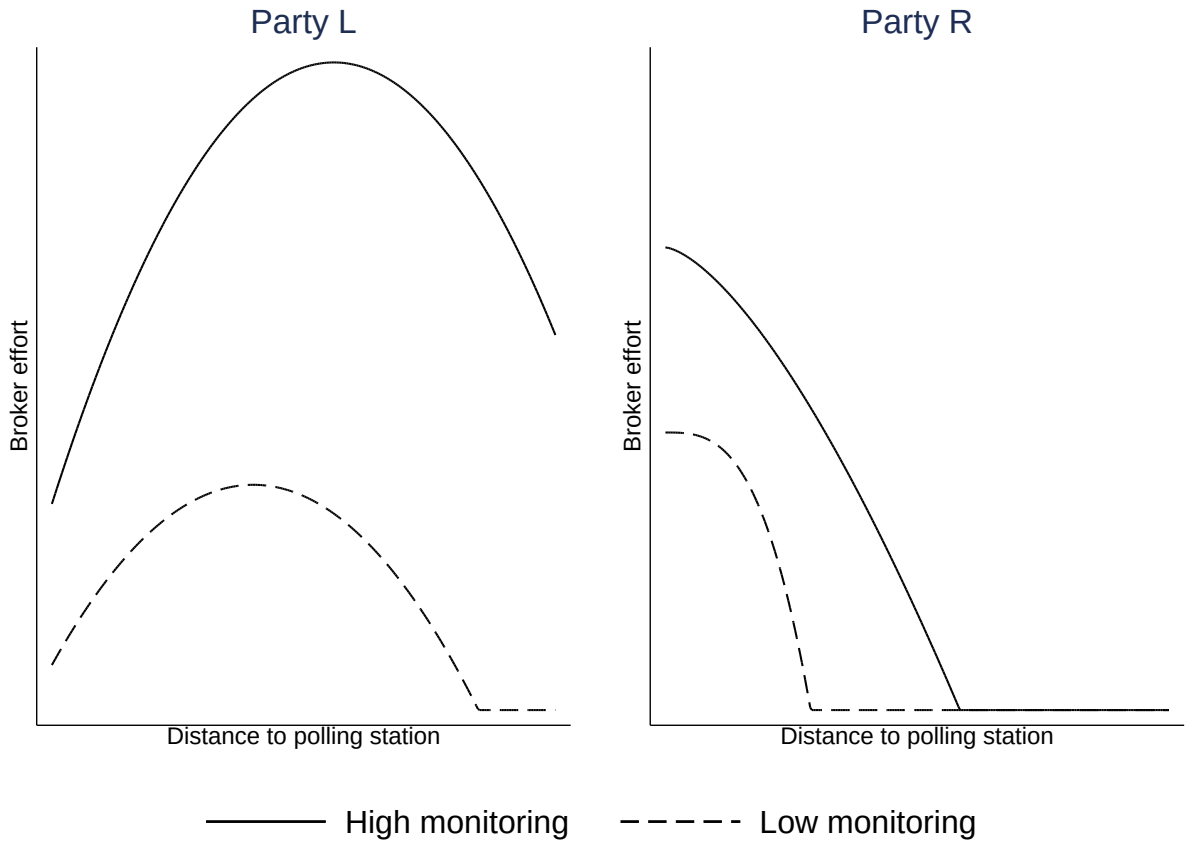

Figure 3: How distance effect's the optimal broker effort demanded by political parties

Note: The optimal effort demanded by party R does not necessarily decrease monotonically with distance from the polling station.

Figure 3 illustrates how the optimal effort contracted by each party changes with distance, depicting the case where R's optimal effort is always declining in d. Given the sign of $P_R'(d)$ is ambiguous, whether turnout buying by party R initially increases with distance for small distances depends upon the bias toward R in the electoral precinct being sufficiently large. The figure shows the importance of our third result by comparing the effort prescribed in a party's contract in electoral precincts with high and low monitoring capacity. Importantly, the first part of our proposition shows that broker effort is higher at every distance when there is greater monitoring capacity, while the third part demonstrates that the gradient with respect to distance is shallower for L (and often for R).

The fourth result summarizes the differential effects across political parties. First, it intuitively says that the party L will engage in relatively more turnout buying than the party whose voters are close to the polling station. When the average bias in an electoral precinct supports the more rural party L, turnout buying will always be greater for L regardless of the distance of voters from the polling station.

3.3 Implications in Mexico

The theoretical model has clear predictions for our Mexican case, to which the model applies well. Most regions of Mexico are essentially two-party races: between 2000 and 2012, only 14% of electoral precincts had a third party with more than 20% of the vote. Furthermore, as consistently shown in U.S. studies (see Brady and McNulty 2011; Gimpel and Schuknecht 2003), Figure 4 demonstrates that turnout declines with the average distance from the polling stations. Finally, as discussed above, there is considerable qualitative evidence pointing to the importance of local brokers in mobilizing votes for a given party.

A central prediction of the model is that turnout buying is more prevalent where parties are more effective at monitoring their brokers. Although monitoring could work in a variety of ways, this paper will examine how the number of polling stations within an electoral precinct affects

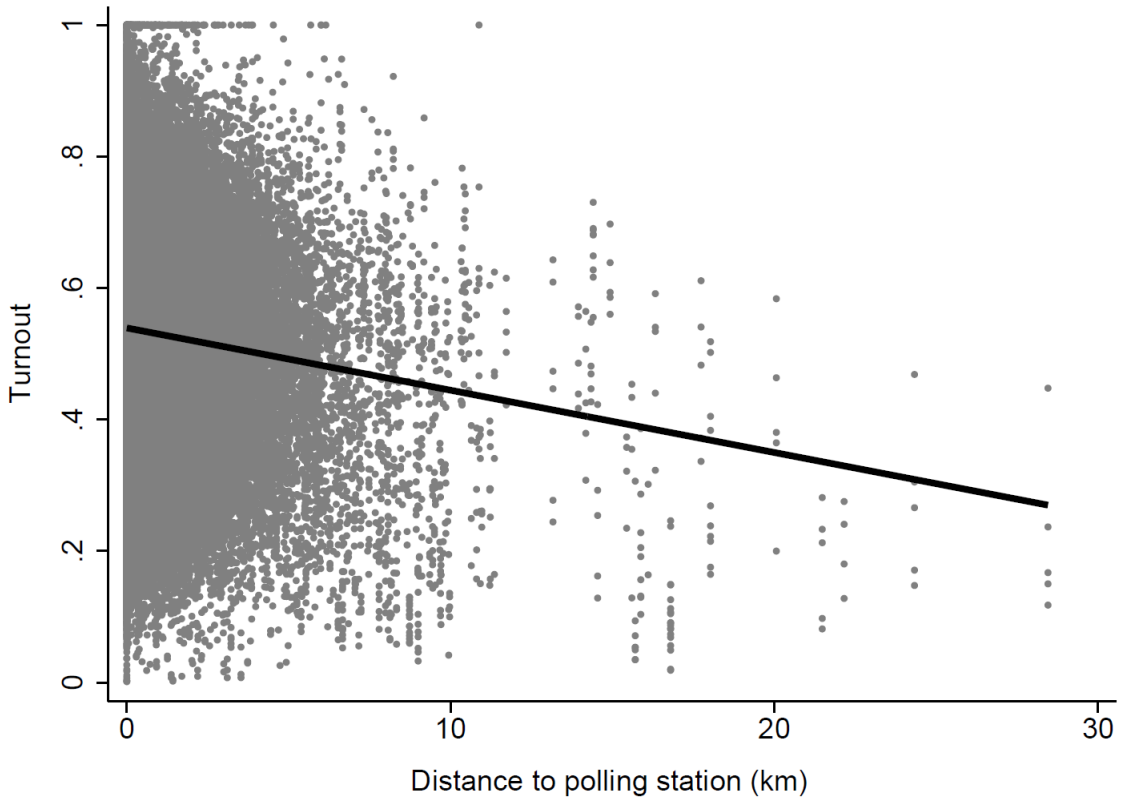

Figure 4: Turnout by average distance of voters to polling station

Notes: The black line is the best linear fit. Grey dots are individual polling stations.

turnout. In particular, we argue that a larger number of polling stations improves monitoring capacity by producing multiple signals of broker performance.[25] Multiple signals allow parties to better differentiate random shocks to their vote share from the systematic turnout buying efforts of their brokers: while precinct-level vote outcomes struggle to differentiate a positive shock from high broker effort, observing two signals of performance always conveys more information unless the shocks are perfectly correlated. For example, if parties expect broker efforts to equally affect voting across all polling stations in a given electoral precinct, then a large difference between polling stations is suggestive that the random shock was the predominant contributor to high (or

[25]More generally, vote buying is greater in smaller communities (Brusco, Nazareno and Stokes 2004; Gingerich and Medina 2013; Rueda 2013).

low) turnout. Regardless of the number of brokers operating in an electoral precinct, additional information always improves monitoring capacity.[26] We therefore hypothesize that:

H1. Turnout increases in the number of polling stations in an electoral precinct.

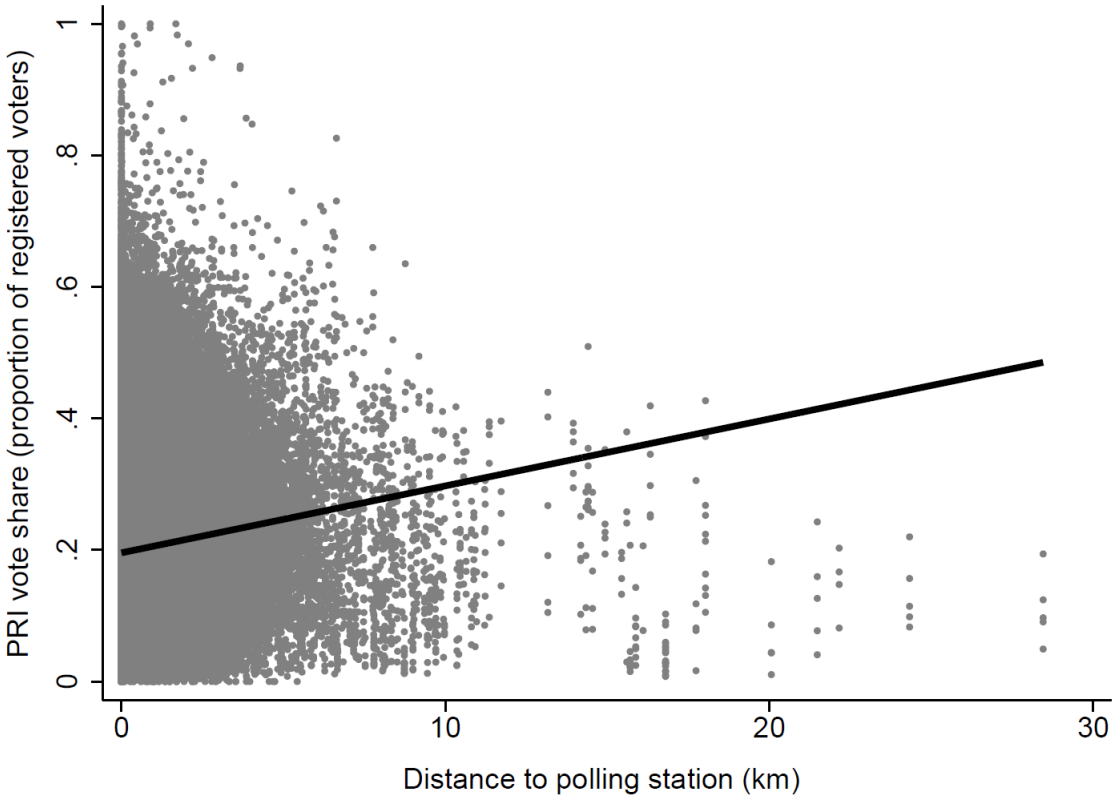

Figure 5: Turnout by average distance of voters to polling station

Note: The black line is the best linear fit. Grey dots are individual polling stations.

We can also provide a clear empirical interpretation for $b(d)$, and thus derive differential predictions across parties. Despite relative declines in support, the PRI has generally remained the largest party since 2000, averaging 33% of the national legislative votes over this period. The

[26] If different brokers operate in different polling stations, we could recast the problem as one of "moral hazard in teams". In that case, polling station level data is a powerful tool for ascertaining relative performance (Holmstrom 1982). Given voters are assigned to polling stations by surname (see below), this type of separation is very unlikely.

PAN has obtained just below 30% of legislative votes. Therefore, on average across regions and elections, $b(d) > 0$ holds in races between the PRI and the more right-wing PAN, while the PRI similarly enjoys an advantage against the PRD even though the left-right comparison is less appropriate. The fourth part of Proposition 2 thus implies that we should expect greatest turnout buying by the PRI:

H2. Improved monitoring capacity increases the number of PRI votes more than PAN or PRD votes on average across types of polling stations.

Beyond the logic of the model, the PRD are less likely to engage in the types of illegal electoral practices associated with turnout buying given they have regularly denounced corruption to distance themselves from the PRI.

Mexican politics is also defined by a clear rural-urban divide. While the PAN and PRD are well-supported in more urban areas, among richer voters, and where clientelistic ties are weaker, the PRI continues to win a large proportion of more rural voters (see Larreguy, Marshall and Snyder Jr. 2014). Figure 5 shows this strong relationship, which clearly indicates that $b'(d) > 0$ generally holds (where the PRI can be regarded as *L* and the PAN or PRD as *R*). In some southern areas where the PAN and PRD are the dominant parties, the rural-urban division is less salient; accordingly, $b'(d) \approx 0$ and any interaction between distance and turnout buying should be weaker because the benefits of mobilizing rural voters are lower.[27] Combining these insights with Proposition 2, we only expect a non-linear relationship between turnout buying and distance for the PRI:

H3. Following an improvement in monitoring capacity, the PRI vote share will first increase in distance from the polling station before decreasing. PAN and PRD vote share will start decreasing in distance for a smaller distance, or exhibit no relationship with distance.

[27]Plotting vote share by distance for the PAN and PRD respectively yield slightly negative and flat slopes.

4 Empirical design

This section first describes the data used to test the hypotheses derived above, before explaining how we use a discontinuity in the number of polling stations in an electoral precinct to identify the causal effects of monitoring on turnout buying.

4.1 Data

Mexico's 300 electoral districts are divided in to around 67,000 electoral precincts. These, as will be explained in more detail as part of our identification strategy below, are in turn composed of polling stations. The average precinct contains 1.97 polling stations. The IFE has collected detailed polling station level data since 2000, including the coordinates of polling stations. We use this data to analyze polling station electoral returns for the 2000, 2003, 2006, 2009 and 2012 national legislative elections. Combined, this produces a maximum sample of 561,256 polling stations.[28]

4.1.1 Dependent variables

We use two main measures of voting behavior. We first measure polling station turnout as a proportion of the total number of voters registered at a given polling station. Turnout includes all votes for political parties or coalitions, including null votes and non-registered votes. To measure the beneficiaries of increased turnout, we measure party vote share as the number of votes for the party as a proportion of the total number of registered voters. By not conditioning on turnout, this outcome is independent of the mobilization efforts of other political parties.

[28] In our analysis, we restrict attention to the *casilla básica* and *casilla contigua* polling stations that are relevant for the electoral rule we exploit. Less than 1% of polling stations are special or extraordinary polling stations, which include temporary residents or were created to address challenging sociocultural or geographic circumstances (see Cantu forthcoming). Due to the existence of such polling stations, the electoral rule for splitting polling stations can be violated in certain precincts. We remove all such polling stations.

4.1.2 Independent variables

The central variable used to capture the monitoring potential of political parties over their brokers is the number of polling stations in a given electoral precinct. Polling stations may contain up to 750 registered voters, while the number of registered voters in a precinct will play a central role in determining the number of voters per polling station (see below). Electoral precincts were originally drawn to contain up to 1,500 voters, but demographic changes have caused this number to rise considerably in some precincts.

To test our main heterogeneous effect, we calculated the average distance (in kilometers) of voters to the polling station to which they are assigned. The average distance entailed computing the electorate-weighted distance to the polling stations among the set of registered voters in each electoral precinct using locality-level population data provided by IFE.[29] Since all the polling stations in a given precinct in our sample are located in the same place, and we cannot distinguish the geographic density of voters registered at different polling stations within an electoral precinct, the weighted distance varies by precinct rather than polling station.[30]

4.2 Identification strategy

In order to identify the effects of changes in monitoring incentives on turnout buying, we leverage variation in the size of the electorate at different types of polling station. In particular, we examine a discontinuity in the number of polling stations in an electoral precinct—and thus the capacity of political parties to monitor their brokers—that arises from Mexico's electoral rules.

Each electoral precinct is served by at least one polling station. Once the registered electorate in

[29]Due to the difficulties of matching localities to polling station coordinates, we only use the voter geographic distributions and polling station coordinates for 2012. Although redistricting occurred in 2005, this affected only districts—not precincts. Thus, our distance metric may suffer minor measurement error, but it is hard to see how any biases could be systematic.

[30]This represents a good approximation given the unlikeliness of surnames, which determine the polling station voters are assigned to, being systematically spatially correlated.

the precinct exceeds 750 (or any such multiple) due to demographic changes, an additional polling station is added and voters are reallocated equally between all polling stations. Figure 6 shows this procedure in our data. The new polling station must be located in the same building or an adjacent building, and voters are assigned alphabetically by surname between polling stations. The addition of a new polling station therefore does not affect the distance that voters must travel to vote. Given surname does not predict voter behavior in Mexico (Cantu forthcoming), the assignment of voters to polling stations is random with respect to our voting outcomes.

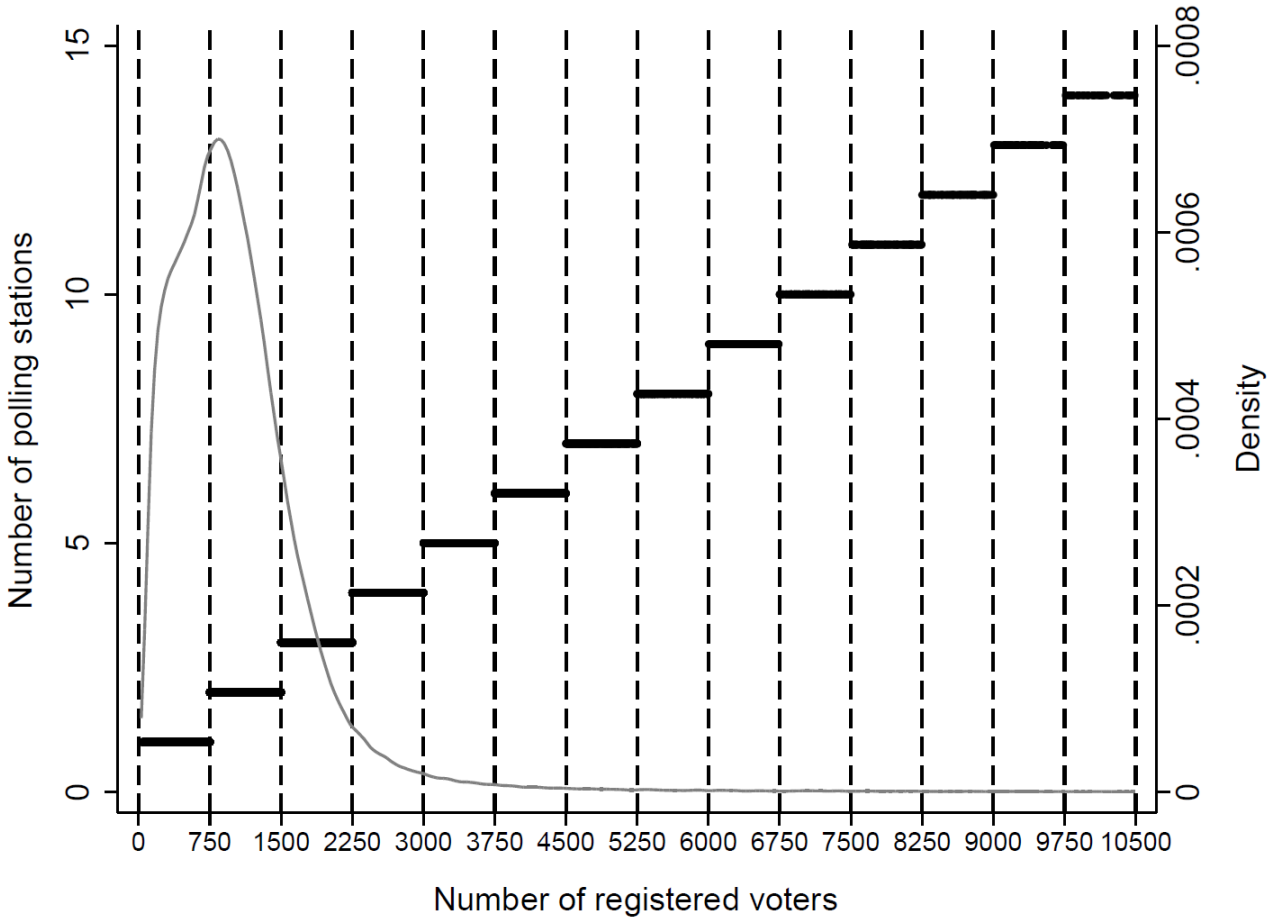

Figure 6: Number of polling stations per electoral precinct

Notes: Points represent electoral precincts. The grey kernel density plot shows the distribution of total registered voter by electoral precinct.

We exploit the electoral rules defining the creation of a new polling station as a source of discontinuous variation in political party monitoring capacity. More specifically, we employ a regression discontinuity (RD) design to compare polling station returns in electoral precincts that just exceeded the threshold required to split into more polling stations with fewer voters to polling stations in precincts that fell just below the threshold. The "running variable" determining whether a precinct is treated with a new polling station is the number of registered voters in the electoral precinct.

Given new polling stations are created at each multiple of 750 voters, there exist many discontinuities (at 750, 1,500, 2,250 etc. voters). We pool all discontinuities by redefining the running variable as the deviation from the nearest multiple of 750 registered voters in a given precinct, which ranges from -375 to 375.[31] Our treatment indicator for an additional polling station in electoral precinct j at time t is defined by:

$$split_{jt} \equiv 1(registered\ voters\ deviation_{jt} > 0). \tag{5}$$

To be clear, our source of variation comes from the precinct level.

It is important to note that we are averaging over discontinuities that could produce different effects by virtue of their differential impact on monitoring. In particular, the change in the number of voters per polling station is smaller at the discontinuities in larger electoral precincts. If polling stations with fewer voters communicate more information (e.g. Gingerich and Medina 2013),[32] an additional signal of broker performance has diminishing monitoring value to political parties. We analyze each discontinuity separately as a robustness check, and show that the results are primarily driven by the largest changes in the smaller precincts that constitute the vast majority of the sample.

[31] The smallest electoral precincts containing less than 375 voters were assigned scores -750 to -376.

[32] Gingerich and Medina (2013) show that for a sufficiently large number of registered voters, the performance signal becomes sufficiently noisy that it becomes suboptimal for brokers to engage in outcome-contingent contracts.

The RD framework allows us to identify the local average treatment effect of an additional polling station under relatively weak assumptions. In particular, identification of causal effects at the discontinuity requires that potential outcomes are continuous across the discontinuity such that as we approach the discontinuity precincts that were not split are effectively identical to those that were split (see Imbens and Lemieux 2008). We now verify the validity of this assumption.

4.2.1 Validity of the RD design

A key concern with any RD design is the possibility of sorting around the discontinuity. In our case, this could occur if electorate sizes or precinct boundaries are subject to political manipulation. The kernel density plot in Figure 6 strongly suggests that there is no systematic difference in the density of precincts at any multiple of 750 voters in the electoral precinct. Confirming this, a McCrary (2008) density test similarly fails to reject the null hypothesis of equal density either side of each discontinuity.[33]

Given the absence of bunching at either side of the discontinuity, it is hard to imagine that electoral precincts with just above 750 registered voters systematically differ from those with just below 751. Nevertheless, it is important to demonstrate that other variables are continuous at the discontinuity. We compare precincts within 20 voters of being split to precincts that exceeded a multiple of 750 by less than 20 voters using contemporaneous IFE electoral data and precinct-level variables from the 2010 Census.[34] Table 1 examines 26 observable political, economic and demographic characteristics of these polling stations, and shows that treated and control units are

[33] At each of the first six discontinuities, where the vast majority of our data is, we cannot reject the possibility that the density of electoral precincts is identical either side of the discontinuity. We used the natural unit bin size and a bandwidth of five voters. The Online Appendix provides histograms showing no change in density around these discontinuities.

[34] Although it is not quite necessary, achieving balance across treated and control units is sufficient for potential outcomes to be continuous in all variables other than our treatment at the discontinuity (Imbens and Lemieux 2008). Intuitively, this is because balance across treated and control units implies continuity as the bandwidth goes to zero.

Table 1: Balance checks—correlation between splitting polling stations and political, economic and demographic characteristics of electoral precincts

	(1) PRI municipal incumbent	(2) PAN municipal incumbent	(3) PRD municipal incumbent	(4) PRI Governor	(5) Area (log)	(6) Voter density (log)	(7) Distance to polling station	(8) Share economically active	(9) Share employed
Split	-0.0018	0.0082	-0.0091*	-0.0058	-0.0269	-0.0245	-0.0126	-0.0025***	-0.0004
	(0.0080)	(0.0060)	(0.0050)	(0.0038)	(0.0272)	(0.041)	(0.014)	(0.0008)	(0.0004)
Observations	27,322	27,322	27,322	27,697	24,199	24,199	27,419	27,457	27,457

	(10) Share medical insurance	(11) Share illiterate	(12) No schooling above 15	(13) Incomplete primary school	(14) Complete primary school	(15) Incomplete secondary school	(16) Complete secondary school	(17) Share owns house	(18) Share basic amenities
Split	-0.0011	0.0001	0.0005	-0.0002	-0.0017	-0.0018	-0.0017	-0.001	-0.0005
	(0.0015)	(0.0006)	(0.0005)	(0.0029)	(0.0022)	(0.0016)	(0.0015)	(0.0014)	(0.004)
Observations	27,457	27,457	27,457	27,457	27,457	27,457	27,457	27,462	27,462

	(19) Share with radio	(20) Share with TV	(21) Share with fridge	(22) Share washing machine	(23) Share with car	(24) Share with telephone	(25) Share cell phone	(26) Share with internet
Split	-0.0031**	-0.0005	-0.0013	-0.0026	-0.0028	-0.0016	-0.0063*	-0.0003
	(0.0015)	(0.0015)	(0.0022)	(0.0026)	(0.0026)	(0.0027)	(0.0034)	(0.0028)
Observations	27,462	27,462	27,462	27,462	27,462	27,462	27,462	27,462

Notes: Each coefficient is estimated separately from an OLS regression including district and year fixed effects. Block-bootstrapped standard errors are clustered by district. * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

well balanced across these variables.[35] Particularly important is the fact that we find no correlation with political variables like incumbency that could plausibly have indicated sorting via political manipulation.

4.2.2 Estimation

We first estimate the causal effect of an additional polling station in an electoral precinct on polling station-level turnout and party vote share. Our non-parametric RD approach includes observations within a bandwidth of 20 voters either side of the discontinuity.[36] Accordingly, we estimate the following simple equation using OLS:

$$Y_{ijdst} = \beta split_{jt} + \mu_d + \xi_t + \varepsilon_{ijdst}, \tag{6}$$

where Y_{ijdst} is one of our two outcome variables at polling station i, and μ_d and ξ_t are district and election fixed effects. By focusing on within-district and within-election variation, we ensure that our results are not driven by national or district-specific election characteristics. Throughout we conservatively cluster standard errors by state s, computed using a block bootstrap based on 1,000 resamples. Although the choice of bandwidth is inevitably somewhat arbitrary, we will show that our results are insensitive to the choice of bandwidth.

To test the distance-specific predictions of our broker monitoring theory, we also estimate the following interaction specifications:

$$Y_{ijdst} = \beta split_{jt} + \sum_{k=1}^{2} \tau_{0k} distance_j^k + \sum_{k=1}^{2} \tau_{1k} \left(distance_j^k \times split_{jt} \right) + \mu_d + \xi_t + \varepsilon_{ijdst}, \tag{7}$$

[35] Plotting these variables as functions of our running variable similarly supports continuity across the discontinuity.

[36] Although there is little trend in our outcomes at the discontinuity to worry about in terms of the "boundary problem" of estimating means either side of the discontinuity, we verify our results equally apply when using a local linear regression approach with a rectangular kernel and including trend terms either side of the discontinuity as a robustness check.

where $distance_j^k$ measures the average distance of voters to the polling station(s) in precinct j. To demonstrate that distance is not simply proxying for another variable, we also control for other interactions at the discontinuity as a robustness check.

5 Results

Our results provide strong support for the theoretical model. We first show a jump in turnout and PRI and PAN vote shares at the discontinuity on average across the sample. Distinguishing our monitoring mechanism from potentially confounding explanations, we then show that the effect of an additional polling station increases and then decreases with distance—but only for the PRI. Finally, we show that our results are highly robust across a wide range of alternative specifications.

5.1 Average effects of an additional polling station

Before estimating equation (6), we first depict our variation graphically. Panel A of Figure 7 pools all discontinuities around multiples of 750 registered voters together, and shows a jump in turnout of nearly one percentage point once a polling station is split. Panels B-D examine party vote share, and suggest that the PRI—traditionally Mexico's most clientelistic and fraudulent political party—and, to a lesser extent, the PAN—which held the Presidency between 2000 and 2012—are the principal beneficiaries of adding a new polling station. Conversely, there is little evidence that the PRD experienced a change in their vote share. We now test these relationships more formally.

The RD regression results in Table 2 support the changes identified in Figure 7. Column (1) shows that, on average, splitting a polling station increases polling station turnout by 0.8 percentage points. This represents nearly a two percent increase in the proportion turning out to vote. This finding is consistent with the monitoring argument (in H1) that parties are better able to monitor their brokers in electoral precincts with more polling stations.

Columns (2)-(4) examine changes in vote share by party, and reiterate that the PRI and PAN

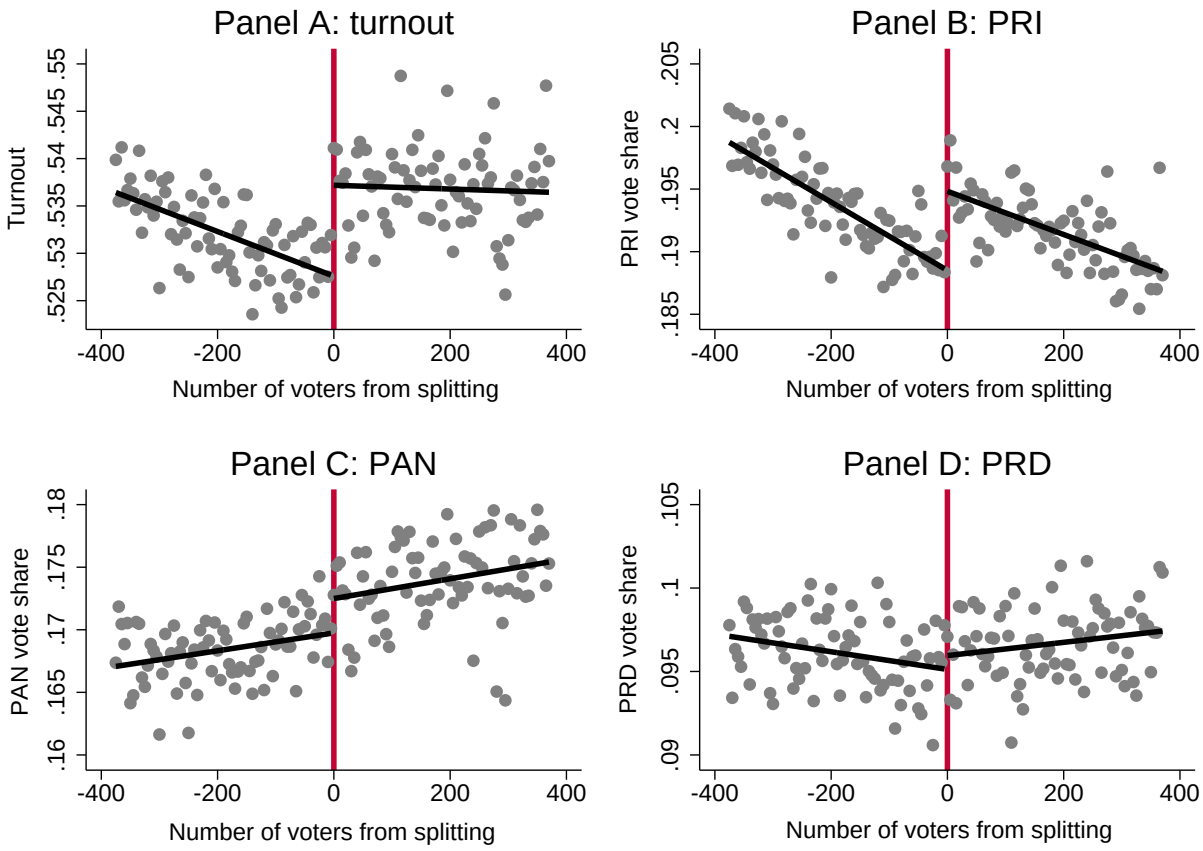

Figure 7: The effect of splitting polling stations on turnout and party vote share

Notes: Points in each graph represent the mean outcome for bins of registered voters of size five. The black line is the best linear fit either side of the discontinuity.

are the main beneficiaries, respectively increasing their vote shares by 0.5 and 0.4 percentage points—this represents 2.5 percent, or almost a standard deviation, increase in their vote shares. The estimate in column (4), which is precisely estimated, confirms that the PRD did not significantly increase their vote share. Consistent with H2, the results imply that turnout buying is most prevalent by the PRI, although the PAN is also engaged in significant turnout buying.

Our estimates imply that the methods of the PAN and PRI are relatively effective. Using the discontinuity as an instrument for turnout, unreported 2SLS estimates reiterate that the PRI vote share increases by 0.62 percentage points for every 1 percentage point increase in overall turnout

Table 2: Effect of split polling station on voting behavior by polling station

	Turnout (1)	PRI (2)	PAN (3)	PRD (4)
Split	0.0079***	0.0049***	0.0040**	-0.0002
	(0.0014)	(0.0010)	(0.0016)	(0.0008)
Observations	27,699	27,699	27,699	27,699
Outcome mean	0.54	0.19	0.17	0.10
Outcome standard deviation	0.14	0.08	0.11	0.09

Notes: All specifications include district and year fixed effects, and estimated using OLS. 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 $p < 0.01$.

while the PAN vote share increases by 0.51 percentage points. Given PRI and PAN turnout buying may be occurring simultaneously in some precincts, these estimates are almost certainly lower bounds on the efficacy of turnout buying's mapping to vote share. While this finding clearly fits with newspaper accounts of the election, it considerably exceeds survey estimates of the proportion of voters who reported that a gift caused them to change their vote (e.g. Gasca 2013, Nichter and Palmer-Rubin forthcoming). Our results therefore suggest that survey measures suffer from considerable social desirability bias.

Although our identification strategy does not permit causal claims away from the discontinuity, Table 3 explores the correlation between the number of polling stations in an electoral precinct and voting in the full sample. The results paint a similar picture, with each additional polling station increasing turnout and support for the PRI. These estimates for an additional polling station are fairly similar to the RD estimates, and support the validity of our estimates away from the discontinuity. Consistent with the qualitative accounts and H2, the results suggest that PRI turnout buying is significantly more prevalent than PAN turnout buying on average across the country.

Table 3: Correlation between number of polling stations and voting behavior in the full sample

	Turnout (1)	PRI (2)	PAN (3)	PRD (4)
Number of polling stations	0.0108***	0.0084***	0.0011	0.0008
	(0.0012)	(0.0017)	(0.0009)	(0.0006)
Registered voters in electoral precinct (1000s)	-0.025***	-0.017***	-0.005***	-0.001*
	(0.002)	(0.003)	(0.001)	(0.001)
Observations	561,256	561,256	561,256	561,256
Outcome mean	0.54	0.20	0.17	0.10
Outcome standard deviation	0.14	0.09	0.11	0.09

Notes: All specifications include district and year fixed effects, and estimated using OLS. Block-bootstrapped standard errors are clustered by district. * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

However, monitoring is not the only explanation that could account for the increase in turnout observed at the discontinuity. Although less likely, especially since there is no change in PRD vote share, increased turnout could reflect an increased incentive to turn out if the expected duration of queuing declines, or if electoral administration improves. Even if monitoring explains the discontinuous change in voting, our results could be capturing vote buying rather than turnout buying. To differentiate our theory, we test the heterogeneous effects by distance predicted by the model.

5.2 Heterogeneous effects of additional polling stations

Table 4 reports estimates of equation (7), and provides clear support for non-linear effects of monitoring capacity in line with H3. The results show that splitting a polling station increases the PRI vote share, but particularly in locations where voters are not too close but not too far from the polling station. Figure 8 illustrates our results graphically, showing that the effect of an additional polling station is maximized where the average voter lives around 1.75km from the polling sta-

Table 4: Effect of split polling station by distance

	Turnout (1)	PRI (2)	PAN (3)	PRD (4)
Split	0.0078***	0.0035***	0.0042***	0.0002
	(0.0015)	(0.0011)	(0.0017)	(0.0010)
Distance	-0.0084*	0.0066*	-0.0110***	-0.0025
	(0.0049)	(0.0035)	(0.0042)	(0.0019)
Distance squared	-0.0001	-0.0007	0.0005	0.0001
	(0.0008)	(0.0007)	(0.0009)	(0.0003)
Split × distance	0.0037	0.0105***	-0.0007	-0.0000
	(0.0047)	(0.0034)	(0.0043)	(0.0027)
Split × distance squared	-0.0011	-0.0030***	0.0007	-0.0004
	(0.0013)	(0.0012)	(0.0010)	(0.0005)
Observations	27,419	27,419	27,419	27,419

Notes: All specifications include district and year fixed effects, and 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 $p < 0.01$.

tion.[37] Around 96% of the RD sample is located on the increasing side of the curve, indicating that the electoral benefits of mobilizing voters for the PRI generally exceed the costs. The effect of increased monitoring capacity above the discontinuity is significantly larger than the effect just below the discontinuity for c.99% of the RD sample.

By adhering closely to our theoretical expectations, our results for the PRI lend strong support to the claim that the average effect at the discontinuity is picking up differences in monitoring capacity rather than any other change associated with new polling stations. Conversely, although the PAN benefits from splitting polling stations where the average distance to the polling station

[37] We also defined indicators for different distances and found similar results, suggesting that the monitoring effect is largest when the distance from the polling station is 2km on average across votes.

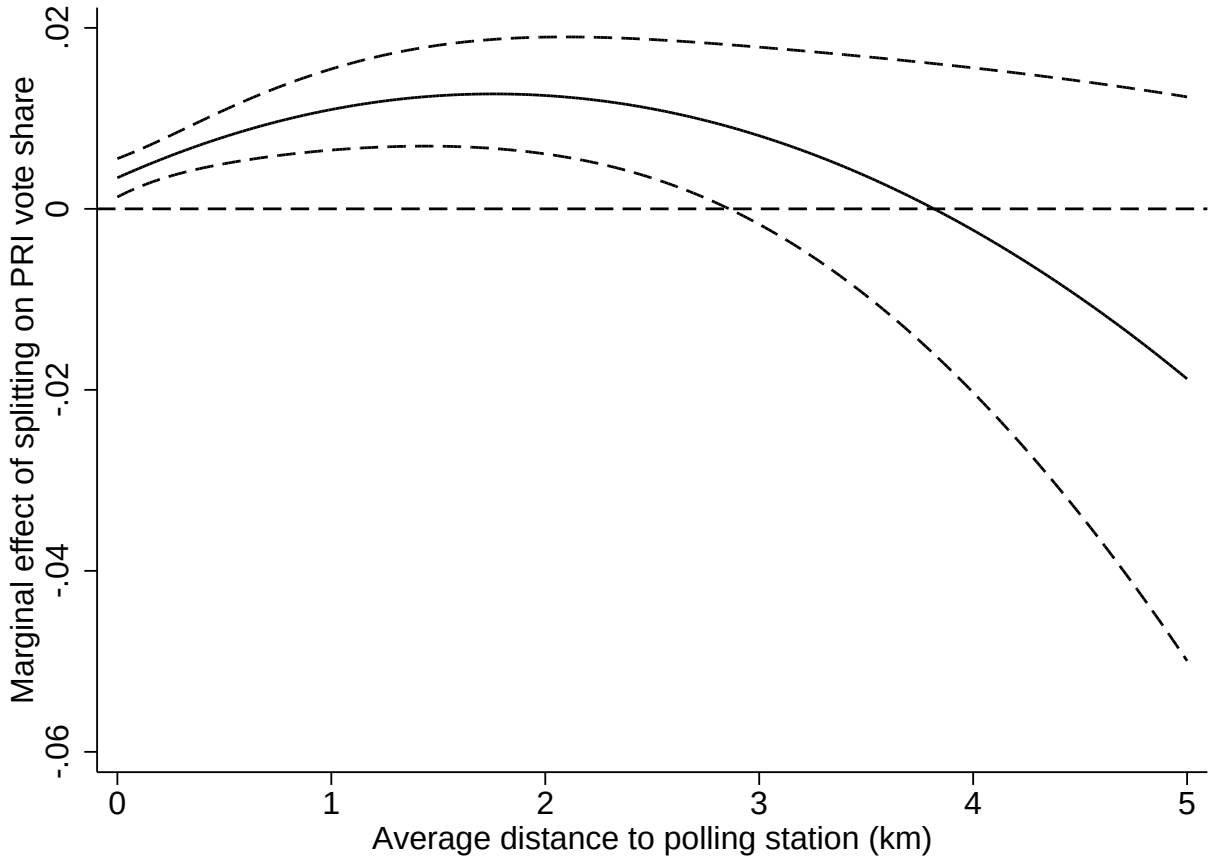

Figure 8: Marginal effect of splitting polling stations on PRI vote share by distance (with 95% confidence interval)

is small, column (3) shows that the PAN vote share declines with distance at the same rate either side of the discontinuity. This fits exactly with the differential predictions of our model across parties, suggesting that the lack of potential PAN voters living further from the polling station cannot support the cost of hiring brokers to reach these areas. Furthermore, column (4) shows no differential effects for the PRD, which does not appear engage in significant turnout buying of the type examined here.

Our heterogeneous effects are not consistent with the most plausible alternative explanations for the effects of an additional polling station. First, if electoral administration reduces the costs

of turning out, we should not expect to observe any interaction with distance unless those worried about queues are disproportionately living further away from the polling station. Even if this were the case, this explanation cannot explain why the relationship holds only for the PRI, or why the effect with distance is non-linear. Second, if monitoring was facilitating vote buying then we should expect the relationship with distance to reverse because there are more potential voters available to the PRI in the less rural areas where there are more PAN and PRD voters.

5.3 Robustness checks

Although we demonstrated that split polling stations occur effectively randomly, there remain several important robustness checks to undertake. First, we show that our results do not depend on the RD specification choices, while the effect continues to hold as we move away from the discontinuity. Second, we show that the interaction with distance at the discontinuity is not confounded by omitted variables. All results are available in the Online Appendix.

To demonstrate that our results are not specification-dependent, we employ a variety of sensitivity analyses. First, Figure 9 considers bandwidths in multiples of five ranging from 5 to 100. The results show that the average effect at the discontinuity is very similar across bandwidth choices, and always statistically significant for turnout and PRI vote share. Consistent with H2, the smaller effects for the PAN are less robust.

Second, the results are robust to the inclusion of linear trends in the running variable on either side of the discontinuity. That this local linear approach yields essentially identical results is not surprising given that our bandwidth choice produces a relatively trendless sample around the discontinuity. This demonstrates that our results are not susceptible to the "boundary problem" that can confound non-parametric RD estimation.

Third, our pooling of different precinct discontinuities masks variation by discontinuity. Specifically, the change at the discontinuity is largest for split polling stations when there are 750 and 1,500 voters in the electoral precinct. For the larger discontinuities, the effect is smaller (as well

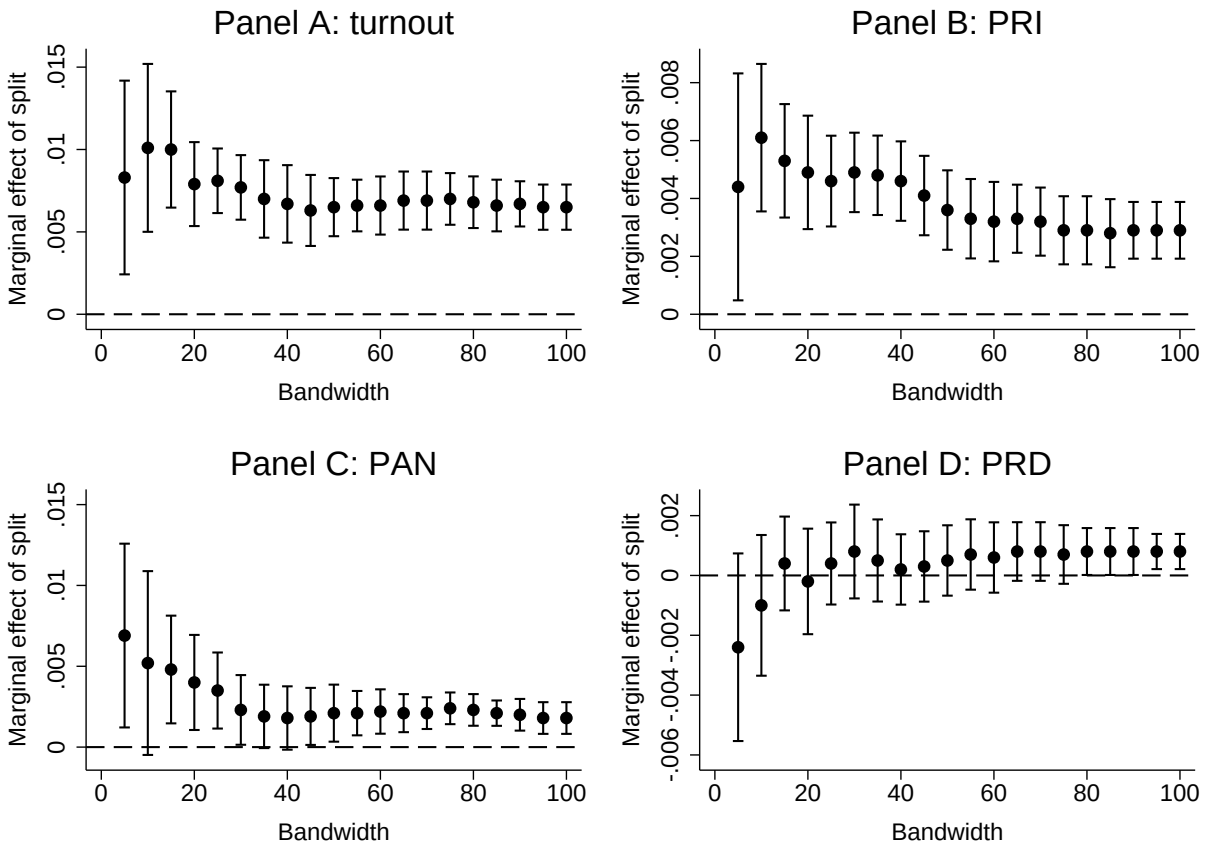

Figure 9: Marginal effect of splitting polling stations by bandwidth choice (with 95% confidence interval)

Notes: All estimates are from specification equivalent to those in Table 2, with the exception that the bandwidth varies from 5 to 100 voters (in five voter intervals) either side of the discontinuity.

as less precisely estimated due to the smaller samples). Supporting Gingerich and Medina (2013), these differences suggest that the largest changes in turnout occur where the change in monitoring capacity is largest. Given larger electoral precincts may differ for other reasons, we treat this observation with caution.

The second main concern is that the interaction with distance to the polling station is actually proxying for another variable. A particular worry is that in using distance, we are simply picking up characteristics like poverty that are correlated with support for the PRI. Although such arguments

are unlikely to entail a non-linear interaction, we included all the variables that we demonstrated balance on as quadratic interactions and find that the quadratic effects of monitoring by distance to the polling station on the PRI vote share are remarkably stable. In no specification were the distance interaction terms statistically insignificant.[38] To further dismiss the concern that spatial correlations with distance are driving our results, we show that the results are similarly robust to the inclusion district-specific discontinuities. Furthermore, to allow for flexible interactions with time, we also show that our results are robust to the inclusion of district-year fixed effects.

6 Conclusion

While electoral clientelism is prevalent in nascent and consolidating democracies, little is known about the conditions under which such electoral strategies can be effectively executed. In particular, the political brokers hired by political parties to deliver votes, using their detailed local knowledge, have received surprising limited attention given that they represent the essential intermediaries linking parties to voters. Emphasizing the moral hazard problem faced by parties seeking to monitor the brokers they hire to procure such votes, this paper provides a new theoretical explanation for variation in the effectiveness of turnout buying.

We find strong support for the central predictions of our theoretical model in Mexico, showing that an additional polling station—which provides parties with valuable information about broker performance—increases the number of PAN and especially PRI votes in that electoral precinct. Furthermore, such votes vary systematically with distance: consistent with our broker-based theory of turnout buying, we find that the increased vote share of the PRI—which is more popular in rural areas—is initially greater in areas where voters live further from the polling station but ultimately subsides as the costs of incentivizing brokers to mobilize voters with high costs of turning out to vote become too high. As expected, the PAN, which does better in more urban areas, does not

[38] Unsurprisingly, the results are similarly robust to the inclusion of just linear interactions.

exhibit such a relationship.

Our evidence that turnout buying systematically varies with monitoring capacity in Mexico has clear implications for countries across Africa, East Asia and Latin America where "party machines" remain strong. Contrasting with most existing research, our analysis points to the importance of the party-broker relationship. Where such a relationship cannot be sustained, mobilizing votes becomes substantially more challenging for clientelistic parties like the PRI or Argentina's Peronists. While our identification strategy provides strong causal evidence for our theory with respect to the number of polling stations in an electoral precinct, we only capture one way in which parties monitor their brokers. That we still identify moderately sized effects suggests our findings represent the tip of an iceberg, and thus there is considerable scope for further exploring this principal-agent problem both theoretically and empirically.

Our findings also have important implications for policy-makers. The clear policy recommendation is that turnout buying can be reduced by scrambling the information available to political parties about broker performance. In particular, providing less disaggregated electoral information—which could be achieved by reducing the number of polling stations—can inhibit turnout buying. The relative ease with which such measures can be implemented differs from prominent recent accounts emphasizing the role of slow-changing structural factors like economic development, political competition and income inequality (e.g. Stokes et al. 2013; Weitz-Shapiro 2012).

Paradoxically, however, the disaggregated data used by political parties to monitor their brokers is also used by election observers to more effectively monitor electoral fraud. Furthermore, reducing the number of polling stations may conflict with attempts to increase political participation by expanding the number of polling stations to reduce the costs of voting. An important outstanding question is thus to evaluate the relative costs and benefits of empowering voters and election observers while simultaneously providing valuable resources for the electoral machines of clientelistic political parties. This question remains highly relevant as countries continue to adopt

such electoral reforms. El Salvador, for example, recently increased its number of polling stations for the 2014 Presidential election.

References

Auyero, Javier. 2000. *Poor People's Politics: Peronist Survival Networks and the Legacy of Evita*. Duke University Press.

Brady, Henry E. and John E. McNulty. 2011. "Turning out to vote: The costs of finding and getting to the polling place." *American Political Science Review* 105(01):115–134.

Brennan, Geoffrey and Alan Hamlin. 1998. "Expressive voting and electoral equilibrium." *Public Choice* 95(1-2):149–175.

Brennan, Geoffrey and Alan Hamlin. 2000. *Democratic Devices and Desires*. Cambridge University Press.

Brusco, Valeria, Marcelo Nazareno and Susan C. Stokes. 2004. "Vote Buying in Argentina." *Latin American Research Review* 39(2):66–88.

Cornelius, Wayne A. 2004. Mobilized voting in the 2000 elections: The changing efficacy of vote buying and coercion in Mexican electoral politics. In *Mexico's Pivotal Democratic Elections: Candidates, Voters, and the Presidential Campaign of 2000*, ed. Jorge I. Dominguez and Chappell Lawson. Stanford University Press pp. 47–65.

Diaz-Cayeros, Alberto, Federico Estévez and Beatrix Magaloni. forthcoming. *Strategies of Vote Buying: Democracy, Clientelism, and Poverty Relief in Mexico*. New York: Cambridge University Press.

Dunning, Thad and Susan C. Stokes. 2008. Clientelism as Persuasion and as Mobilization. In *American Political Science Association annual meeting*.

Finan, Frederico and Laura Schechter. 2012. "Vote-Buying and Reciprocity." *Econometrica* 80(2):863–881.

Gans-Morse, Jordan, Sebastian Mazzuca and Simeon Nichter. 2013. "Varieties of clientelism: Machine politics during elections." *American Journal of Political Science* 58(2):415–432.

Gasca, Lauro Mercado. 2013. Clientelismo electoral: compra, coacción y otros mecanismos que limitan el voto libre y secreto en México. In *Fortalezas y Debilidades del Sistema Electoral Mexicano (2000-2012): Ideas para una nueva reforma electoral, la perspectiva estatal e internacional*, ed. Luis Carlos Ugalde and Gustavo Rivera Loret de Mola. Available here pp. 236–275.

Gimpel, James G. and Jason E. Schuknecht. 2003. "Political participation and the accessibility of the ballot box." *Political Geography* 22(5):471–488.

Gingerich, Daniel W. and Luis Fernando Medina. 2013. "The Endurance and Eclipse of the Controlled Vote: A Formal Model of Vote Brokerage Under the Secret Ballot." *Economics and Politics* 25(3):453–480.

Gonzalez-Ocantos, Ezequiel, Chad Kiewiet De Jonge, Carlos Meléndez, Javier Osorio and David W. Nickerson. 2012. "Vote buying and social desirability bias: Experimental evidence from Nicaragua." *American Journal of Political Science* 56(1):202–217.

Hicken, Allen. 2011. "Clientelism." *Annual Review of Political Science* 14:289–310.

Holmstrom, Bengt. 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13(2):324–340.

Imbens, Guido W and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142(2):615–635.

Kitschelt, Herbert and Steven I. Wilkinson. 2007. Citizen-Politician Linkages: An Introduction. In *Patrons, Clients, and Policies: Patterns of Democratic Accountability and Political Competition*. Cambridge: Cambridge University Press pp. 1–49.

Larreguy, Horacio. 2013. "Monitoring Political Brokers: Evidence from Clientelistic Networks in Mexico." Working paper.

Larreguy, Horacio, John Marshall and James M. Snyder Jr. 2014. "Political advertising in consolidating democracies: Radio ads, clientelism, and political development in Mexico.".

Levitsky, Steven. 2014. "Brokers beyond Clientelism: A New Perspective on Brokerage through the Argentine Case.".

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico*. New York: Cambridge University Press.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142(2):698–714.

Nichter, Simeon. 2008. "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot." *American Political Science Review* 102(1):19–31.

Nichter, Simeon and Brian Palmer-Rubin. forthcoming. Clientelism, Declared Support and Mexico's 2012 Campaign. In *Mexico's Evolving Democracy: A Comparative Study of the 2012 Elections*, ed. Jorge I. Dominguez, Kenneth G. Greene, Chappell Lawson and Alejandro Moreno. John Hopkins University Press.

Palfrey, Thomas R. and Howard Rosenthal. 1985. "Voter participation and strategic uncertainty." *The American Political Science Review* 79(1):62–78.

Robinson, James A. and Thierry Verdier. 2013. "The Political Economy of Clientelism." *Scandinavian Journal of Economics* 115(2):260–291.

Rueda, Miguel R. 2013. "Election Aggregates and the Choice of Electoral Manipulation Strategies." Working paper.

Smith, Alastair and Bruce Bueno de Mesquita. 2012. "Contingent prize allocation and pivotal voting." *British Journal of Political Science* 42(2):371–392.

Stokes, Susan C. 2005. "Perverse accountability: A formal model of machine politics with evidence from Argentina." *American Political Science Review* 99(3):315–325.

Stokes, Susan C., Thad Dunning, Marcelo Nazareno and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. Cambridge University Press.

Szwarcberg, Mariela. 2012*a*. "Revisiting clientelism: A network analysis of problem-solving networks in Argentina." *Social Networks* 34(2):230–240.

Szwarcberg, Mariela. 2012*b*. "Uncertainty, Political Clientelism, and Voter Turnout in Latin America: Why Parties Conduct Rallies in Argentina." *Comparative Politics* 45(1):88–106.

Ugalde, Luis Carlos and Gustavo Rivera Loret de Mola. 2013. Clientelismo electoral: coacciòn y compra del voto en México. In *Fortalezas y Debilidades del Sistema Electoral Mexicano (2000-2012): Ideas para una nueva reforma electoral*, ed. Centro de Estudios Espinosa Yglesias. pp. 35–40.

Vicente, Pedro C. and Leonard Wantchekon. 2009. "Clientelism and Vote Buying: Lessons from Field Experiments in African Elections." *Oxford Review of Economic Policy* 25(2):292–305.

Weitz-Shapiro, Rebecca. 2012. "What wins votes: Why some politicians opt out of clientelism." *American Journal of Political Science* 56(3):568–583.

Appendix

Proofs

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

$$ph'(e_i^*)P_i(d) - C_e(d, e_i^*) = 0. \tag{8}$$

The concavity of h and convexity of C in e ensure the solution is a maxima. If $ph'(e_i^*)P_i(d) - C_e(d, e_i^*) < 0$, then $e_i^* = 0$. The binding IC constraint then determines the optimal wage w_i^*. By virtue of satisfying the IC constraint, the broker optimally chooses $e_b^*(i) = e_i^*$. ■

Proof of Proposition 2: The results are proved using the implicit function theorem. The first result follows from:

$$\frac{\mathrm{d}e_i^*}{\mathrm{d}p} = \frac{h'(e_i^*)P_i(d)}{C_{ee}(d, e_i^*) - ph''(e_i^*)P_i(d)} > 0, \tag{9}$$

which holds by virtue of $h' > 0$ and the concavity of h and convexity of C in e. Because votes (and thus turnout) are an increasing function of effort, we can apply the chain rule to yield the stated results.

The second result follows from the following distance derivative:

$$\frac{\mathrm{d}e_i^*}{\mathrm{d}d} = \frac{ph'(e_i^*)P_i'(d) - C_{de}(d, e_i^*)}{C_{ee}(d, e_i^*) - ph''(e_i^*)P_i(d)}. \tag{10}$$

The denominator is clearly positive (from above). When $P_i'(d) < 0$ (and thus $B(d, e_i^*) < 0$), e_i^* is monotonically decreasing in d. However, there always exists a $d > 0$ where $P_i'(d) > 0$ for L; such a $d > 0$ may also exists for R. When $P_i'(d) > 0$, the numerator's sign is ambiguous. Let $\overline{d}_i$ define: $ph'(e_i^*)P_i'(\overline{d}_i) = C_{de}(\overline{d}_i, e_i^*)$. Given $\lim_{d\to} B(d, e_i^*)$, such a $\overline{d}_i$ exists. This $\overline{d}_i$ is unique since $P'(d)$ and $C_{de}(d, e_i^*)$ cross at most once as they change with d when $D_B(d, e_i^*) < 0$.

The cross-partial with p shows when this marginal effect is larger for higher monitoring capacity:

$$\frac{\mathrm{d}^2 e_i^*}{\mathrm{d}d\mathrm{d}p} = \frac{C_{ee}(d, e_i^*)P_i'(d)h'(e_i^*) - C_{de}(d, e_i^*)P_i(d)h''(e_i^*)}{[C_{ee}(d, e_i^*) - ph''(e_i^*)P_i(d)]^2}. \tag{11}$$

$P_i'(d) > 0$ is a sufficient condition for $\frac{\mathrm{d}^2 e_i^*}{\mathrm{d}d\mathrm{d}p} > 0$; this always holds for L, but is ambiguous for R. As above, these implications for e_i^* carry over to Π_i^*, albeit now with a level shifter for $\overline{d}_i$ due to the additional effect $e_i^* P_i'(d)$ for $\frac{\mathrm{d}e_i^*}{\mathrm{d}d}$ and $\frac{\partial e_i^*}{\partial p} P_i'(d)$ for $\frac{\mathrm{d}(e_i^*)^2}{\mathrm{d}d\mathrm{d}p}$ which arises from the chain rule.

The last claim follows from inspection of equations (2) and (3), after noting that $P_L'(d) > P_R'(d)$ and, if $b(d) > 0$, $P_L(d) > P_R(d)$. ■

Variable definitions and summary statistics

Our main variables are defined below. Summary statistics are provided in Table 5 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 to polling station. We use the set of IFE localities (which differ from INEGI ones since the IENGI groups people in bigger localities), and calculate the Euclidean distance (in kilometers) of each locality to the polling station. To calculate the average distance, we take sum all those distances weighting by the IFE locality registered population divided by the total registered population in the electoral precinct. This was calculated using 2012 data.

PRI/PAN/PRD municipal incumbent. Indicator coded one if the mayor of the municipality is from the PRI/PAN/PRD.

PRI Governor. Indicator coded one if the Governor of the state comes from the PRI.

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.

No schooling above 15. Percentage of electoral precinct population above 15 with no schooling. 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.

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.

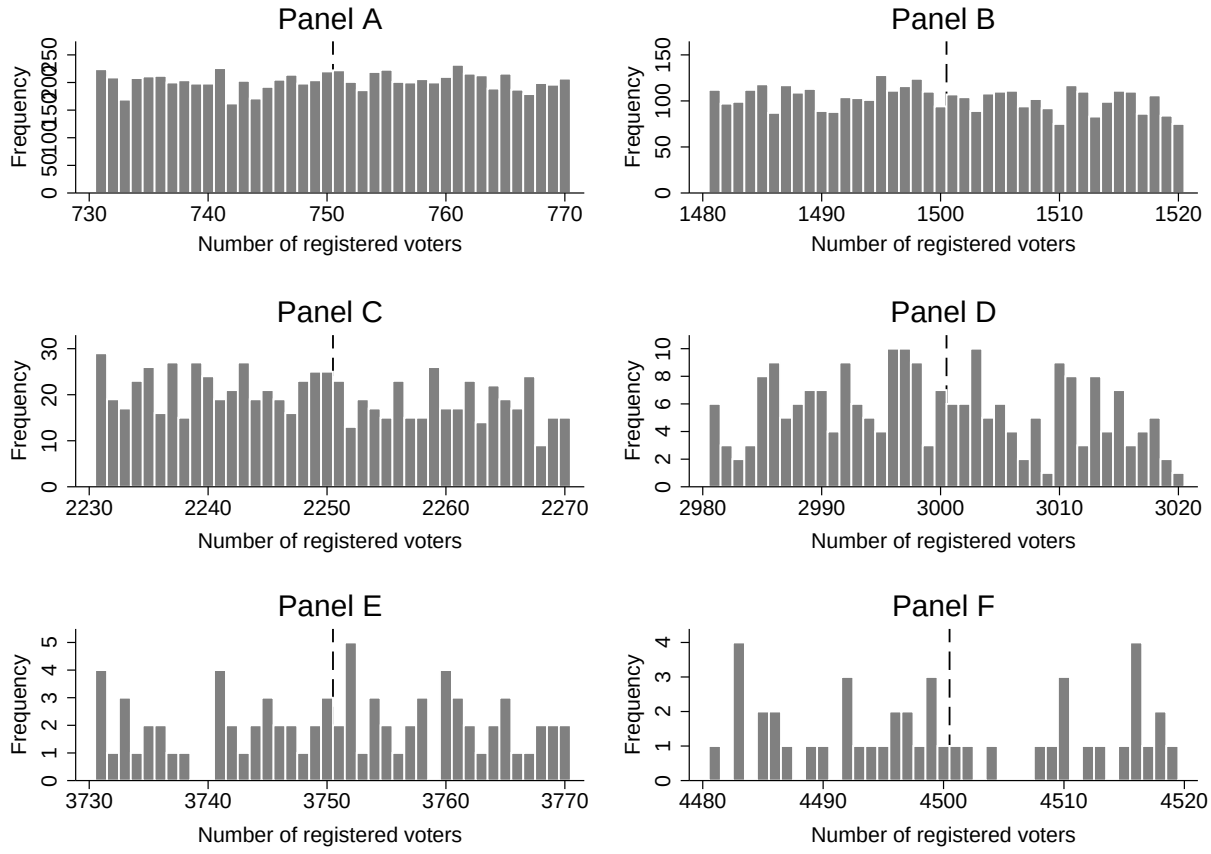

Figure 10: Histograms of the number of electoral precincts either side of the discontinuity

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

Robustness checks

Figure 10 shows that there is no evidence of bunching around the first six discontinuities. Note that there are relatively few electoral precincts with large electorates.

Tables 6-9 report the additional robustness checks cited in the main paper. 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

Table 5: Summary statistics: BSAS and LFS 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 to polling station	27,417	0.22	0.65	0	18.03	512,774	0.22	0.67	0	24.33

and only cease to be statistically significant in the case of the PAN average effect. This increase is not especially surprising given that Figure 7 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 examines how the effect of splitting polling stations varies at different discontinuities. Consistent with theory, the effect generally declines as the marginal addition of polling stations declines. Beyond precincts with around 2,250 voters, the sample size declines considerably. Table 8 show the results for the interactions with distance when controlling for our balancing variables—with the exceptions of distance, population density and area (all of which are highly correlated and capture the same concept). This table also includes the district-discontinuity fixed effects in specification (24). Across all these specifications the results are highly robust and stable. Finally, Table 9 shows that the results are highly robust to the inclusion of district-year fixed effects.

Table 6: Local linear regression estimates

	Turnout (1)	PRI (2)	PAN (3)	PRD (4)
Panel A: Average effects				
Split	0.0117***	0.0070***	0.0049	0.0005
	(0.0033)	(0.0022)	(0.0033)	(0.0014)
Observations	27,699	27,699	27,699	27,699
Panel B: Heterogeneous effects				
Split	0.0118***	0.0053**	0.0052	0.0009
	(0.0032)	(0.0022)	(0.0032)	(0.0014)
Distance	-0.0085*	0.0066*	-0.0110***	-0.0025
	(0.0045)	(0.0034)	(0.0038)	(0.0019)
Distance squared	-0.0001	-0.0007	0.0005	0.0001
	(0.0007)	(0.0007)	(0.0009)	(0.0002)
Split $\times$ distance	0.0037	0.0106***	-0.0007	-0.0000
	(0.0047)	(0.0034)	(0.0043)	(0.0029)
Split $\times$ distance squared	-0.0011	-0.0030**	0.0007	-0.0004
	(0.0013)	(0.0012)	(0.0009)	(0.0006)
Observations	27,419	27,419	27,419	27,419

Notes: All specifications include a linear term in the running variable and its interaction with the split discontinuity, and district and year fixed effects, and is 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: Average effect of splitting polling stations by number of polling stations

	Turnout	PRI	PAN	PRD
	(1)	(2)	(3)	(4)
Panel A: 750-voter discontinuity				
Split	0.0053**	0.0037**	0.0045**	-0.0013
	(0.0021)	(0.0019)	(0.0018)	(0.0015)
Observations	12,226	12,226	12,226	12,226
Panel B: 1,500-voter discontinuity				
Split	0.0047**	0.0018	0.0016	0.0014
	(0.0022)	(0.0017)	(0.0017)	(0.0011)
Observations	10,183	10,183	10,183	10,183
Panel C: 2,250-voter discontinuity				
Split	0.0122**	0.0004	0.0060	0.0016
	(0.0056)	(0.0046)	(0.0052)	(0.0036)
Observations	2,745	2,745	2,745	2,745

Notes: All specifications include a linear term in the running variable and its interaction with the split discontinuity, and district and year fixed effects, and is 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 municipal incumbent	(2) PAN municipal incumbent	(3) PRD municipal incumbent	(4) PRI Governor	(5) Share economically	(6) Share employed active	(7) Share medical insurance	(8) Share illiterate
Split × distance	0.0116***	0.0114***	0.0116***	0.0106***	0.0096***	0.0106***	0.0106***	0.0109***
	(0.0033)	(0.0034)	(0.0034)	(0.0035)	(0.0042)	(0.0034)	(0.0034)	(0.0032)
Split × distance sq.	-0.0031***	-0.0031***	-0.0031***	-0.0030***	-0.0026**	-0.0030**	-0.0030**	-0.0027***
	(0.0011)	(0.0012)	(0.0011)	(0.0012)	(0.0010)	(0.0012)	(0.0012)	(0.0010)
Observations	27,043	27,043	27,043	27,043	27,406	27,406	27,406	27,406

	(9) No schooling above 15	(10) Incomplete primary school	(11) Complete primary school	(12) Incomplete secondary school	(13) Complete secondary school	(14) Share with households	(15) Share basic amenities	(16) Share with radio
Split × distance	0.0116***	0.0096***	0.0100***	0.0099***	0.0097***	0.0099***	0.0109***	0.0109***
	(0.0034)	(0.0035)	(0.0036)	(0.0038)	(0.0032)	(0.0032)	(0.0035)	(0.0032)
Split × distance sq.	-0.0029**	-0.0026**	-0.0029**	-0.0029***	-0.0029***	-0.0030**	-0.0028***	-0.0029***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0013)	(0.0011)	(0.0011)
Observations	27,406	27,406	27,406	27,406	27,406	27,411	27,411	27,411

	(17) Share with TV	(18) Share with fridge	(19) Share washing machine	(20) Share with car	(21) Share with telephone	(22) Share cell phone	(23) Share with internet	(24) District-discontinuity fixed effects
Split × distance	0.0120***	0.0102***	0.0097***	0.0105***	0.0091**	0.0094***	0.0094***	0.0120***
	(0.0036)	(0.0037)	(0.0034)	(0.0038)	(0.0035)	(0.0035)	(0.0033)	(0.0038)
Split × distance sq.	-0.0030***	-0.0028**	-0.0026**	-0.0029**	-0.0026***	-0.0025**	-0.0026**	-0.0031***
	(0.0011)	(0.0012)	(0.0011)	(0.0038)	(0.0011)	(0.0010)	(0.010)	(0.0008)
Observations	27,411	27,411	27,411	27,411	27,411	27,411	27,411	26,703

Notes: Each coefficient is estimated separately from an OLS regression including district and year fixed effects, in addition to quadratic interactions of the variable listed at the top of each column with the split indicator. 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: Inclusion of district-year fixed effects

	Turnout (1)	PRI (2)	PAN (3)	PRD (4)
Panel A: Average effects				
Split	0.0085***	0.0045***	0.0041**	0.0003
	(0.0012)	(0.0012)	(0.0016)	(0.0008)
Observations	27,699	27,699	27,699	27,699
Panel B: Heterogeneous effects				
Split	0.0083***	0.0033***	0.0042**	0.0006
	(0.0015)	(0.0011)	(0.0018)	(0.0012)
Distance	-0.0067	0.0058	-0.0091***	-0.0015
	(0.0047)	(0.0035)	(0.0032)	(0.0021)
Distance squared	-0.0003	-0.0006	0.0004	0.0000
	(0.0007)	(0.0007)	(0.0006)	(0.0003)
Split $\times$ distance	0.0015	0.0118***	-0.0038	-0.0005
	(0.0049)	(0.0039)	(0.0032)	(0.0035)
Split $\times$ distance squared	-0.0010	-0.0032**	0.0010	-0.0005
	(0.0015)	(0.0013)	(0.0006)	(0.0007)
Observations	26,703	26,703	26,703	26,703

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 $p < 0.01$.