Jumping Ship or Jumping on the Bandwagon: When Do Local Politicians Support National Candidates?*

Guillermo Rosas
Department of Political Science
Washington University in St. Louis
grosas@wustl.edu

Adrián Lucardi
Department of Political Science
ITAM
adrian.lucardi@itam.mx

Abstract

Local politicians are often expected to mobilize voters on behalf of copartisan candidates for national office. Yet this requirement is difficult to enforce because the effort of local politicians cannot be easily monitored and the promise of rewards in exchange for help is not fully credible. Using a formal model, we show that the incentives of local politicians to mobilize voters on behalf of their party depend on the proportion of copartisan officials in a district. Having many copartisan officials means that the party is more likely to capture the district, but the effort of each local politician is less likely either to be noticed by higher-level officials or to make a difference on the election outcome, thus discouraging lower-level officials from exerting effort. We validate these claims with data from federal elections in Mexico between 2000 and 2012. In line with the argument, the results show that political parties fail to draw great mobilization advantages from simultaneously controlling multiple offices.

Keywords: party organization, voter mobilization, local politics

Accepted version – forthcoming at *Political Science Research and Methods*

*We are grateful to George Avelino, Dawn Brancati, Mario Chacon, Justin Fox, Ben McClelland, Matias Mednik, Margit Tavits, and the *PSRM* reviewers and editors for very helpful comments. Guillermo Rosas acknowledges support from Washington University's Weidenbaum Center on the Economy, Government, and Public Policy and to the Hertie School of Governance for hosting him during the academic year 2015–2016. Adrián Lucardi thanks the Asociación Mexicana de Cultura A.C. for financial support. To view supplementary material for this article, please visit `http://dx.doi.org/10.1017/psrm.2014.11`.

What are the conditions under which *local politicians* support copartisan candidates running for *national office*? In many countries around the world, local politicians like mayors and governors are expected to campaign and mobilize voters on behalf of copartisan candidates to national office. This seems natural: What else, after all, could local politicians want if not to see their party victorious in all electoral arenas? Why would they ever choose *not* to mobilize if this means lowering the chances of scoring a party victory at the national level?

And yet, anecdotal evidence casts doubt on the notion that local politicians automatically extend a helping hand to copartisan national candidates. In 2006 the left-of-center candidate to the Mexican presidency, Andrés Manuel López Obrador, lost by a razor-thin margin (0.6 percentage points) an election that he was expected to win. The post-mortem analysis revealed that some copartisan governors and mayors had been reluctant to mobilize voters on his behalf in the states and cities they governed.[1] In Costa Rica, the 2014 presidential elections were not concurrent with local elections for the first time in more than a decade, making party leaders realize that, "though lack of local elections implies a lighter workload on election day, it also implies that we need to convince local leaders to put up a fight when they stand to win nothing for themselves."[2] In the 2009 midterm congressional election in Argentina, Néstor Kirchner attempted to win the race in the province of Buenos Aires by packing the Peronist list with the local governor and several mayors, hoping that fear of an embarrassing personal defeat at the local level would induce these politicians to mobilize the Peronist vote in their districts — that is, Kirchner exploited coattail effects precisely because he knew that local politicians would not exert effort to mobilize voters if their own careers

[1] Ricardo Alemán, "Itinerario Político", *El Universal*, April 15, 2007. See also the appeal of Josefina Vázquez Mota, candidate for *Partido Acción Nacional* to the Mexican presidency in 2012, to mayors from her party to promote her candidacy vigorously, "since you are closest to the people" (Mauricio Torres, "Vázquez Mota pide a los alcaldes del PAN promoverla en 'cada rincón'." *CNN México*, June 2, 2012; authors' translation).

[2] Gerardo Ruiz Ramón, "Partidos cubrirán con dirigencia local vacío que dejan elecciones municipales". *La Nación*, September 25, 2013; authors' translation.

were not on the line.[3] The Venezuelan *Mesa de Unión Democrática*, a coalition of parties opposed to the Chávez government, followed a similar rationale in setting primary elections for mayoral candidates to coincide with the 2012 presidential race.[4]

These anecdotes suggest that local politicians can and do shirk from helping copartisan candidates even when they share similar ideological preferences and partisan objectives, and that shirking can take place even in close races where a small increase in effort could have had a substantial effect on the outcome. Yet even if local politicians are eager to see their party victorious, mobilizing voters to increase the party's electoral chances is costly, and since the consequences of these efforts are not easy to ascertain — how to know if voters turned out on election day because they wanted to rather than due to the effort of local party branches? — they may not be compensated.

In other words, intra-party life is plagued with agency problems, yet so far the literature has paid little attention to the incentives that local politicians face to exert effort on behalf of copartisans. In this paper, we propose a formal model about the conditions under which local elected politicians exert effort in favor of copartisan candidates running for national office, particularly when their own re-election is not at stake. We focus on the interaction between two kinds of local politicians — "governors" and "mayors" — based on the following characteristics: First, both mayors and governors are agents of a political party, in the sense that they are expected to exert effort to increase the local vote share of party candidates to national office. Second, mayors can only mobilize voters at the "city level" whereas governors can mobilize voters at the "state level." Third, we assume that governors can foster the political careers of mayors. Thus, we characterize intra-party politics as a game where governors may monitor mayors to screen them for potential promotion and to defuse free-riding in pursuit of the common goal of helping a copartisan candidate win.

[3] "The glass empties for the Kirchners," *The Economist*, June 18, 2009.

[4] "Oposición en Venezuela lista para unas históricas primarias", *BBC Mundo*, February 9, 2012.

The main implication of our argument is that, counterintuitively, the presence of a copartisan governor can have a *dampening effect* on mayoral mobilization on behalf of a party's candidates. This happens because mayors have incentives to free-ride on each others' effort: the more copartisan mayors there are in a district, the more likely it is that the effort of other mayors will suffice to win the district, and the less likely that a given mayor will be promoted due to his effort. In practical terms, this means that a party with a large number of local agents is not necessarily in a much better position to elect national-level candidates than a party that has just a handful of such agents. Using observational data from Mexican congressional elections between 2000 and 2012, we find evidence consistent with this dampening effect. Specifically, we show that having more copartisan mayors in a district improves a party's probability of winning that district only when such party does not control the state governorship. The effect is especially marked for Mexico's historic hegemonic party, the PRI, which is widely credited as having a strong political machine. We interpret these results as a product of collective action problems derived from relatively limited opportunities of promotion within a party.

1 Related literature

The question we analyze — when do local politicians support copartisan candidates to national office — concerns hierarchical relations among politicians within the same political party. Various literatures analyze similar concerns. An extensive literature on *voter mobilization* has inspected the effects of local campaigning and canvassing on the electoral fortunes of political parties, especially through the promotion of turnout (Merriam and Gosnell 1924, Gosnell 1937, Patterson and Caldeira 1983, Caldeira, Patterson and Markko 1985, Caldeira, Clausen and Patterson 1990, Huckfeldt and Sprague 1992, Green, Gerber and Nickerson 2003, Green and Gerber 2004). These studies suggest that party effort is deployed with an eye to winning elections, but the question that this literature addresses is not *whether* the local party organization will choose to boost the party's vote share (it will!), but *which* local politicians should be awarded resources to aid mobilization efforts on behalf of the party (e.g. Holbrook and McClurg 2005).

The literature on *coattail effects* has also recognized that copartisan candidates in different electoral arenas benefit from each others' efforts (Ferejohn 1986, Jones 1997, Burns 1999, Samuels 2000, Carey 2003, Hogan 2005, Golder 2006, Rodden and Wibbels 2011, Zudenkova 2011, Magar 2012, Meredith 2013). These effects, however, result from *externalities* of the effort that candidates make *to secure their own victories.* The idea is that a local candidate that manages to turn out a large number of sympathizers to vote for her might end up bolstering the vote share of copartisan candidates elected concurrently to other posts: citizens vote a straight ticket based on a party label, and the mechanism operates even if the more recognizable candidate makes *no explicit attempts* to improve the vote share of copartisan candidates running concurrently.[5]

A growing literature on *multi-level parties* explores how local and national politicians interact under various institutional arrangements (see e.g. Koelble 1996, Van Houten 2009), how a discernible "local vote" might emerge as a consequence of certain national-level determinants (Morgenstern and Swindle 2005, Morgenstern and Vázquez-D'Elía 2007, Morgenstern, Swindle and Castagnola 2009), how coordination of electoral strategies among local and national politicians shapes national party systems (Caramani 2004, Chhibber and Kollman 2004, Gibson 2012), and how national elites interact with local machines in clientelistic party systems (see Leiras 2006, Magaloni, Díaz-Cayeros and Estévez 2007, Weitz-Shapiro 2012, Albertus 2013, Stokes et al. 2013, Gans-Morse, Mazzuca and Nichter 2014). This literature refocuses research on what one might call the micro-foundations of solidarity among copartisans, suggesting that we cannot analyze within-party cooperation as self-enforcing. Indeed, systematic attempts to measure the effect of local copartisan politicians on the elec-

[5] Ames (1994) analyzes city mayors in Brazil that endorse presidential candidates, an instance of a purposeful cue; it is not clear, however, whether endorsements increase a presidential candidate's local vote share as a consequence of the cue or because endorsing mayors actually exert effort in mobilizing voters.

toral returns of national candidates suggest the possibility of agency problems in the relation between local and national politicians. In the absence of such agency problems, estimates of a mayoral effect would be consistently positive, but a large and growing set of studies suggest that the effect of local politicians on the electoral returns of copartisans may be positive, nil, or even negative (Broockman 2009, Erikson, Folke and Snyder 2015, Folke and Snyder 2012, Novaes 2018).

Yet even though scholars do not always take within-party cooperation for granted, few have attempted to formalize the agency dilemmas that we highlight in this paper. Zudenkova (2011) mentions the possibility that copartisan politicians may free-ride on each others' campaigning efforts, but this is in a formal model of coattails in which coattails arise because candidates care about their copartisans' performance and voters evaluate copartisan candidates jointly. Shachar and Nalebuff (1999) suggest that US governors have an incentive to mobilize in favor of copartisan presidential candidates when they anticipate that their state will be pivotal in the electoral college. Kemahlıoğlu (2012) studies tensions between local and national politicians that obtain as the former build support networks across localities: national politicians have an incentive to fund the creation of these networks — they are helpful to mobilize voters — but are also wary that local politicians can use these networks to challenge their leadership. Stokes et al. (2013) model clientelistic brokers as opportunistic agents with private information, who recruit partisan rather than swing voters in order to signal their mobilization capacity — thus complicating their party's effort to target voters efficiently. Persico, Rodríguez Pueblita and Silverman (2011) and Benton (2007) build models of intra-party strife in which within-party factions appear when politicians compete to obtain patronage or local public goods for their districts. In contrast, our contribution illuminates some of the mechanisms that prevent local politicians from supporting copartisans even when factionalism is not a relevant concern.

2 A theory of local mobilization in multi-level parties

We develop a model in which politicians at different government levels — mayors in cities and governors in states — act as agents of a copartisan candidate who is running for a congressional seat. We seek to understand the incentives that push local politicians to exert effort on behalf of the national candidate. We focus on the strategic interaction between one *governor* (G) and up to three *copartisan mayors*, which strikes us as the easiest way to capture variation in the proportion of mayors in a district that are copartisans of the governor.

This political geography captures conditions typical of many political systems with multi-level governance. Mayors govern cities that are nested within electoral districts. Governors govern states that can comprise several electoral districts, though for simplicity we assume that state and district are coterminous. We model mayors as politicians that court gubernatorial support to further their careers.[6] Gubernatorial support appears as a decision to reward the mayor for a job well done; the reward broadly captures any kind of gubernatorial support for the mayor's future career, such as endorsement for a re-election run, promotion within the party ranks, or nomination for a candidacy to higher office. We do not consider voters' behavior explicitly, though voters appear indirectly in the model through a parameter that captures the party's "normal vote" in the district. Our goal is to study the behavior of local politicians in a context in which they have the means to benefit their copartisans but lack an incentive to mobilize voters because their own reelection is not at stake. For this reason, we assume that local and national elections are *not* concurrent.[7]

---

[6] Our use of terms like mayors, governors, cities, and states is meant to be general; many electoral districts are nested within cities, rather than the other way around. However, electoral districts are often composed of smaller units, like wards, where local politicians face incentives similar to those encountered by mayors in our model.

[7] This assumption is not particularly limiting, as we observe non-concurrent or partially-concurrent local/national elections in many countries, including Austria, Brazil, Canada, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Germany, India, Mexico, Peru, Russia, South Africa, Spain, Switzerland, United Kingdom, United States, Uruguay and Venezuela.

Governors and mayors can increase the probability that their copartisan candidate will carry the district by mobilizing voters, which requires effort. We treat the effort exerted by mayors and governor as additive, so that the probability that the national candidate will carry a district when all agents cooperate is strictly larger than when only some exert effort.[8] Players have utilities defined over *effort* and *reward*. Specifically, mayors obtain reward $r>0$ if promoted by the governor, which can only occur if the party wins the district. The cost of exerting effort depends on a mayor's type. With probability $\tau \in (0,1)$, a mayor is of *good type* and pays no cost for exerting effort ($f_g(e)=0$), while *bad-type* mayors incur a cost from effort ($f_b(e)=f>0$). The governor is interested in promoting good mayors because these are adroit politicians capable of helping the party in the future. To capture this incentive, G receives a payoff of $R=1$ for promoting a good type, and $R=0$ otherwise.

To decide whom to promote, the governor can invest effort in monitoring mayoral behavior. Monitoring detracts from campaigning directly on behalf of the party's candidate, introducing a sharp trade-off in the governor's actions. Thus, at the beginning of the game the governor decides between two options that we model as mutually exclusive for the sake of simplicity: to *campaign* on behalf of her party (c) or to *monitor* the behavior of mayors (m). Specifically, G must choose between $\{m=1,\ c=0\}$ with cost $g>0$; $\{m=0,\ c=1\}$ with cost $k>0$; or $\{m=c=0\}$, which is costless. Admittedly, real governors might choose to split effort between campaigning and monitoring or to campaign once monitoring reveals information about a mayor's behavior. However, gubernatorial effort is limited, and the stark assumption that effort can be deployed either to monitor or to campaign puts in sharp relief the consequences of adopting either action without missing important characteristics of the strategic environment. Through monitoring, G can ascertain whether a mayor exerted effort or not. For simplicity, we assume that if the governor chooses to monitor, she picks a

[8] The assumption of additive returns to effort disallows potential team-effort complementarities, which are not the focus of this paper.

copartisan mayor at random,[9] and receives a perfectly reliable signal of whether the mayor exerted effort, though not of the mayor's *type*. At the end of the game, the governor chooses to reward up to one copartisan mayor, though she may also decide not to reward anyone.

The outcome of interest is the probability that the party's candidate will win the district, which depends on the effort of the governor and mayors (c and e, respectively), the district's proclivity to vote for the opposition ($v \in [-1, 1]$), and a pro-opposition shock ε with known uniform distribution along the $[-\frac{1}{2\psi}, \frac{1}{2\psi}]$ interval, with $\psi \in (0, 1/5]$.[10] With these assumptions in place, the probability of victory in the district is:

$$\begin{aligned}
\Pr(\text{Win}) &\equiv \Pr\left(\varepsilon \quad \leq \quad \overbrace{I[c]s}^{\substack{\text{Governor's}\\ \text{campaigning}\\ \text{effort}}} \quad + \quad \overbrace{\Sigma_{i\in\mathcal{I}} I[e_i]\left(\frac{1-s}{3}\right)}^{\substack{\text{Copartisan mayors'}\\ \text{campaigning effort}}} \quad - \quad \overbrace{\Sigma_{o\in\mathcal{O}} I[e_o]\left(\frac{1-s}{3}\right)}^{\substack{\text{Opposition mayors'}\\ \text{campaigning effort}}} \quad - \quad v\right) \\
&= \int_{-\frac{1}{2\psi}}^{I[c]s+\Sigma_{i\in\mathcal{I}} I[e_i]\frac{1}{3}(1-s)-\Sigma_{o\in\mathcal{O}} I[e_o]\frac{1}{3}(1-s)-v} h(\varepsilon)\, d\varepsilon \\
\Pr(\text{Win}) &= \frac{1}{2} + \psi\left(I[c]s + \Sigma_{i\in\mathcal{I}} I[e_i]\left(\frac{1-s}{3}\right) - \Sigma_{o\in\mathcal{O}} I[e_o]\left(\frac{1-s}{3}\right) - v\right), \qquad (1)
\end{aligned}$$

where i and o denote whether mayors belong to the *incumbent* governor's party or to the *opposition*, respectively, and $I[\cdot]$ is an indicator function specifying whether a player engages in a certain action, such as campaigning (c=1) or exerting effort (e=1). The contribution of gubernatorial effort to the likelihood of victory is $s \in (0, 1)$; each mayor that exerts effort contributes an additional $\frac{1-s}{3}$ to the likelihood of victory if he belongs to the incumbent party and $-\left(\frac{1-s}{3}\right)$ if he belongs to the opposition.[11] This captures in a simple way two relevant aspects of party competition: First, s indicates the relative efficacy of gubernatorial versus mayoral effort; higher values of s imply that the governor is more effective than the mayors at mobilizing votes on behalf of the party if she chooses to campaign. Second, the

[9] It makes no sense to monitor an opposition mayor, as the governor cannot promote him within her party's ranks.

[10] We assume small enough values for the exogenous parameter ψ such that an adverse shock could *potentially* produce an electoral defeat even when all local politicians campaign for the party in a pro-party district.

[11] Each mayor contributes one-third of the total mayoral effort $1-s$, since each oversees one-third of the electoral district and one-third of the district's voters.

efforts exerted by opposition- and incumbent-party mayors are countervailing.

Equation 1 shows, as one would expect, that the probability of carrying the district decreases with the opposition's normal vote (v) and the effort of opposition mayors, and increases with copartisan effort. More importantly, *no actor single-handedly determines the probability of winning; victory is a collective outcome determined by the aggregation of individual choices.* The predictability of the race, measured by ψ, interacts with these quantities: as ψ increases, the probability of winning is more tightly linked to effort. The probability of copartisan victory is much lower when the race is very predictable ($\psi \rightarrow 1/5$), and the district leans toward the opposition ($v \rightarrow 1$), and opposition mayors are of good type and thus exert effort ($e_o = 1 \; \forall o$).

The sequence of moves is as follows. Before the election, Nature chooses mayoral types, then bad-type mayors choose e and the governor chooses m or c simultaneously (if the governor chooses m, she then observes e for one copartisan mayor; her decision to monitor is *not* conditional on the mayor's action). After the outcome of the election is revealed, the governor decides which mayor to reward only if the party wins the district (the governor's decision to reward *is* conditional on the mayor's action). Finally, the mayor's type is revealed and payoffs are collected. Notice that the only strategic players in the game are the *governor* and *bad-type mayors* in the governor's party: good types always exert effort and bad types in the opposition never exert effort because they would never be rewarded by the governor.

We discuss equilibria for a scenario in which all three mayors belong to the governor's party and refer tangentially to equilibria in scenarios with two and one copartisan mayors (Appendix A provides detailed proofs for all scenarios).[12] We first characterize the governor's posterior beliefs about a monitored mayor's type. Recall that, if she chooses to monitor, the governor monitors a copartisan mayor chosen at random. We first show that the governor's optimal strategy is promotion conditional on learning that a monitored copartisan mayor

[12]In a scenario with three copartisan mayors, Equation 1 simplifies to $\Pr(\text{Win}) = \frac{1}{2} + \psi\left(I[c]s + \Sigma_{i \in \mathcal{I}} I[e_i] \frac{1-s}{3} - v\right)$.

exerted effort:

Lemma 1 *When two or more mayors belong to the governor's party, and the governor decides to monitor a copartisan mayor, the following rule is optimal: (i) promote the monitored mayor if he exerted effort, and (ii) promote some other randomly-chosen copartisan mayor if the monitored mayor did not exert effort.*[13]

Recall that the governor cannot learn a monitored mayor's type, only whether he exerted effort. If monitoring reveals lack of effort by the mayor, the governor can infer that the mayor is of bad type, as good types find effort costless. If monitoring reveals effort, the governor might still be in the presence of a bad type masquerading as a good one. However, the probability that a mayor observed to exert effort is of good type is equal to or larger than τ, and thus promoting him is better than promoting some other mayor at random (see Appendix A).[14]

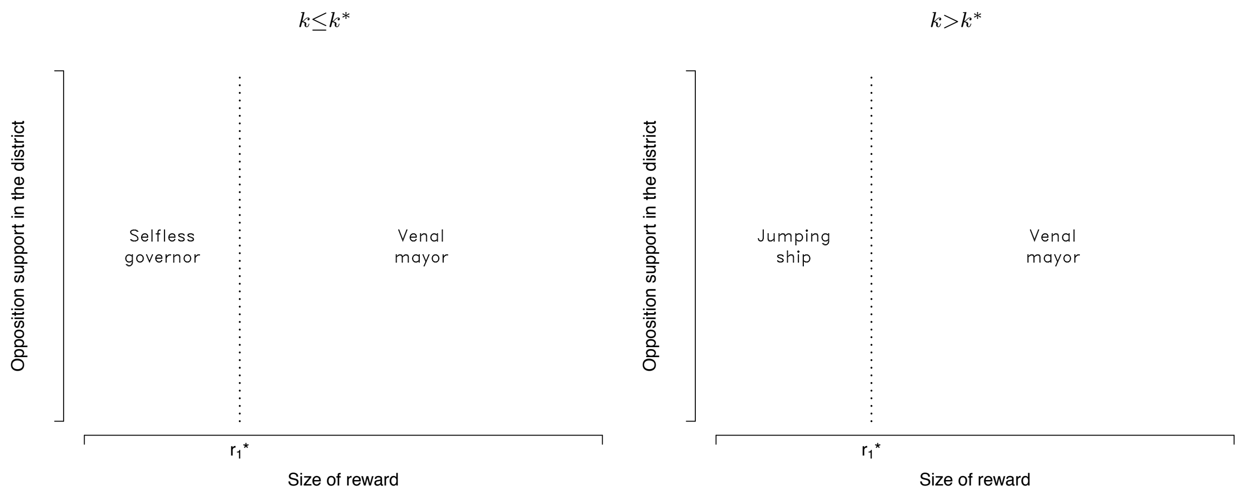
We now turn to a description of the different equilibria. We proceed by iterative elimination of suboptimal strategies, emphasizing the threshold values of reward (r), cost of campaigning (k), and opposition normal vote (v) for which (some) players have dominant strategies. We then find optimal responses for other players. When there are no equilibria in pure strategies, we show that equilibria exist in mixed strategies. Figure 1 guides our discussion of the different equilibria that exist when there are one, two, and three copartisan mayors in a district.[15] Notice that the scenarios with two and three copartisan mayors are similar, but in the one-copartisan-mayor scenario there are no equilibria in mixed strategies.

[13]Lemma 1 requires that we restrict attention to "symmetric equilibria" among all bad types, a common restriction in this kind of models (cf. Palfrey and Rosenthal 1985).

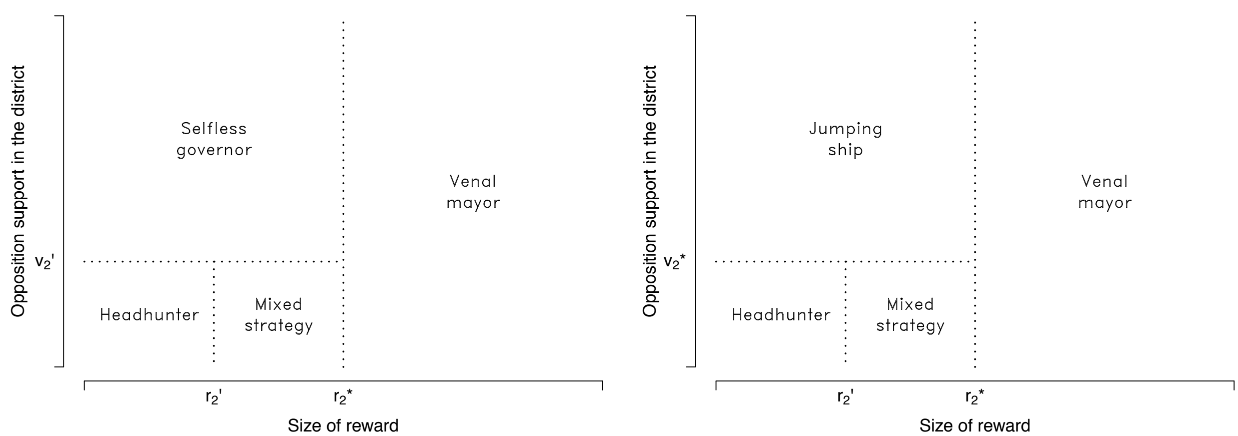
[14]Strictly speaking, we find perfect Bayesian Nash equilibria of the game, which requires that we characterize gubernatorial beliefs about mayoral types. The proof of Lemma 1 in the Appendix shows how this belief is determined by Bayes's rule and how the decision to promote is sequentially rational given this belief.

[15]The scenario with no copartisan mayors is trivial because G cannot promote anybody, and thus lacks incentives to monitor.

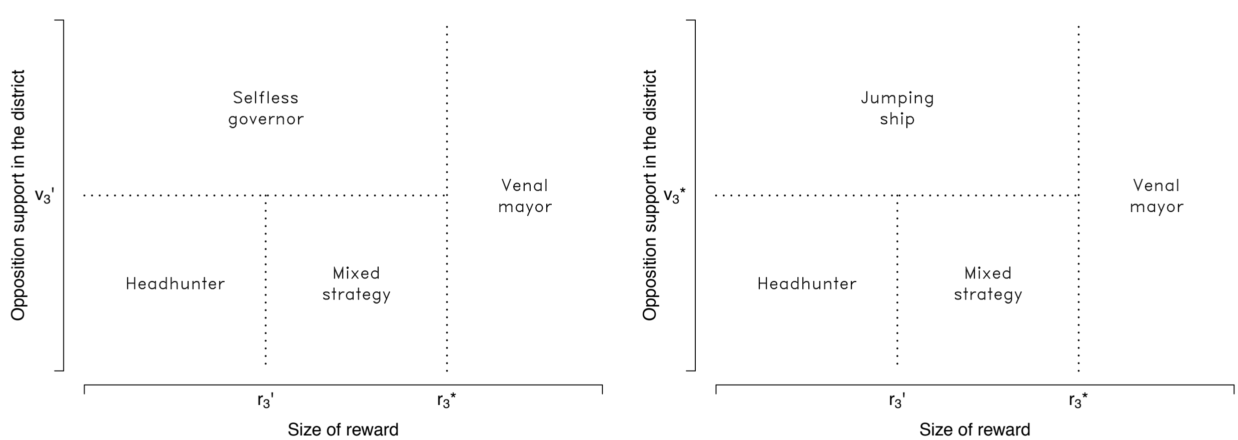
Figure 1: Equilibria of the *local mobilization game* under three alternative scenarios. The horizontal and vertical axes indicate the values of *reward size* (r) and *opposition's support in the district* (v) that serve as cutpoints among the different equilibria. Note that $r_3^* > r_2^* > r_1^*$.



(a) One copartisan mayor



(b) Two copartisan mayors



(c) Three copartisan mayors

Proposition 1 *A* venal mayor *equilibrium exists when the reward for promotion is high* ($r_3^* \geq 9f/[\psi(1-s)(3-2\tau)]$). *Under these circumstances, the allure of a large reward motivates bad type mayors to mobilize (*e=1*) regardless of the governor's actions. The governor's best response is to campaign (*c=1*) if the cost of doing so is low (*$k_3^* \leq \psi s\tau$*) and do nothing (*c=m=0*) otherwise.*

The threshold r_3^* establishes a sufficient condition for mayoral mobilization, as can be seen along the horizontal axis in Figure 1. As in other delegation games (e.g. Besley 2007), a "disciplined" bad-type can masquerade as a good type, which is beneficial for the party in the short-term (it brings votes) but detrimental in the long run because the governor could end up rewarding a bad type that will weaken the party. We say such disciplined behavior is "venal" because the bad-type mayor only works on behalf of the party when he has a shot at getting a reward. Specifically, the size of the effort-inducing reward r_3^* depends on (i) the relative mobilizing capacity of the governor (s), (ii) the abundance of good types in the population (τ), (iii) the cost of exerting effort (f); and (iv) the uncertainty surrounding the election outcome (ψ). If governors have a mobilizing advantage (high s) or if the expected incidence of good types is high (high τ), the bad type will reason that his contribution is unlikely to lead to victory (his relative mobilization capacity is low and other mayors are probably of good type, therefore already campaigning and likely to get the reward if they are monitored). Given the low chance of being rewarded, bad types will only bother to exert effort if the reward is very large, or the cost of exerting effort is very low.[16] Similarly, as the translation of effort into outcomes becomes more reliable (i.e., as ψ increases), the size of the effort-inducing reward diminishes: the cost of effort is independent of the electoral outcome, so an increase in the reliability of effort will induce more of it. A *venal mayor* equilibrium also exists in scenarios with one or two copartisan candidates, though the effort-inducing reward in these circumstances is smaller (i.e., $r_1^* < r_2^* < r_3^*$; see Appendix A). This

[16]In the limit, if f=0 (i.e., effort is costless), a mayor will exert effort as long as $r_3^* \geq 9 \cdot 0/[\psi(1-s)(3-2\tau)] \Rightarrow r_3^* \geq 0$, which is consistent with the claim that good-type mayors — for whom, by assumption, effort is costless — always exert effort.

finding follows from the higher probability that a bad-type mayor will obtain the reward when there is less competition from other copartisan mayors, and we will make use of it in deriving a testable implication of the model.

When the reward is low ($r<r_3^*$), we find two equilibria in pure strategies in which bad mayors *do not exert effort*, and one equilibrium in mixed strategies in which bad mayors *exert effort only some of the time*. The first of the pure-strategy equilibria is a *headhunter monitoring* equilibrium:

Proposition 2 *A* headhunter monitoring *equilibrium obtains when the reward is low ($r<r_3'$), but the probability of carrying the district in the absence of effort is high (either $v<v_3^*$ or $v<v_3'$). Bad-type mayors do not campaign (e=0) because the reward is too low, and the governor monitors (m=1) to find a good candidate for promotion.*

The intuition for this result is that mayors know there is a high probability that a copartisan will carry the district *even if they do nothing* (because the opposition normal vote, v, is low enough). Since the reward for exerting effort is low, bad types lack an incentive to help regardless of the governor's actions. But the governor can still monitor mayors to find a good type to reward after the election. In fact, *monitoring* and *low rewards* go hand-in-hand: because the reward is low, bad types do not consider it worth their while to try to distinguish themselves by exerting effort. After all, if the governor monitors some other bad type and finds him wanting, there is still a chance they might get the reward without having exerted effort. The requirement that v be low enough is also intuitive. Doling out the reward requires winning the district first; if the election is lost, finding that a mayor is of good type is worthless. Thus, when the opposition enjoys an electoral advantage the governor has incentives to campaign rather than to monitor the mayors. But if the governor's party can expect to win, the benefits of identifying a good mayor outweigh those of campaigning: in these circumstances, effort exerted by a monitored mayor is superfluous in winning the district, but it still reveals useful information about the mayor's type. Appendix A shows that a similar logic applies in a scenario with two copartisan mayors, but not in the one-mayor scenario, as the latter has no incentive to monitor a mayor for whom, were he found to be a bad type, there is no potential substitute.

The second equilibrium in pure strategies in a low reward environment is one where a *selfless governor* "takes one for the team" while mayors coast. This equilibrium can only be sustained if the cost of campaigning is low enough; otherwise, both the probability of carrying the district and the potential mayoral reward are so low that local politicians — *both* mayors *and* governor — lack incentives to mobilize on behalf of copartisan candidates and choose to *jump ship* instead.

Proposition 3 *A* selfless governor *equilibrium exists where the governor campaigns even though (i) rewards are low and thus unlikely to motivate mayoral effort ($r<r_3^*$) and (ii) the normal opposition vote is relatively high ($v\geq \max\{v_3', v_3^*\}$). This equilibrium requires a low campaigning cost for the governor ($k\leq k_3^*$). If k is too large, the* selfless governor *equilibrium cannot be sustained and cooperation by local politicians unravels, leading to a* jumping ship *equilibrium in which no player exerts effort.*[17]

The main factor driving a governor to act selflessly is the low cost of campaigning. By choosing to campaign on behalf of her party, the governor has a shot at rewarding a mayor — an eventual ally. The cost k that the governor is willing to pay increases with the amount of electoral certainty (ψ), the efficiency of her own effort (s), and the distribution of good types in the population (τ), though this cost does not vary across scenarios (i.e., $k_1^*=k_2^*=k_3^*$; see Appendix A). The governor will only act selflessly if the party is at least minimally competitive (relatively low v), because this gives her a shot at promoting a mayor believed to be of good type. If the party is not minimally competitive, neither the governor nor the mayors have incentives to campaign at all, leading to a "jumping ship" equilibrium.

[17] Rundlett and Svolik (2016) suggest a similar logic to explain why local operatives in Russia may refrain from orchestrating electoral fraud on behalf of a hegemonic party. In their case, local agents do not engage in fraud if the chance of losing the election is too large (they can get caught and be punished by the victorious opposition); however, if the chance of winning is high enough, they participate willingly to increase the chance of a reward. The implication is that fraud is thus over- or undersupplied.

Finally, when the effort-inducing reward increases beyond r'_3 but still remains below r^*_3, players adopt a *mixed-strategy* equilibrium:

Proposition 4 *A* mixed strategy *equilibrium obtains when the reward is neither too high nor too low ($r'_3<r<r^*_3$), and the probability of carrying the district in the absence of effort is sufficiently high (either $v<v^*_3$ or $v<v'_3$). In this case, bad-type mayors sometimes exert effort and governors sometimes monitor, with probabilities chosen so as to make the other actor indifferent.*

Intuitively, if bad mayors coast, the governor prefers to monitor them in order to identify a good one for promotion; but since the reward is large enough to induce bad mayors to exert effort if they expect to be monitored, the governor will prefer not to monitor anybody, and so on. To put it differently, the governor only wants to monitor when bad mayors do not exert effort, while bad mayors only want to exert effort when the governor is monitoring them. Thus governor and mayors will choose a combination of actions to make each other indifferent between exerting effort or not, and monitoring or not. Appendix A shows how the mixing probabilities depend on the cost of monitoring (g), the cost of exerting effort (f), and the size of the reward (r). The Appendix also shows that a similar logic applies to a scenario with two mayors, but if there is a single copartisan mayor a mixed strategy equilibrium does not exist because the governor never has an incentive to monitor.

Observable implications of the model

Ideally, we would evaluate these propositions by carefully measuring the amount of effort that mayors exert in favor of copartisan candidates and by gauging the level of involvement of governors in a campaign or in monitoring mayoral behavior. Neither individual nor team effort are readily observable — neither by us nor by party leaders, which is an important part of our argument and the starting point of this paper.[18] However, we can inspect whether

[18]Szwarcberg (2014) suggests that politicians can measure the effort of political operatives by looking at the size of the rallies they organize. However, this requires the ability to

Table 1: Four possible combinations of copartisan officials

		Governor	
		Copartisan	Opposition
# Copartisan mayors	Few	I	II
in district	Many	III	IV

a congressional candidate carries the election *at the district level*, which is the outcome of interest in our theory.

Thus, we focus on comparing a congressional candidate's probability of winning across electoral *districts* that comprise several *municipalities*. As shown in Table 1, these districts vary along two dimensions: (i) the proportion of municipalities within the district that are governed by copartisan mayors and (ii) the governor's copartisan status. All other things equal, both factors should have a positive effect on the probability that a party captures a district: on the one hand, the governor may campaign or monitor mayors, thus inducing them to exert effort; on the other, the more mayors there are, the more likely it is that some of them are of good type, and thus will campaign regardless of the governor's behavior. Yet this only holds as long as we consider the governor or the mayors *in isolation*; if players interact with each other, some may try to pass the responsibility for mobilizing voters to others and thus increasing the number of copartisans in a district does not necessarily produce a corresponding increase in the party's chances of carrying the district.

In other words, we claim that having a copartisan governor will exert a *dampening effect* on the degree of mayoral mobilization in a district.The intuition behind this claim is the following. When there is no copartisan governor (quadrants II or IV in Table 1), only good mayors will exert effort. To the extent that the probability of having a good mayor is increasing on the number of mayors in a district, having many mayors will increase a party's electoral performance in the district. When there is a copartisan governor, on the other

observe each and every rally in every location. Furthermore, such behavior may not lead to optimal mobilization effort because brokers often have incentives to recruit "cheap" copartisans rather than (relatively) "expensive" swing voters (Stokes et al. 2013).

hand, this logic continues to hold for good mayors, but now bad-quality mayors may also exert effort if conditions are right. Yet these conditions are very specific: bad-quality mayors *never* exert effort in three of the five equilibria derived above (*selfless governor*, *headhunter*, or *jumping ship*), and in other (*mixed strategies*) they exert effort only some of the time; only in the *venal mayor* equilibrium do bad-type mayors consistently exert effort on behalf of the party. And moving from the top to the bottom of Figure 1 shows that for a given reward size, the area for which the *venal mayor* equilibrium holds — that is, the range of combinations of *reward size* and *opposition support* — is decreasing on the number of mayors in a district. To put it differently, bad-quality mayors never have as many incentives to exert effort than when they are "alone" with the governor.

Such incentives are reinforced by the governor's behavior. As Figure 1 shows, the governor never has incentives to campaign if $k > k^*$, which does not depend on the number of mayors. Yet whether the governor campaigns or monitors does depend on the number of mayors in a district. Figure 1 indicates that when there is a single mayor, the governor always campaigns (as long as $k > k^*$); as the number of mayors increases, however, the values of the parameters for which the governor campaigns decreases, because the governor prefers to monitor mayors in order to find a good one for promotion. Of course, this is good enough for the governor, and may be good for the party *in the long run*, insofar as it fosters the promotion of good politicians; but in the current election, it means that the governor will campaign less, and thus the party's probability of winning the district will be lower than otherwise.

Summing up: throwing more copartisan officials into a district would be unquestionably good if such officials always exerted effort on behalf of their party, but the model shows that the *simultaneous* provision of effort by the governor and the mayors is both (a) relatively rare; and (b) decreasing on the number of mayors in a district. Intuitively, this results from the combination of two mechanisms. On the one hand, mayors face a cooperation problem: the more numerous they are, the more likely it is that the party will win the district, but the less likely that any given mayor will be promoted; thus, unless the reward is very large, bad mayors prefer to coast rather than work for the party. On the other hand, the governor finds it less attractive to campaign but more rewarding to monitor mayors in order to identify a good one for promotion. That is, when there is a copartisan governor, increasing the number of mayors will increase the number of good types that always exert effort, but it will also discourage both the governor and bad-type mayors from campaigning, thus (partially) offsetting the benefits of having more good mayors.

Note that we do not claim that having a governor is "bad" for a party; on the contrary, we expect a congressional candidate to do better when there is a copartisan governor. Rather, our claim is that a copartisan governor should *dampen* the effect of having more copartisan mayors in a district. Formally, let the probability that party j captures district d depend on both the number of copartisan mayors in d, M_{jd}, and whether there is a copartisan governor ($G_{jd} = 1$), such that

$$\Pr(\text{Win}_{jd}{=}1|G_{jd}{=}1) = \beta_0 + \beta_1 M_{jd}$$

$$\Pr(\text{Win}_{jd}{=}1|G_{jd}{=}0) = \gamma_0 + \gamma_1 M_{jd}$$

Our claim is that $\beta_0 > \gamma_0$, which is an expectation for which little theoretical insight into political behavior is needed, but also $\beta_1 < \gamma_1$, which is not such an obvious claim. The latter difference corresponds to the dampening effect that follows from our formal argument.

Before considering the empirical evidence, we comment briefly on the plausibility of the assumptions behind our formal argument. First, our argument applies more clearly to strong party organizations that are able to promote politicians within the party ranks. It would certainly be more difficult to see these dilemmas in parties that lack minimal control over decisions to nominate candidates for different local and national positions. With this caveat in mind, the gist of our argument holds if we assume that the congressional candidate, rather than the governor, performs the task of monitoring. Even more so than with the governor, the congressional candidate would face a trade-off between campaigning and monitoring. We anticipate that the argument would hold as well if, more realistically, we allow for the possibility that the governor may monitor and/or reward several mayors. What matters in our model is the characterization of rewards (gubernatorial support) as limited, an assumption that is valid as long as the number of mayors that aspire to improve their careers is larger than the number of promotion opportunities within the party, which seems reasonable in most contexts.

3 Estimating the electoral impact of copartisan mayors and governors

We validate our empirical claims using data from congressional elections in Mexico between 2000 and 2012. Mexico provides an ideal scenario for examining these predictions, for three reasons. First, several Mexican states have electoral calendars that are not concurrent with national elections, which helps us distinguish the effect of mobilization in support of a copartisan candidate from the "reverse coattail" effect of mobilization to secure control of the mayor's office; indeed, we restrict the analysis to states where local and federal elections take place in different years. Second, Mexican governors play an important role in promoting their copartisan's future careers (Rosas and Langston 2011, Kerevel 2015). Third, during the period under study the Mexican constitution barred *all* elected officials from seeking consecutive reelection at the end of their term. This ensures that any positive effect that governors and mayors exert on the probability of victory of their copartisans is a consequence of their effort on behalf of copartisan candidates, rather than an attempt to secure their own reelections in simultaneous races. Similarly, the probability of winning the district cannot be attributed to an incumbent legislator's advantage, because legislators are also limited to one term in office. Finally, Mexican municipalities are for the most part nested within single-member districts, which in turn are nested within states. Specifically, the Mexican lower chamber is elected using a mixed-member system: 200 legislators are elected in five proportional-representation districts, the remaining 300 in single-member districts. We focus exclusively on the single-member constituencies, most of which comprise several municipalities.[19]

[19] Voters cast a single ballot to select their SMD representative. The SMD ballots are then aggregated at the PR district level to determine the distribution of PR seats (see Calvo

Our study is based on a set of 461 district-elections throughout five federal elections scheduled triennially from 2000 to 2012. The number of municipalities per district ranges from 1 to 40, with a median of 9 and an average of 9.5. Each district-election yields two observations — one for the party that controls the governorship at the time of the election and one for the runner-up party in the previous gubernatorial election — which we separate into different sets to estimate two models — one for candidates with copartisan governors and the other for candidates without copartisan governors.[20] We follow this approach because even though Mexico has a three-party system at the national level, subnational competition is mostly bipartisan during the period under observation: in most states the *Partido Revolucionario Institucional* (PRI), the former hegemonic party, competes against a single large opposition party, either the center-right *Partido Acción Nacional* (PAN) or the center-left *Partido de la Revolución Democrática* (PRD). For example, the PRI has governed the state of Chihuahua between 1998 and 2016, having beaten PAN candidates in what were practically two-party races. Hence, districts in Chihuahua furnish two sets of observations: one set for PRI mayors and congressional candidates operating under a copartisan PRI governor, and one set for PAN mayors and congressional candidates that operate under an opposition (i.e., PRI) governor.

The outcome is an indicator of whether the congressional candidate carried district d and the parameter of interest is the effect of the *proportion of copartisan mayors* in the district (i.e., β_1 and γ_1). Many factors extraneous to our model have an effect on both mayoral behavior and a congressional candidate's probability of victory; moreover, we cannot feasibly measure a host of factors within our model (*size of reward*, *gubernatorial capacity*) that would need to remain constant in order to verify the existence of a dampening effect. In short,

and Abal Medina 2002).

[20]The actual number of observations in the copartisan governor sample is 453, not 461. Attrition occurs because of the few cases in which PAN and PRD form an alliance to elect a mayor. In these cases, we cannot know the partisan identity of the mayor, and we are forced to drop the district to which the municipality belongs.

we cannot claim that the assignment of *proportion of copartisan mayors* across districts is exogenous, which is an important limitation in our analysis.[21] We do control for observable confounders in our models: the party's *margin of victory* in the municipality in the previous district-level race, the municipality's *poverty* level, and a quadratic polynomial of the party's *normal vote share* in the district.[22]

Table 2 displays ordinary least squares estimates of the effect of the proportion of copartisan mayors in a district on the probability that the congressional candidate will win the district based on two sets of observations (for candidates with and without copartisan governors) in a linear probability model.[23] Models 1 and 3 control for *margin of victory*, *normal vote share*, *poverty*, and year fixed effects; Models 2 and 4 in addition include state fixed effects. In Panel (a), we consider results based on the full sample of districts and municipalities they contain, which include governors and mayors from the three main political parties. In line with expectations, comparing the intercept estimates in Models 1–2 against those in Models 3–4 confirms that, when the proportion of copartisan mayors in the district is nil, the probability that a congressional candidate will carry a district in a state with a copartisan governor is larger than the probability that she will win the election if the state is led by a non-copartisan governor. In addition, we find *some* evidence of a dampening effect, but this evidence is not entirely consistent with our claims. To see this, consider that the point

[21] Nonetheless, in Appendix B, we present results from a potential identification strategy based on Folke's (2014) pioneering idea of employing the number of seats that a party barely wins or loses in a high-magnitude proportional representation district as an instrument for the total number of seats that a party captures in a district.

[22] *Poverty* is an aggregate measure compiled every five years by Mexico's National Population Council; it is the first principal component of a number of correlates of poverty, such as a municipality's share of houses with unpaved floors, a municipality's mean wage, etc. A party's *normal vote share* in district d at time t is the average of the party's vote share in that district at $t-1$, *and* the (population-weighted) average vote share at $t-1$ in all municipalities that comprise the district.

[23] Table A7 in the Appendix shows that using a probit link produces substantively identical results.

Table 2: LPM estimates of *proportion of copartisan mayors* on a congressional candidate's *probability of victory* in Mexico, 2000-2012.

	Copartisan governor		*Opposition governor*	
	(1)	(2)	(3)	(4)
(a) Full Sample				
Proportion of copartisan mayors	0.26	0.17	0.28	0.29
	(0.10)*	(0.09)†	(0.13)*	(0.12)*
Margin of victory (lagged)	0.07	-0.26	0.49	0.43
	(0.18)	(0.24)	(0.24)*	(0.20)*
Poverty	0.09	0.17	-0.02	-0.09
	(0.03)*	(0.04)*	(0.03)	(0.02)*
Intercept	0.54	0.65	0.25	0.23
	(0.06)*	(0.07)*	(0.05)*	(0.05)*
RMSE	0.40	0.38	0.39	0.38
num. obs.	453	453	461	461
(b) PRI-only Sample				
Proportion of copartisan mayors	0.02	0.05	0.53	0.43
	(0.08)	(0.08)	(0.11)*	(0.12)*
Margin of victory (lagged)	0.35	-0.09	-0.00	0.10
	(0.17)*	(0.21)	(0.46)	(0.53)
Poverty	0.12	0.19	0.13	0.06
	(0.03)*	(0.04)*	(0.03)*	(0.04)
Intercept	0.66	0.38	0.14	0.16
	(0.05)*	(0.11)*	(0.04)*	(0.07)*
RMSE	0.37	0.36	0.34	0.33
num. obs.	336	336	146	146
Previous vote share	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
State effects	No	Yes	No	Yes

* p<0.05; † p<0.10. OLS estimates (standard errors clustered by state in parentheses.)

estimate of the effect parameter of *proportion of copartisan mayors* is positive in districts with *copartisan governors* (Models 1–2), though the effect is only statistically significant at the 95% confidence level in Model 1, which excludes state-level fixed effects. In contrast, the positive estimates for the effect parameters of *proportion of copartisan mayors* in districts with non-copartisan governors (Models 3–4) are of larger magnitude and statistically significant at the 95% confidence level. The dampening effect we expect requires a difference-of-means test between the coefficients for *proportion of copartisan mayors* in Models 1 and

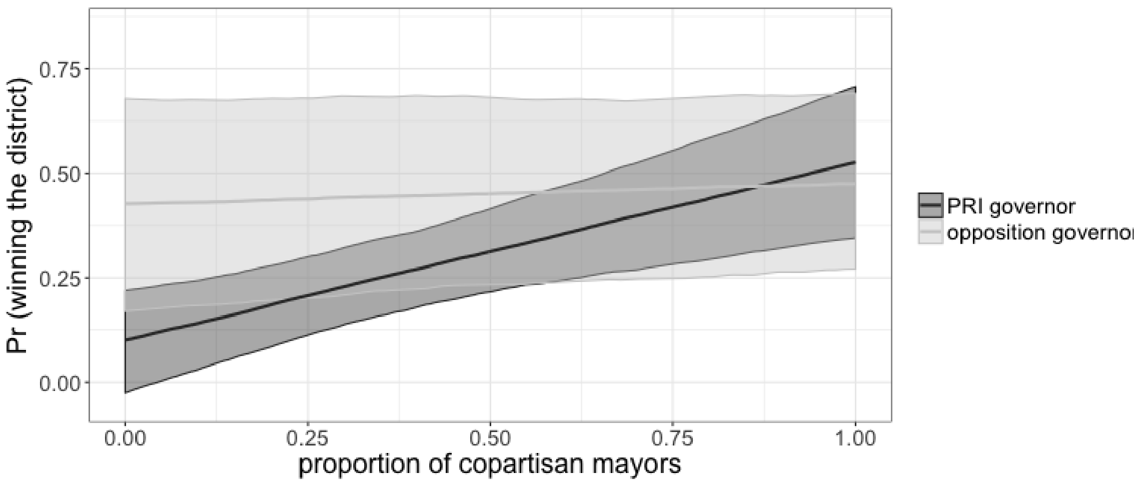


Figure 2: Predicted probability and 95% confidence intervals that the PRI will capture a district, conditional on *proportion of copartisan mayors* and the presence of a copartisan governor. All other predictors are held constant at their means. Values are based on the models reported in columns 2 and 4 of Table 2b.

3 or Models 2 and 4. These tests indicate that the differences between the coefficients in Models 1–2 and 3–4 are not statistically significant, i.e., we cannot discard the possibility that the coefficients may in fact be equal. Nonetheless, a simple difference-of-coefficients test is inappropriate in this case because the samples are *not independent*, as the same district enters the first sample with a copartisan governor for one candidate and the second sample with an opposition governor for a second candidate. Thus, outcomes across samples are negatively correlated: a better performance for the governor's party likely implies a worse performance for the runner-up's, and vice versa.

To break this dependence across samples, Panel (b) in Table 2 replicates the same four models, but limiting the analysis to PRI congressional candidates. As Mexico's erstwhile hegemonic party, the PRI came in first or second place in almost all gubernatorial elections held between 2000 and 2012 (see Langston (2017) for a comprehensive study of the PRI).[24] This ensures that the counterfactuals for districts in which the PRI was the runner-up in the

[24]Of the 74 gubernatorial elections that took place in Mexico between 2000 and 2012, the PRI did not finish in the first or second place only four times: twice in the Federal District (2000 and 2006), once in Michoacán (2007) and once in Morelos (2006). In Chiapas in 2012, the PRI did not win the governorship but was allied to the Green Party, which did.

previous gubernatorial election are districts in which the PRI controlled the governorship, and vice versa. Although this brings about a considerable reduction in sample size, especially for districts not governed by the PRI, results in Table 2b are consistent with the expected dampening effect: when there is a PRI governor, the coefficient of *proportion of copartisan mayors* is essentially zero, while in opposition-governed states the estimated coefficient ranges between 0.43 (SE=0.12) and 0.53 (SE=0.11), and is statistically significant at conventional levels. Moreover, difference-of-means tests between the coefficients of Models 1–2 and 3–4 are statistically significant. To illustrate this finding, Figure 2 shows the probability that the PRI will carry a district conditional on the existence of a copartisan or opposition governor and on the proportion of the district population that is governed by a copartisan mayor, while holding all other variables at their means: when there is a copartisan governor, the probability of winning is large, but it remains essentially flat as the proportion of copartisan mayors increase. When the governor is not a *priísta*, on the other hand, increasing the proportion of copartisan mayors has a large effect on the probability that the PRI will win the district.

Robustness checks

We check the robustness of the dampening effect in a number of ways. First, Table A8 in the Appendix shows that the proportion of copartisan mayors has no effect on a congressional candidate's *vote share* or *margin of victory* at the district level. This is not entirely surprising, as none of these races are strictly bipartisan. Because the ballot in the single-member district aggregates into at-large districts where further seats are awarded by proportional representation, Duverger's law does not operate fully. It is perfectly possible for a party to carry a district even with decreases in vote share, as long as the first loser in the previous election sees her vote share competed away even more by the second (and further) losers.

Second, our estimate of the effect of copartisan mayors may be indirectly capturing the effect of municipality size. For example, if the proportion of copartisan municipalities is large, but this obtains because there are only 3–4 municipalities in the district, 2–3 of which are controlled by copartisans, mayors would be mobilizing because they perceive themselves as pivotal, not because of the set of incentives highlighted by our model.[25] However, Table A10 in the Appendix shows that controlling for the number of municipalities in the district barely changes our estimates.

Third, a reasonable concern is that "ceiling effects" imply that congressional candidates competing in districts with copartisan neighbors already have a high vote share and, consequently, there is little that even highly-motivated neighboring mayors could do to increase chances of victory. We counter that (i) we have controlled for the party's normal vote, and (ii) these vote shares are for the most part not particularly large (between 0.35 and 0.55 is typical). Furthermore, if ceiling effects were present, we should find that after taking into account the presence (or lack thereof) of a copartisan governor, then *any* mayor, regardless of the number of copartisan mayors that surround him, would do his best to mobilize in favor of a copartisan congressional candidate. This criticism is unfounded because the probability that a congressional candidate will win *the mayor's own municipality* is actually lower, regardless of whether there is a copartisan governor (see Table A11 in the Appendix). We estimate this causal effect in a plain-vanilla regression-discontinuity design where the forcing variable is simply the mayor's margin of victory in her own electoral race. Prima facie, it seems paradoxical that *a group of mayors* can affect the probability of victory of a copartisan congressional candidate while *any single mayor* fails to achieve this feat. Consider however that the regression-discontinuity design pools *all* mayors together, *regardless of how many copartisan mayors exist in a district* (and, of course, as long as they were elected in very close races). Some of these will come from districts with very few copartisan mayors, others will come from districts with very many copartisan mayors, and their exertions on behalf of copartisans give rise to a single estimate of an average treatment effect. In contrast, when we look at *groups of mayors*, we consider explicitly the proportion of municipalities controlled

[25]As mentioned above, we weight municipalities by population, so a municipality that comprises half of a district's population receives a weight of 0.5, and so on.

by copartisan mayors in any given district. It is this conditional dampening effect that our theory uniquely predicts.

An alternative explanation for the dampening effect is that the efforts of copartisan governors and mayors are partial substitutes in their impact on voters. From the point of view of the average voter, the marginal effect of the first exhortation received from an incumbent politician might be larger than the marginal effect of the second exhortation received from another incumbent politician from the same political party. That is, in the absence of a copartisan governor, the first exhortation from a mayor would be likely to have a larger marginal effect on a voter's decision than the first exhortation from a mayor in a district with a copartisan governor. In other words, if the existence of a copartisan governor has already moved most voters to make up their minds on whether to vote and for whom, the additional effort of a copartisan mayor will be superfluous. This, rather than the collective action problem underscored by our theory, may be driving the dampening effect.[26]

One way to arbitrate between our explanation and the partial substitute hypothesis is to consider the average mayoral effect on *turnout* depending on whether (a) there is a copartisan governor; and (b) there are *few* or *many* copartisan mayors in the district. If the "partial substitute" hypothesis holds water, we should see that individual mayors exert a larger effect on turnout in the absence of a copartisan governor than in her presence (this is also an expectation of our theory), but this effect should *not* depend on the number of copartisan mayors in the other municipalities in the district: this is because mayors have jurisdiction over a single municipality, and thus the mobilizing efforts of two mayors cannot be partial substitutes for each other. To see whether this is the case, we use a regression discontinuity design to estimate the causal effect of municipal incumbency on turnout in four different sub-samples, depending on whether (a) there is a copartisan governor or not; and (b) there are "few" or "many" copartisan mayors in the district, excluding the municipality of interest. Specifically, we code the mayor of municipality m in district d as facing "many"

[26]We thank an anonymous reviewer for suggesting this possibility.

Table 3: RD estimates of the effect of PRI incumbency at the municipal level on *turnout* in federal elections (measured at the municipal level), 2000-2012

		Governor	
		Copartisan	Opposition
# Copartisan mayors in district	Few	0.070	–0.015
		(0.028)*	(0.025)
	Many	–0.049	0.005
		(0.022)*	(0.039)
p-value for test of equal means		0.001	0.682

* $p<0.05$; † $p<0.10$. Coefficients report the effect of PRI incumbency at the municipal level on *turnout* in federal elections in four different subsamples, depending on whether (a) the PRI controls the governorship; and (b) there are "few" (less than 60% of the district's population governed by copartisan mayors) or "many" PRI mayors in a district. Sample sizes from top-left, clockwise, are 545, 701, 224 and 879. The running variable is *margin of victory* in the municipal election held at $t-1$. Bias-corrected estimates are based on a local linear regression fitted separately at both sides of the threshold and employing a triangular kernel. The bandwidth is calculated according to the automatic selection procedure proposed by Calonico, Cattaneo and Titiunik (2014). Standard errors in parentheses.

copartisan mayors if at least 60% of d's population is governed by *other* mayors from the same party.[27] We restrict the analysis to the PRI because this is the only party for which we have enough observations for each subsample.[28]

Table 3 displays the results. In districts with opposition governors, we basically see no mayoral effect on turnout: the estimated coefficients are not only statistically indistinguishable from zero, but are small in substantive terms as well. Consistent with our logic, mayors have little incentive to mobilize since there is no local agent that can monitor and reward

[27]This reflects the fact that, in districts where the PRI controls at least one municipality, the median value of the (population-weighted) proportion of copartisan mayors in *other* municipalities is 0.58. For the PAN and PRD the corresponding values are much lower: 0.18 and 0.11, respectively.

[28]We have 1783 observations for the copartisan governor/many mayors subsample; 1409 for the copartisan governor/few mayors subsample; 371 for the opposition governor/many mayors combination; and 1074 for the opposition governor/few mayors subsample. However, RD analyses restrict the sample to close elections, so the estimates reported in Table 3 are based on fewer observations. For the PAN and PRD, on the other hand, the corresponding sample sizes are 112, 358, 1495 and 2672; and 494, 322, 851 and 2970, respectively; this means that when the sample is restricted to close elections, only a handful of observations remain for some combinations.

their efforts. In contrast, our explanation suggests that the mayoral effect on turnout will be positive but only in those municipalities governed by copartisan mayors surrounded by few copartisan municipalities (0.07, SE=0.03). In these places, the mayor mobilizes support precisely to show competence and in anticipation that his effort will be noticed since there are relatively less copartisan mayors to consider for promotion. In the presence of many additional copartisan mayors, we actually estimate a negative mayoral effect on PRI turnout in districts governed by a PRI governor (−0.05, SE=0.02). This suggests that mayors do not bother motivating turnout, which we attribute to their anticipation that the election will likely be won *and* that the governor would probably not notice their efforts anyway.

4 Conclusion

Under what conditions will local politicians exert effort on behalf of national candidates? To answer this question, we developed a theory based on three stylized facts about the relationship between local and national candidates. First, collaboration cannot be taken for granted. In federal and unitary polities alike, electoral districts and administrative circumscriptions may be nested in ways that generate collective action dilemmas. Second, monitoring the behavior of local party agents and elected politicians such as mayors is costly. Third, the existence of within-party hierarchies means that politicians higher in the party's chain of command can soften free-riding incentives because they can affect to some degree the careers of politicians that are lower in the hierarchy. In our setup, governors can promote good-type mayors, but our theory does not preclude alternative mechanisms; basically, if mayors depend at least somewhat on gubernatorial support to further their own careers, this gives governors (or other superiors in the party hierarchy) a lever that they can use to motivate local effort. In short, we have purported to understand when and how collaboration between local and national politicians within the same party becomes self-enforcing.

The electoral performance data we have analyzed is consistent with our claims: First, the proportion of copartisan mayors in a district is a statistically and substantively important predictor of whether a congressional candidate will win the district race. Second, the positive effect of having a large proportion of copartisan mayors in the district is attenuated in the presence of a copartisan governor.

Our theory makes three important contributions to our understanding of the relationship between copartisans at different levels of competition. First, it provides an explanation for why not all local politicians mobilize, even in close elections. In the setup that we analyze, mayors are less likely to exert effort if they compete against other copartisan mayors for a limited set of promotion opportunities and if they anticipate that their party cannot win the electoral district. Local politicians have the largest incentive to mobilize when victory cannot be taken for granted, but is reachable, and they do not have to compete with other copartisans for a shot at promotion. From the point of view of national party leaders, this is an infelicitous situation, as they would prefer to see all local politicians exert effort. Second, our theory suggests that governors and mayors seldom mobilize together; on the contrary, if there are reasons to expect that the party will carry an election easily, the governor prefers to monitor the mayors' effort rather than campaign. Third, empirical studies on whether candidates are helped by copartisans holding local offices tend to focus on one such office at a time — notably the governorship (Erikson and Titiunik 2015, Folke and Snyder 2012) or the local executive (Novaes 2018, Klašnja and Titiunik 2017) — and thus ignore the issue of whether controlling several offices at the same time is especially advantageous.[29] We have argued and provided evidence that, at least from the point of view of electoral performance, holding more offices is not necessarily advantageous for a political party.

This paper suggests at least two important avenues for further research. Most obviously, we need to extend our model in order to incorporate rules that feature prominently in other polities. Foremost among these is the possibility of mayoral reelection: to what extent would

[29]One exception is Grimmer et al. (2011), who explicitly argue, in contrast with our claims, that the probability that a candidate will win a close congressional race should be increasing in the proportion of state-level offices controlled by her party: the governorship, the state electoral administration, and the local and upper houses of the legislature.

reelection induce effort on behalf of copartisan candidates as a means of signaling strength and thus discouraging potential entrants from running against a mayoral incumbent? There is also the question of ideology: what would happen if the governor had an ideological affinity with certain mayors but not others? Ideology might introduce perverse incentives, discouraging effort from both aligned and unaligned mayors — the former because they expect to be promoted regardless of their performance, the latter because they do not expect to be promoted anyway — but it might also introduce a trade-off in the governor's utility function, e.g. if she has to weight the costs and benefits of promoting an ideologically-distant copartisan that is more likely to be of good type than (potentially) bad-type mayors that are closer in ideological terms.

Finally, we have omitted a discussion about the normative implications of partisan mobilization. Elected local politicians such as governors and mayors wear two hats. First and foremost, they are — or are supposed to be — public servants that are accountable to their constituents, and as such are expected to devote ample time and effort to fulfilling their governance roles. But they are also agents of the party that nominated and supported them for the position. These partisan and governmental roles can be reinforcing — for example, a mayor that provides good governance strengthens his party's reputation — but they can also be at loggerheads if acting as the party's agent means taking time away from governing functions. From a normative perspective, understanding what institutions strengthen the alignment between the *partisan* and *representative* roles of local politicians represents an important issue for further research.

References

Albertus, Michael. 2013. "Vote Buying with Multiple Distributive Goods." *Comparative Political Studies* 46(9):1082–1111.

Ames, Barry. 1994. "The Reverse Coattails Effect: Local Party Organization in the 1989 Brazilian Presidential Election." *American Political Science Review* 88(1):95–111.

Benton, Allyson L. 2007. "The Strategic Struggle for Patronage: Political Careers, State Largesse, and Factionalism in Latin American Parties." *Journal of Theoretical Politics* 19(1):55–82.

Besley, Timothy. 2007. *Principled Agents? The Political Economy of Good Government.* Lindahl Lectures Oxford, UK: Oxford University Press.

Broockman, David E. 2009. "Do Congressional Candidates Have Reverse Coattails? Evidence from a Regression Discontinuity Design." *Political Analysis* 17:418–434.

Burns, Peter F. 1999. "The Gubernatorial Coattail Hypothesis Considered: Determining the Effects of Gubernatorial Races on Senatorial Contests." *Southeastern Political Review* 27(1):151–172.

Caldeira, Gregory A., Aage R. Clausen and Samuel C. Patterson. 1990. "Partisan Mobilization and Electoral Participation." *Electoral Studies* 9(3):191–204.

Caldeira, Gregory A., Samuel C. Patterson and Gregory A. Markko. 1985. "The Mobilization of Voters in Congressional Elections." *Journal of Politics* 47(2):490–509.

Calonico, Sebastian, Matias D. Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6):2295–2326.

Calvo, Ernesto and Juan Manuel Abal Medina. 2002. "Institutional Gamblers: Majoritarian Representation, Electoral Uncertainty, and the Coalitional Costs of Mexico's Hybrid Electoral System." *Electoral Studies* 21(3):453–471.

Caramani, Daniele. 2004. *The Nationalization of Politics. The Formation of National Electorates and Party Systems in Western Europe*. Cambridge University Press.

Carey, John M. 2003. Presidentialism and Representative Institutions. In *Constructing Democratic Governance in Latin America*, ed. Jorge I. Domínguez. Baltimore, MD: Johns Hopkins University Press pp. 11–42.

Chhibber, Pradeep K. and Ken Kollman. 2004. *The Formation of National Party Systems. Federalism and Party Competition in Canada, Great Britain, India, and the United States.* Princeton: Princeton University Press.

Erikson, Robert S., Olle Folke and James M. Snyder. 2015. "A Gubernatorial Helping Hand? How Governors Affect Presidential Elections." *Journal of Politics* 77(2):491–504.

Erikson, Robert S. and Rocío Titiunik. 2015. "Using Regression Discontinuity to Uncover the Personal Incumbency Advantage." *Quarterly Journal of Political Science* 10:101–119.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50:5–25.

Folke, Olle. 2014. "Shades of Brown and Green: Party Effects in Proportional Election Systems." *Journal of the European Economic Association* 12(5):1361–1395.

Folke, Olle and James M. Snyder. 2012. "Gubernatorial Midterm Slumps." *American Journal of Political Science* 56(4):931–948.

Gans-Morse, Jordan, Sebastián Mazzuca and Simeon Nichter. 2014. "Varieties of Clientelism: Machine Politics during Elections." *American Journal of Political Science* 58(2):415–432.

Gibson, Edward L. 2012. *Boundary Control. Subnational Authoritarianism in Federal Democracies*. Cambridge Studies in Comparative Politics Cambridge, MA: Cambridge University Press.

Golder, Matt. 2006. "Presidential Coattails and Legislative Fragmentation." *American Journal of Political Science* 50(1):34–48.

Gosnell, Harold F. 1937. *Machine Politics: Chicago Model*. Chicago, IL: University of Chicago Press.

Green, Donald, Alan S. Gerber and David Nickerson. 2003. "Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments." *Journal of Politics* 65:1083–1096.

Green, Donald P. and Alan S. Gerber. 2004. *Get Out the Vote! How to Increase Voter Turnout*. Washington, DC: Brookings Institution Press.

Grimmer, Justin, Eitan Hersh, Brian Feinstein and Daniel P. Carpenter. 2011. "Are Close Elections Random?" Stanford University, unpublished manuscript.

Hogan, Robert E. 2005. "Gubernatorial Coattail Effects in State Legislative Elections." *Political Research Quarterly* 58(4):587–597.

Holbrook, Thomas M. and Scott D. McClurg. 2005. "The Mobilization of Core Supporters: Campaigns, Turnout, and Electoral Composition in United States Presidential Elections." *American Journal of Political Science* 49(4):689–703.

Huckfeldt, Robert and John Sprague. 1992. "Political Parties and Electoral Mobilization: Political Structure, Social Structure, and the Party Canvass." *American Political Science Review* 86(1):70–86.

Jones, Mark P. 1997. "Federalism and the Number of Parties in Argentine Congressional Elections." *Journal of Politics* 59(2):538–549.

Kemahlıoğlu, Özge. 2012. *Agents or Bosses? Patronage and Intra-Party Politics in Argentina and Turkey*. Colchester, Essex: ECPR Press.

Kerevel, Yann. 2015. "(Sub)national Principals, Legislative Agents. Patronage and Political Careers in Mexico." *Comparative Political Studies* 48(8):1020–1050.

Koelble, Thomas A. 1996. "Economic Theories of Organization and the Politics of Institutional Design in Political Parties." *Party Politics* 2(2):251–263.

Langston, Joy. 2017. *Democratization and Authoritarian Party Survival: Mexico's PRI*. Oxford, UK: Oxford University Press.

Leiras, Marcelo. 2006. Parties, Provinces and Electoral Coordination: A Study on the Determinants of Party and Party System Aggregation in Argentina, 1983–2005 PhD thesis University of Notre Dame Notre Dame, IN.

Magaloni, Beatriz, Alberto Díaz-Cayeros and Federico Estévez. 2007. The Erosion of Party Hegemony, Clientelism and Portfolio Diversification: The Programa Nacional de Solidaridad (Pronasol) in Mexico. In *Patrons or Policies*, ed. Herbert Kitschelt and Steven Wilkinson. Cambridge University Press chapter 6.

Magar, Eric. 2012. "Gubernatorial Coattails in Mexican Congressional Elections." *Journal of Politics* 74(2):383–399.

McCarty, Nolan and Adam Meirowitz. 2007. *Political Game Theory. An Introduction*. Analytical Methods for Social Research Cambridge, MA: Cambridge University Press.

Meredith, Marc. 2013. "Exploiting Friends-and-Neighbors to Estimate Coattail Effects." *American Political Science Review* 107(4):742–759.

Merriam, Charles E. and Harold F. Gosnell. 1924. *Non-Voting: Causes and Methods of Control*. Chicago, IL: University of Chicago Press.

Morgenstern, Scott and Javier Vázquez-D'Elía. 2007. "Electoral Laws, Parties, and Party Systems in Latin America." *Annual Review of Political Science* 10:143–168.

Morgenstern, Scott and Stephen M. Swindle. 2005. "Are Politics Local? An Analysis of Voting Patterns in 23 Democracies." *Comparative Political Studies* 38(2):143–170.

Morgenstern, Scott, Stephen M. Swindle and Andrea Castagnola. 2009. "Party Nationalization and Institutions." *Journal of Politics* 71(4):1322–1341.

Novaes, Lucas M. 2018. "Disloyal Brokers and Weak Parties." *American Journal of Political Science* 62(1):84–98.

Palfrey, Thomas R. and Howard Rosenthal. 1985. "Voter Participation and Strategic Uncertainty." *American Political Science Review* 79(1):62–78.

Patterson, Samuel C. and Gregory A. Caldeira. 1983. "Getting Out the Vote: Participation in Gubernatorial Elections." *American Political Science Review* 77(3):675–689.

Persico, Nicola, José Carlos Rodríguez Pueblita and Dan Silverman. 2011. "Factions and Political Competition." *Journal of Political Economy* 119(2):242–288.

Rodden, Jonathan and Erik Wibbels. 2011. "Dual Accountability and the Nationalization of Party Competition: Evidence from Four Federations." *Party Politics* 17(5):629–653.

Rosas, Guillermo and Joy Langston. 2011. "Gubernatorial Effects on the Voting Behavior of National Legislators." *Journal of Politics* 73(2):477–493.

Rundlett, Ashlea and Milan W. Svolik. 2016. "Deliver the Vote! Micromotives and Macrobehavior in Electoral Fraud." *American Political Science Review* 110(1):180–197.

Samuels, David J. 2000. "The Gubernatorial Coattails Effect: Federalism and Congressional Elections in Brazil." *Journal of Politics* 62(1):240–253.

Shachar, Ron and Barry Nalebuff. 1999. "Follow the Leader: Theory and Evidence on Political Participation." *American Economic Review* 89(2):525–547.

Sovey, Allison J. and Donald P. Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55(1):188–200.

Staiger, Douglas and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65(3):557–586.

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

Szwarcberg, Mariela. 2014. "Political Parties and Rallies in Latin America." *Party Politics* 20(3):456–466.

Klašnja, Marko and Rocío Titiunik. 2017. "The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability." *American Political Science Review* 111(1):129–148.

Van Houten, Pieter. 2009. "Multi-level Relations in Political Parties. A Delegation Approach." *Party Politics* 15(2):137–156.

Weitz-Shapiro, Rebecca. 2012. "What Wins Votes: Why Some Politicians Opt Out of Clientelism." *American Journal of Political Science* 56(3):568–583.

Zudenkova, Galina. 2011. "A Political Agency Model of Coattail Voting." *Journal of Public Economics* 95(11–12):1652–1660.

Jumping Ship or Jumping on the Bandwagon: When Do Local Politicians Support National Candidates?

Online Appendix

A Proofs and derivations

Proof of Lemma 1. G monitors one mayor at random, and upon monitoring she receives one of two signals: either the mayor is exerting effort or not. Because we restrict the analysis to "symmetric equilibria", we only consider situations in which all bad-type mayors employ the same decision rule. Under these assumptions, we first derive the posterior probability that a mayor is of good type, conditional on *not* observing effort:

$$\begin{aligned}\Pr(\text{good}|e=0) &= \frac{\Pr(e=0|\text{good})\Pr(\text{good})}{\Pr(e=0|\text{good})\Pr(\text{good})+\Pr(e=0|\text{bad})\Pr(\text{bad})}\\ &= \frac{0\cdot\tau}{0\cdot\tau+(1-\tau)(1-\sigma)}=0,\end{aligned}\tag{2}$$

where σ is the probability that a bad type will exert effort. When G fails to observe effort from a monitored mayor, she can infer with certainty that the mayor is a bad type; if G then decides to promote some other mayor, the probability of selecting a good type is $\tau>0$.

On the other hand, when G observes effort from a monitored mayor, she updates her beliefs about the mayor's type in the following manner:

$$\begin{aligned}\Pr(\text{good}|e=1) &= \frac{\Pr(e=1|\text{good})\Pr(\text{good})}{\Pr(e=1|\text{good})\Pr(\text{good})+\Pr(e=1|\text{bad})\Pr(\text{bad})}\\ &= \frac{\tau}{\tau+(1-\tau)\sigma}\geq\tau,\end{aligned}\tag{3}$$

with the inequality holding strictly as long as $\sigma<1$. Thus, upon observing that a mayor is exerting effort, the governor cannot do better than promoting him. ■

Proof of Proposition 1. The argument proceeds in two parts. We begin by showing that the mayor's optimal strategy is to campaign in favor of the copartisan candidate as long as $r\geq r^*$. We then show that, conditional on the mayor choosing e=1, the governor will also campaign on behalf of the copartisan candidate if $k \leq k^*$. The first part follows from comparing expected utilities $\mathsf{E}(U|e=1,c=m=0)$ and $\mathsf{E}(U|e=0,c=m=0)$ (we use U for the mayor's utility, V for the governor's utility). The only complication arises from the fact that these utilities comprise probabilities for different events, namely, the number of incumbent party mayors that exert effort because they are of good type. In a scenario where we need to guess the types of the other two candidates in the district, the probabilities that 0, 1, or 2 other mayors are of good type are, respectively, $(1-\tau)^2$, $2\tau(1-\tau)$, and τ^2, and the expected reward is $r/3$. With these quantities in hand, we can derive the proof of Proposition 1 for a scenario with three copartisan mayors:

$$\begin{aligned}
&\mathsf{E}(U|e=1,c=m=0) \geq \mathsf{E}(U|e=0,c=m=0)\\
&\underbrace{\left[3\frac{1-s}{3}\psi - v\psi + \frac{1}{2}\right]}_{\Pr(\text{Win}|e=1)}\frac{r}{3} - f \geq \underbrace{\left[(1-\tau)^2\left(\frac{1}{2}-v\psi\right)\right]}_{\Pr(\text{Win}|3\text{ bad types})}\frac{r}{3} +\\
&\qquad + \underbrace{\left[2\tau(1-\tau)\left(\frac{1}{2}+\frac{1-s}{3}\psi - v\psi\right)\right]}_{\Pr(\text{Win}|2\text{ bad types})}\frac{r}{3} +\\
&\qquad + \underbrace{\left[\tau^2\left(\frac{1}{2}+2\frac{1-s}{3}\psi - v\psi\right)\right]}_{\Pr(\text{Win}|1\text{ bad type})}\frac{r}{3}\\
&r_3^* \geq \frac{9f}{\psi(1-s)(3-2\tau)}.
\end{aligned} \tag{4}$$

The calculus for scenarios with one and two copartisan mayors is similar. Note that when there are two mayors, a bad-type mayor knows the probability that the other mayor is also a bad-type to be $(1-\tau)$, and he also needs to factor in the probability that opposition mayors are of good type, which explains the additional terms that depend on τ in the probabilities of victory. The calculus of expected utilities for the case with two copartisan mayors follows:

$$\begin{aligned}
&\mathsf{E}(U|e=1,c=m=0) \geq \mathsf{E}(U|e=0,c=m=0)\\
&\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{2} - f \geq \tau\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{2}\\
&\qquad + (1-\tau)\left[\psi\left(-\tau\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{2}\\
&r_2^* \geq \frac{6f}{\psi(1-s)(2-\tau)}.
\end{aligned} \tag{5}$$

A bad-type mayor in a 1-copartisan-mayor scenario campaigns if $\mathsf{E}(U|e=1,c=m=0)\geq \mathsf{E}(U|e=0,c=m=0)$, i.e., if

$$\left(\psi\left(\frac{1-s}{3}\right)-\psi\left[v+2\tau^2\left(\frac{1-s}{3}\right)+2\tau(1-\tau)\left(\frac{1-s}{3}\right)\right]+\frac{1}{2}\right)r-f\geq \left(\frac{1}{2}-\psi\left[v+2\tau^2\left(\frac{1-s}{3}\right)+2\tau(1-\tau)\left(\frac{1-s}{3}\right)\right]\right)r, \quad (6)$$

which obtains when $r_1^*\geq\dfrac{3f}{\psi(1-s)}$. Note that $r_1^*<r_2^*<r_3^*$.

The second part of the proof is as follows: First, if $r>r^*$, the reward is large enough that bad mayors always have an incentive to exert effort. In consequence, G does not have an incentive to monitor, which would at best reveal that the mayors are exerting effort. However, G has an incentive to campaign if the cost is low enough. This follows from comparison of expected utilities $\mathsf{E}(V|e=1,c=1)$ and $\mathsf{E}(V|e=1,c=0)$. In the scenario with three copartisan mayors, the comparison is as follows:

$$\begin{gathered}\mathsf{E}(V|e=1,c=1)\geq\mathsf{E}(V|e=1,c=0)\\ \underbrace{\left[s\psi+3\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]}_{\Pr(\text{Win}|e=1,c=1)}\tau-k\geq\underbrace{\left[3\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]}_{\Pr(\text{Win}|e=1,c=0)}\tau \\ k_3^*\leq\psi s\tau.\end{gathered}\tag{7}$$

Note that τ enters these expected utilities as the probability that the mayor that the governor taps for promotion is actually of good type. The calculus for the other two scenarios is similar. When there are two copartisan mayors, G campaigns if

$$\begin{gathered}\left[s\psi+(2-\tau)\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]\tau-k\geq\left[(2-\tau)\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]\tau\\ k_2^*\leq\psi s\tau,\end{gathered}\tag{8}$$

and, with a single copartisan mayor, G campaigns if

$$\begin{gathered}\left[s\psi+(1-2\tau)\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]\tau-k\geq\left[(1-2\tau)\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]\tau\\ k_1^*\leq\psi s\tau.\end{gathered}\tag{9}$$

Note that $k_1^*=k_2^*=k_3^*=k^*$, which means that effort-inducing cost k^* does not depend on the number of copartisan mayors in the district.

We finally show that bad-type mayors have no incentive to deviate from campaigning when the governor also campaigns. The mayor's calculus in a scenario with three copartisan mayors is as follows:

$$\begin{aligned}
\mathsf{E}(U|e=1,c=1) &\geq \mathsf{E}(U|e=0,c=1) \\
\left[s\psi+3\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]\frac{r}{3}-f &\geq \left[(1-\tau)^2\left(\frac{1}{2}+s\psi-v\psi\right)\right]\frac{r}{3}+ \\
&\quad+\left[2\tau(1-\tau)\left(\frac{1}{2}+s\psi+\left(\frac{1-s}{3}\right)-v\psi\right)\right]\frac{r}{3}+ \\
&\quad+\left[\tau^2\left(\frac{1}{2}+s\psi-2\left(\frac{1-s}{3}\right)\psi-v\psi\right)\right]\frac{r}{3} \\
r_3^{\dagger} &\geq \frac{9f}{\psi(1-s)(3-2\tau)} \equiv r_3^{*}.
\end{aligned} \tag{10}$$

Inspection of the relevant utilities in the other two scenarios confirms that $r_2^{\dagger}=r_2^{*}$ and $r_1^{\dagger}=r_1^{*}$. With two copartisan mayors:

$$\begin{aligned}
\left[\psi\left(s+\left(2-\tau\right)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{2}-f &\geq \tau\left[\psi\left(s+\left(1-\tau\right)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{2} \\
&\quad+(1-\tau)\left[\psi\left(s-\tau\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{2} \\
r_2^{\dagger} &\geq \frac{6f}{\psi(1-s)(2-\tau)} \equiv r_2^{*};
\end{aligned} \tag{11}$$

and with one copartisan mayor:

$$\begin{aligned}
\left[\psi\left(s+\left(1-2\tau\right)\left(\frac{1-s}{3}\right)-\nu\right)+\frac{1}{2}\right]r-f &\geq \left[\psi\left(s-2\tau\left(\frac{1-s}{3}\right)-\nu\right)+\frac{1}{2}\right]r \\
r_1^{\dagger} &\geq \frac{3f}{\psi(1-s)} \equiv r_1^{*}.
\end{aligned} \tag{12}$$

We confirm that threshold r^* alone determines whether bad-type mayors will have an incentive to mobilize in pure-strategy equilibria. ■

Proof of Proposition 2. We first prove that the governor will have no incentives to campaign (i.e., she will choose $c=0$) whenever $k \geq k^*$. We will then show that G might still have an incentive to monitor under these circumstances.

Consider first the scenario with three copartisan mayors:[30]

$$\begin{gathered}
\mathsf{E}(V|e=0,c=0) \geq \mathsf{E}(V|e=0,c=1) \\
3\tau(1-\tau)^2\left[\left(\frac{1-s}{3}\right)\psi - v\psi + \frac{1}{2}\right]\frac{1}{3} + 3\tau^2(1-\tau)\left[2\left(\frac{1-s}{3}\right)\psi - v\psi + \frac{1}{2}\right]\frac{2}{3} + \\
\tau^3\left[3\left(\frac{1-s}{3}\right)\psi - v\psi + \frac{1}{2}\right] \quad \geq \\
3\tau(1-\tau)^2\left[\psi\left(s + \frac{1-s}{3} - v\right) + \frac{1}{2}\right]\frac{1}{3} + 3\tau^2(1-\tau)\left[\psi\left(s + 2\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right]\frac{2}{3} + \\
\tau^3\left[\left(s + 3\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right] - k \\
k \geq \psi s \tau \equiv k^*.
\end{gathered} \tag{13}$$

Similarly, a comparison of utilities in the 2- and 1-copartisan-mayor scenarios reveals that G will not campaign if $k \geq k^*$. For the 2-copartisan-mayor scenario:

$$\begin{gathered}
\mathsf{E}(V|e=0,c=0) \geq \mathsf{E}(V|e=0,c=1) \\
\tau^2\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right] + 2\tau(1-\tau)\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right]\frac{1}{2} \geq \\
\tau^2\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right) + s - v\right) + \frac{1}{2}\right] + 2\tau(1-\tau)\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right) + s - v\right) + \frac{1}{2}\right]\frac{1}{2} - k \\
k \geq \psi s \tau \equiv k^*;
\end{gathered} \tag{14}$$

while in the 1-copartisan mayor scenario:

$$\begin{gathered}
\mathsf{E}(V|e=0,c=0) \geq \mathsf{E}(V|e=0,c=1) \\
\psi\left(\frac{1}{2} - v - 2\tau\left(\frac{1-s}{3}\right)\right)\tau \geq \psi\left(\frac{1}{2} - v - 2\tau\left(\frac{1-s}{3}\right) + s\right)\tau - k \\
k \geq \psi s \tau \equiv k^*.
\end{gathered} \tag{15}$$

We have shown that G never has an incentive to campaign if $k > k^*$. We now consider whether there exists an incentive to monitor under different values of k. We consider two

[30]Here, $3\tau(1-\tau)^2$ captures the probability that exactly 1 copartisan mayor will be of good type, therefore exerting effort. The probabilities of exactly 2 and exactly 3 good types are similarly defined as $3\tau^2(1-\tau)$ and τ^3, respectively.

cases, depending on whether k is larger or smaller than k^*.

Case 1: $k > k^*$. We will define v_3^* as the value of v that makes G (weakly) prefer monitoring in a 3-copartisan mayor scenario:[31]

$$\begin{aligned}
&\mathsf{E}(V|m=1,e=0) \geq \mathsf{E}(V|c=m=0,e=0)\\
&\qquad (1-\tau)^3\left(-v\psi+\frac{1}{2}\right)\cdot 0+3\tau(1-\tau)^2\left(\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right)\frac{2}{3}+\\
&\qquad\qquad +3\tau^2(1-\tau)\left(2\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right)\cdot 1+\tau^3\left(3\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right)\cdot 1-g\geq\\
&\qquad (1-\tau)^3\left(-v\psi+\frac{1}{2}\right)\cdot 0+3\tau(1-\tau)^2\left(\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right)\frac{1}{3}+\\
&\qquad\qquad 3\tau^2(1-\tau)\left(2\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right)\frac{2}{3}+\tau^3\left(3\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right)\cdot 1\\
&\qquad\qquad v_3^* \leq \frac{1}{2\psi}-\frac{g}{\tau(1-\tau)\psi}+\left(\frac{1-s}{3}\right)(1+\tau).
\end{aligned} \tag{16}$$

The value of v_2^* corresponding to a scenario with two mayors follows:

$$\begin{gathered}
\mathsf{E}(V|m=1,e=0) \geq \mathsf{E}(V|c=m=0,e=0)\\
\tau^2\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\cdot 1+2\tau(1-\tau)\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\cdot 1-g\geq\\
\tau^2\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\cdot 1+2\tau(1-\tau)\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{1}{2}\\
v_2^* \leq \frac{1}{2\psi}-\frac{g}{\tau(1-\tau)\psi}+\left(\frac{1-s}{3}\right)(1-\tau).
\end{gathered} \tag{17}$$

In a 1-copartisan-mayor scenario, G never has an incentive to monitor. The comparison of relevant utilities reveals that the governor would monitor only if this action were costless — there is no value of v that provides G with an incentive to monitor a potential bad-type for

[31] The structure of these expected utilities looks daunting, but each term contains the product of three elements: (a) the probability that there are exactly n good-type copartisan mayors; (b) the probability of winning given that there are exactly n good-type copartisan mayors; and (c) the posterior probability of promoting a good-type mayor conditional on G choosing to monitor or not.

whom there are no substitutes because there are no other copartisan mayors to promote:

$$\begin{aligned}\mathsf{E}(V|m=1,e=0) &\geq \mathsf{E}(V|c=m=0,e=0)\\ \tau\left[\left(\frac{1-s}{3}\right)\psi-v\psi-2\tau\left(\frac{1-s}{3}\right)+\frac{1}{2}\right]-g &\geq \tau\left[\left(\frac{1-s}{3}\right)\psi-v\psi-2\tau\left(\frac{1-s}{3}\right)+\frac{1}{2}\right]\\ 0 &\geq g. \end{aligned} \tag{18}$$

The last statement contradicts the assumption that $g>0$. Values of $v\leq v^*$ lead to equilibria in which G might be better off choosing $m=1$ rather than choosing $c=m=0$. The *headhunter monitoring* equilibrium further requires a very low reward, which removes any motivation that bad types may have to exert effort in the presence of monitoring. The threshold r' under which mayors will not exert effort follows from comparing utilities, first for the scenario with three mayors:

$$\begin{aligned}\mathsf{E}(U|e=1,m=1) &= \mathsf{E}(U|e=0,m=1)\\ \left(3\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right)\frac{r}{3}-f &\geq 2\tau(1-\tau)\left(\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right)\frac{r}{6}+(1-\tau)^2\left(\frac{1}{2}-\psi v\right)\frac{r}{3}\\ r'_3 &= \frac{3f}{\psi\left(\frac{1-s}{3}\right)[3-\tau+\tau^2]+\tau\left(\frac{1}{2}-v\psi\right)}. \end{aligned} \tag{19}$$

The relevant comparison of utilities for the scenario with two mayors follows:

$$\begin{aligned}\mathsf{E}(U|e=1,m=1) &= \mathsf{E}(U|e=0,m=1)\\ \left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{2}-f &\geq (1-\tau)\left[\psi\left(-\tau\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{2}\\ r'_2 &= \frac{2f}{\psi\left(\frac{1-s}{3}\right)(2-\tau^2)+\tau\left(\frac{1}{2}-v\psi\right)}. \end{aligned} \tag{20}$$

Because parameters s and τ range between 0 and 1, and because of the restrictions on v and

ψ (which imply that $1/2-v\psi$ is positive), it is easy to verify that $r_2' \leq r_2^*$ and $r_3' \leq r_3^*$:

$$\begin{aligned}
r_3^* &\geq r_3' \\
\frac{9f}{\psi(1-s)(3-2\tau)} &\geq \frac{3f}{\psi\left(\frac{1-s}{3}\right)[3-\tau+\tau^2]+\tau\left(\frac{1}{2}-v\psi\right)} \\
\frac{3}{\psi(1-s)(3-2\tau)} &\geq \frac{1}{\frac{1}{3}\psi(1-s)[3-\tau+\tau^2]+\tau\left(\frac{1}{2}-v\psi\right)} \\
3\cdot\frac{1}{3}\psi(1-s)[3-\tau+\tau^2]+3\tau\left(\frac{1}{2}-v\psi\right) &\geq \psi(1-s)(3-2\tau) \\
\psi(1-s)(3-\tau+\tau^2)-\psi(1-s)(3-2\tau) &\geq -3\tau\left(\frac{1}{2}-v\psi\right) \\
\psi(1-s)(\tau+\tau^2) &\geq -3\tau\left(\frac{1}{2}-v\psi\right) \\
\psi(1-s)\tau(1+\tau) &\geq -3\tau\left(\frac{1}{2}-v\psi\right) \\
\underbrace{\psi(1-s)(1+\tau)}_{>0} &\geq \underbrace{-3}_{<0}\underbrace{\left(\frac{1}{2}-v\psi\right)}_{>0}.
\end{aligned} \tag{21}$$

The proof that $r_2^* \geq r_2'$ is analogous. Finally, there are values of r under which a bad-type would exert effort in a scenario where he is the single copartisan mayor and G monitors. However, since G never has an incentive to monitor in these circumstances, effort by a bad-type mayor cannot be sustained in equilibrium.

Case 2: $k<k^*$. We established that G will prefer $c=1$ over $c=m=0$, but we have not considered the possibility that she will prefer $m=1$ over $c=1$, which could happen when k is low. With three copartisan mayors, this occurs whenever $v>v_3'$:

$$\mathsf{E}(V|m=1,e=0) \geq \mathsf{E}(V|c=1,e=0)$$

$$\begin{aligned}
&3\tau(1-\tau)^2\left[\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]\frac{2}{3}+3\tau^2(1-\tau)\left[2\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]+ \\
&\qquad \tau^3\left[3\left(\frac{1-s}{3}\right)\psi-v\psi+\frac{1}{2}\right]-g \geq \\
&3\tau(1-\tau)^2\left[\psi\left(s+\frac{1-s}{3}-v\right)+\frac{1}{2}\right]\frac{1}{3}+3\tau^2(1-\tau)\left[\psi\left(s+2\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{2}{3}+ \\
&\qquad \tau^3\left[\psi\left(s+3\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]-k
\end{aligned}$$

$$v_3' \leq \left(\frac{1-s}{3}\right)(1+\tau) - \frac{s}{1-\tau} + \frac{1}{2\psi} + \frac{k-g}{\tau(1-\tau)\psi}. \tag{22}$$

With two copartisan mayors, G prefers $m=1$ to $c=1$ when:

$$\begin{gathered}
\mathsf{E}(V|m=1,e=0) \geq \mathsf{E}(V|c=1,e=0) \\
\tau^2\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]+2\tau(1-\tau)\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]-g \geq \\
\tau^2\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)+s-v\right)+\frac{1}{2}\right]+2\tau(1-\tau)\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right)+s-v\right)+\frac{1}{2}\right]\frac{1}{2}-k \\
v_2' \leq \left(\frac{1-s}{3}\right)(1-\tau) - \frac{s}{1-\tau} + \frac{1}{2\psi} + \frac{k-g}{\tau(1-\tau)\psi}.
\end{gathered} \tag{23}$$

Note that $v_3' > v_2'$. Finally, G has no incentive to choose monitoring over campaigning in scenarios with one mayor. To see this, consider G's utility calculus:

$$\begin{gathered}
\mathsf{E}(V|m=1,e=0) \geq \mathsf{E}(V|c=1,e=0) \\
\tau\left(\psi\left(\left(\frac{1-s}{3}\right)-v-2\tau^2\left(\frac{1-s}{3}\right)-2\tau(1-\tau)\left(\frac{1-s}{3}\right)\right)+\frac{1}{2}\right)-g \geq \\
\tau\left(\psi\left(s+\left(\frac{1-s}{3}\right)-2\tau^2\left(\frac{1-s}{3}\right)-2\tau(1-\tau)\left(\frac{1-s}{3}\right)\right)+\frac{1}{2}\right)-k \\
k-\psi s\tau \geq g.
\end{gathered} \tag{24}$$

Because we are inspecting a situation where $k<k^*$, we know that $k-\psi s\tau<0$, which means that the statement $k-\psi s\tau \geq g$ cannot be true since g is strictly positive. ■

Proof of Proposition 3. From Proposition 1, we know that exerting effort is not a dominant strategy for mayors when $r<r^*$. We will now assume that $v \geq \max\{v_3', v_3^*\}$, i.e., that the opposition's normal vote in the district is relatively high (which need not mean that $v>0$). From Proposition 2 we know that the governor has no incentive to monitor when opposition support is high, and therefore mayors have no incentive to exert effort. The governor may still have an incentive to campaign, though, and this will happen under the following circumstances:

$$\mathsf{E}(V|c=1,e=0) \geq \mathsf{E}(V|c=0,e=0)$$

$$\tau^3\left[\psi\left(s+3\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]+3\tau^2(1-\tau)\left[\psi\left(s+2\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{2}{3}+$$
$$3\tau(1-\tau)^2\left[\psi\left(s+\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{1}{3}-k\geq$$
$$\tau^3\left[\psi\left(3\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]+3\tau^2(1-\tau)\left[\psi\left(2\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{2}{3}+$$
$$3\tau(1-\tau)^2\left[\psi\left(\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{1}{3}$$
$$k_3''\leq s\psi\tau=k^*. \tag{25}$$

For a scenario with two mayors:

$$\mathsf{E}(V|c=1,e=0)\geq\mathsf{E}(V|c=0,e=0)$$
$$\tau^2\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)+s-v\right)+\frac{1}{2}\right]+2\tau(1-\tau)\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right)+s-v\right)+\frac{1}{2}\right]\frac{1}{2}-k\geq$$
$$\tau^2\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]+2\tau(1-\tau)\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{1}{2}$$
$$k_2''\leq s\psi\tau=k^*. \tag{26}$$

And in the one-mayor scenario:

$$\mathsf{E}(V|c=1,e=0)\geq\mathsf{E}(V|c=0,e=0)$$
$$\tau\left[\psi\left(s+\left(\frac{1-s}{3}\right)-2\tau^2\left(\frac{1-s}{3}\right)-2\tau(1-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]-k\geq$$
$$\tau\left[\psi\left(\left(\frac{1-s}{3}\right)-2\tau^2\left(\frac{1-s}{3}\right)-2\tau(1-\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]$$
$$k_1''\leq s\psi\tau=k^*. \tag{27}$$

The next step is to check that bad-type mayors still lack an incentive to exert effort when the governor campaigns. This follows from comparison of utilities:

$$\begin{aligned}\mathsf{E}(U|e=1,c=1) &\geq \mathsf{E}(U|e=0,c=1)\\ \left[\psi\left(s+3\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{3}-f &\geq \tau^2\left[\psi\left(s+2\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{3}+\\ &\quad 2\tau(1-\tau)\left[\psi\left(s+\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{3}\\ &\quad +(1-\tau)^2\left[\psi(s-v)+\frac{1}{2}\right]\frac{r}{3}\\ r &\geq \frac{9f}{\psi(1-s)(3-2\tau)}=r_3^*.\end{aligned}$$

Bad types still lack an incentive to exert effort in a scenario with two copartisan mayors:

$$\begin{aligned}\mathsf{E}(U|e=1,c=1) &\geq \mathsf{E}(U|e=0,c=1)\\ \left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right)+s-v\right)+\frac{1}{2}\right]\frac{r}{2}-f &\geq \tau\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right)+s-v\right)+\frac{1}{2}\right]\frac{r}{2}\\ &\quad +(1-\tau)\left[\psi\left(-\tau\left(\frac{1-s}{3}\right)+s-v\right)+\frac{1}{2}\right]\frac{r}{2}\\ r &\geq \frac{6f}{\psi(1-s)(2-\tau)}=r_2^*,\end{aligned} \tag{28}$$

and the same is true in the one-mayor scenario:

$$\begin{aligned}\mathsf{E}(U|e=1,c=1) &\geq \mathsf{E}(U|e=0,c=1)\\ &\left[\psi\left(s+(1-2\tau)\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]r-f \geq\\ &\left[\psi\left(s-2\tau\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]r\\ r &\geq \frac{3f}{\psi(1-s)}=r_1^*.\end{aligned} \tag{29}$$

These derivations confirm that the bad-type mayor will mobilize only if $r \geq r^*$, but this contradicts the condition $r < r^*$ under which the *selfless governor* equilibrium exists.

Finally, we assumed that the *selfless governor* equilibrium can obtain when the normal vote for the opposition is sufficiently low, i.e., when $v < \max\{v', v^*\}$. We now show that $k < k^* \Rightarrow v' < v^*$ and $k > k^* \Rightarrow v' > v^*$. We do this by showing that $v'_3 = v^*_3 \Leftrightarrow k = k^*_3$ (the proof that $v_2' = v_2^* \Leftrightarrow k = k_2^*$ is basically identical):

$$
\begin{aligned}
v_3^* &= v_3' \\
\frac{1}{2\psi} - \frac{g}{\tau(1-\tau)\psi} + \left(\frac{1-s}{3}\right)(1+\tau) &= \frac{1}{2\psi} + \frac{k-g}{\tau(1-\tau)\psi} + \left(\frac{1-s}{3}\right)(1+\tau) - \frac{s}{1-\tau} \\
\frac{k}{\tau(1-\tau)\psi} &= \frac{s}{1-\tau} \\
k &= \psi s \tau. \qquad (30)
\end{aligned}
$$

■

Proof of Proposition 4. The first step is to show that no combination of pure strategies constitutes an equilibrium, and thus the game can *only* have equilibria in mixed strategies. We derive this step for the scenario with three mayors, but the proof for the scenario with two mayors follows along similar lines. We proceed in three parts: (i) Proposition 2 implies that if $v < v^*$ or $v < v'$ and bad mayors play $e = 0$, the governor always prefers monitoring to campaigning. This rules out all equilibria in which the governor always plays $c = 1$, as well as any equilibrium in which $e = 0$ and $m = 0$. (ii) From Proposition 1 we know that an equilibrium in which bad mayors always play $e = 1$ while the governor always plays $m = 0$ is only possible if $r > r_3^*$; therefore, the pair of strategies $(e = 1, m = 0)$ cannot be an equilibrium either. (iii) Assume that bad mayors always play $e = 1$. In this case, the governor will respond with $m = 1$ iff

$$
\begin{aligned}
E(V|m=1, e=1) &> E(V|m=0, e=1) \\
\left[\psi\left(3\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right]\tau - g &> \left[\psi\left(3\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right]\tau \\
0 &> g, \qquad (31)
\end{aligned}
$$

which contradicts the assumption that $g > 0$, and thus rules out the pair $(e = 1, m = 1)$ as an equilibrium choice. (iv) Finally, for $(e = 0, m = 1)$ to be an equilibrium, it must be the case that $E(U|m = 1, e = 0) > E(U|m = 1, e = 1)$; but from the proof of Proposition 2, we know that this can only be the case if $r < r_3'$. Thus, no combination of pure strategies can be sustained in equilibrium when $r_3' < r < r_3^*$ *and* $v < v_3^*$ or $v < v_3'$ A similar proof holds for a scenario with two mayors. In the case of a single mayor, Propositions 2 and 3 imply that the governor never has an incentive to monitor, which means that $m=1$ cannot be part of a mixed strategy equilibrium; in the one-copartisan scenario, no equilibrium in mixed strategies is thus possible. This completes the proof that under the set of assumptions established for Proposition 4, there are no equilibria in pure strategies.

The second step requires showing that there is *at least one* equilibrium in mixed strategies under the conditions specified by Proposition 4. McCarty and Meirowitz (2007) show that Bayesian normal form games are special cases of normal form games, and thus they have at least one (Bayesian) Nash equilibrium in mixed strategies (see Proposition 6.1, pp. 168-9). Specifically, the requirement for such an equilibrium to exist is that (i) the set of players, (ii) the set of feasible strategies and (iii) the set of players' types are all finite. Under the parameter restrictions specified by the assumptions of Proposition 4, the game satisfies these conditions, and thus it must have *at least one* equilibrium in mixed strategies.

The third step requires showing that there is *at most one* equilibrium in mixed strategies. To do so, we first show that the probabilities with which G and the bad type mix depend on r, f and g. Let π and $(1-\pi)$ be the probabilities with which G mixes between monitoring $(m=1,\ c=0)$ and not monitoring $(m=0,\ c=0)$, respectively. Let π^* be the value of π that makes bad types indifferent between exerting and not exerting effort, i.e., $E(U|e=1, m=\pi) = E(U|e=0, m=\pi)$. With three copartisan mayors, the equality takes the following form:

$$\begin{aligned}
&\left[\psi\left(3\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right]\frac{r}{3}-f= \\
&\qquad 2\tau(1-\tau)\left(\psi\left(\frac{1-s}{3}-v\right)+\frac{1}{2}\right)\frac{\pi r}{6}+(1-\tau)^2\left(-\psi v+\frac{1}{2}\right)\frac{\pi r}{3}+ \\
&+\left[\tau^2\left(\psi\left(2\left(\frac{1-s}{3}\right)-v\right)+\frac{1}{2}\right)+2\tau(1-\tau)\left(\psi\left(\frac{1-s}{3}-v\right)+\frac{1}{2}\right)+(1-\tau)^2\left(-\psi v+\frac{1}{2}\right)\right](1-\pi)\frac{r}{3}
\end{aligned} \tag{32}$$

After algebraic manipulation, the equation reduces to the following expression:

$$\pi_3^* = \frac{\frac{3f}{\psi r} - (3-2\tau)\frac{1-s}{3}}{\left[\frac{1-s}{3}(1+\tau) + \frac{1}{2\psi} - v\right]\tau} \tag{33}$$

Since the denominator is always positive, π_3^* will be in the unit-range if

$$\left[\left(\frac{1-s}{3}\right)(1+\tau) + \frac{1}{2\psi} - v\right]\tau > \frac{3f}{\psi r} - (3-2\tau)\left(\frac{1-s}{3}\right) > 0. \tag{34}$$

Similarly, in the 2-mayor case the equilibrium value of π is given by

$$\begin{aligned}
&\left[\psi\left((2-\tau)\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right]\frac{r}{2} - f = \\
&\quad \tau\left[\psi\left((1-\tau)\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right](1-\pi)\frac{r}{2} + (1-\tau)\left[\psi\left(-\tau\left(\frac{1-s}{3}\right) - v\right) + \frac{1}{2}\right]\frac{r}{2} \\
&\qquad \pi\tau\left[(1-\tau)\left(\frac{1-s}{3}\right) + \frac{1}{2\psi} - v\right] = \frac{2f}{r\psi} - (2-\tau)\left(\frac{1-s}{3}\right),
\end{aligned} \tag{35}$$

which reduces to:

$$\pi_2^* = \frac{\frac{2f}{r\psi} - (2-\tau)\frac{1-s}{3}}{\left[(1-\tau)\frac{1-s}{3} + \frac{1}{2\psi} - v\right]\tau}. \tag{36}$$

Since the denominator is always positive, π_2^* will be in the unit range as long as

$$\left[\left(\frac{1-s}{3}\right)(1-\tau) + \frac{1}{2\psi} - v\right]\tau > \frac{2f}{r\psi} - (2-\tau)\left(\frac{1-s}{3}\right) > 0. \tag{37}$$

On the other hand, let ρ and $(1-\rho)$ be the probabilities with which the mayor mixes between exerting and not exerting effort ($e=1$ and $e=0$, respectively) in order to make G indifferent between monitoring and not monitoring, i.e., in order to yield $\mathsf{E}(V|m=1, e=\rho) = \mathsf{E}(V|m=0, e=\rho)$. With three copartisan mayors:[32]

[32] We consider only symmetric equilibria (see fn. 12 in the main text), which means that when one bad type campaigns, all bad types campaign. Yet, inside the first curly bracket in Expression 38 we have terms for effort exerted by two bad types (ρ^2), one bad type ($2\rho(1-\rho)$), or none ($(1-\rho)^2$). This is because, in keeping with the analysis of symmetric equilibria, all bad types are playing a mixed strategy in which each of them exerts effort, independently, with probability ρ.

$$3\tau(1-\tau)^2\Big\{\Big[\psi\Big(3\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{1}{3}\rho^2+\Big[\psi\Big(2\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{1}{2}2\rho(1-\rho)$$
$$+\Big[\psi\Big(\frac{1-s}{3}-v\Big)+\frac{1}{2}\Big]\frac{2}{3}(1-\rho)^2\Big\}$$
$$+3\tau^2(1-\tau)\Big\{\Big[\psi\Big(3\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{2}{3}\rho+\Big[\psi\Big(2\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big](1-\rho)\Big\}$$
$$+\tau^3\Big\{\psi\Big(3\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big\}-g=$$
$$3\tau(1-\tau)^2\Big\{\Big[\psi\Big(3\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{1}{3}\rho^2+\Big[\psi\Big(2\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{1}{3}2\rho(1-\rho)$$
$$+\Big[\psi\Big(\frac{1-s}{3}\Big)+\frac{1}{2}\Big]\frac{1}{3}(1-\rho)^2\Big\}$$
$$+3\tau^2(1-\tau)\Big\{\Big[\psi\Big(3\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{2}{3}\rho+\Big[\psi\Big(2\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{2}{3}(1-\rho)\Big\}$$
$$+\tau^3\Big\{\psi\Big(3\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big\} \quad (38)$$

After rearranging terms, this expression reduces to:

$$\rho^2\underbrace{(-1)(1-\tau)\Big(\frac{1-s}{3}\Big)}_{a<0}+\rho\underbrace{(-1)\Big[\frac{1}{2\psi}-v+2\Big(\frac{1-s}{3}\Big)\tau\Big]}_{b<0}+\underbrace{\frac{1}{2\psi}-v+\Big(\frac{1-s}{3}\Big)(1+\tau)-\frac{g}{\tau(1-\tau)\psi}}_{c}=0,$$

and thus the equilibrium value of ρ will be given by

$$\rho_3^*=\frac{\overbrace{-b}^{(+)}\pm\overbrace{\sqrt{b^2-4ac}}^{(+)}}{\underbrace{2a}_{(-)}}, \quad (39)$$

which can *at most* take one value within the $(0,1)$ interval. On the other hand, in the 2-mayor case the equilibrium value of ρ is given by

$$\tau^2\Big[\psi\Big((2-\tau)\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]+2\tau(1-\tau)\Big[\psi\Big((2-\tau)\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{1}{2}\rho$$
$$+2\tau(1-\tau)\Big[\psi\Big((1-\tau)\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big](1-\rho)-g=$$
$$\tau^2\Big[\psi\Big((2-\tau)\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]+2\tau(1-\tau)\Big[\psi\Big((2-\tau)\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{1}{2}\rho$$
$$+2\tau(1-\tau)\Big[\psi\Big((1-\tau)\Big(\frac{1-s}{3}\Big)-v\Big)+\frac{1}{2}\Big]\frac{1}{2}(1-\rho),$$

$$(40)$$

which reduces to

$$\rho\left[v-\frac{1}{2\psi}-(1-\tau)\left(\frac{1-s}{3}\right)\right]=\frac{g}{\tau(1-\tau)\psi}+v-\frac{1}{2\psi}-(1-\tau)\left(\frac{1-s}{3}\right)$$

$$\rho_2^*=\frac{g}{\tau(1-\tau)\psi\underbrace{\left[v-\frac{1}{2\psi}-(1-\tau)\left(\frac{1-s}{3}\right)\right]}_{<0}}+1, \tag{41}$$

which takes a unique value. ■

B A potential identification strategy

We propose a causal identification strategy based on Folke's (2014) pioneering idea of employing the number of seats that a party barely wins or loses in a high-magnitude proportional representation district as an instrument for the total number of seats that a party captures in a district. Adapting this insight to Mexico's political system, we use the proportion of municipalities that a party *barely wins or loses* in a district as an instrument for the proportion of copartisan mayors that the party controls in that district.[33] The rationale behind this approach is that as the distance between the winner and first loser of a mayoral race goes to zero, the partisan status of the winning mayor is assigned *as if* randomly (this is the same principle exploited by regression discontinuity designs). In any given district, a varying number of municipalities may be closely won or lost, yielding a measure of the number of municipalities whose partisan status could be seen as exogenously determined. While a party's electoral performance in a district cannot be considered exogenous, a given party will win more close municipal races in some districts than in others, introducing variation in the proportion of copartisan mayors it controls in similar districts. That is, our identification strategy is based on comparing districts that differ in the proportion of municipalities controlled by copartisans solely because that party won (or lost) a relatively large number of close municipal races. Instrumenting for the proportion of copartisan mayors in this way satisfies the exclusion restriction as long as we control for the proportion of close races in a district, which cannot be characterized as exogenous.

Consider the following example of two districts, A and B, each with four municipalities numbered 1 through 4. Table A1 shows hypothetical vote shares for mayoral candidates in the eight municipalities. In District A, the PRI has three mayors and the PAN has one, whereas in District B one municipality belongs to the PRI, one to the PAN, and two to the PRD. We want to estimate the effect of the number of copartisan mayors on the vote share for party j's candidate at the district level, that is: Vote $\text{Share}_j = f(\#\text{Mayors}_j)$.

[33] We alternatively define close elections as those decided by less than 5 or 2.5 percent points.

Table A1: Vote shares for mayoral candidates in eight municipalities

	District A		
	PRI	PAN	PRD
A1	37.0	36.5	26.5
A2	33.2	33.4	33.3
A3	60.0	20.0	20.0
A4	50.2	49.7	0.1
# captured municipalities	3	1	0

	District B		
	PRI	PAN	PRD
B1	50.0	25.0	25.0
B2	49.0	1.7	49.3
B3	22.0	42.0	36.0
B4	30.0	0.0	70.0
# captured municipalities	1	1	2

To instrument #Mayors, we consider the number of copartisan mayors — i.e., "captured municipalities" — in the district that won their elections by a margin of less than five percentage points. Parties that barely lost a mayor (by a margin of 5pp or less) get a $-\frac{1}{2}$ and parties that barely won get $\frac{1}{2}$.[34] The instrument at the district level is the sum of barely lost/barely won mayors, *but the summands are weighted by the municipality's share of the district population.* Table A2 shows the values of the instrumental variable for the three parties in the two districts (w_d is the municipality's share of the total district population, so $\Sigma_d w_d = 1$):

Table A2: Construction of instrumental variable in eight municipalities

	District A		
	PRI	PAN	PRD
A1	$\frac{1}{2}w_1$	$-\frac{1}{2}w_1$	0
A2	$-\frac{1}{2}w_2$	$\frac{1}{2}w_2$	$-\frac{1}{2}w_2$
A3	0	0	0
A4	$\frac{1}{2}w_4$	$-\frac{1}{2}w_4$	0
Instrument	$\frac{w_1-w_2+w_4}{2}$	$\frac{-w_1+w_2-w_4}{2}$	$\frac{-w_2}{2}$

	District B		
	PRI	PAN	PRD
B1	0	0	0
B2	$-\frac{1}{2}w_2$	0	$\frac{1}{2}w_2$
B3	0	0	0
B4	0	0	0
Instrument	$\frac{-w_2}{2}$	0	$\frac{w_2}{2}$

We use the district-level instrument in a 2SLS regression setup. The first-stage is

$$\#\text{Mayors}_{d,t} = \alpha + \lambda \sum c_{m,t-1} \cdot w_{m,d} + \gamma \sum v_{m,t-1} \cdot w_{m,d} + g\Big(\sum \mathbf{V}_{m,d,t-1}\Big) + \delta_t + \varepsilon_{d,t},$$

[34] We follow Folke (2014) in writing (minus) $\frac{1}{2}$ rather than (minus) 1 because we include an additional control variable that takes the value of $\frac{1}{2}$ whenever there was a close race in a district. Thus, we have $\frac{1}{2} + \frac{1}{2} = 1$ if there was a close municipal race that party j won, and $\frac{1}{2} - \frac{1}{2} = 0$ if there was a close race that party j lost.

where d indexes districts, m indexes municipalities (nested within districts), t indexes election years, c takes the value of $\frac{1}{2}$ if the immediately preceding election (at $t-1$) in municipality m was close, v takes the value of (minus) $\frac{1}{2}$ if the party (lost) won the close election in municipality m at $t-1$, $w_{m,d}$ indicates municipality m's share of district d's population, $g(\sum \mathbf{V}_{m,d,t-1})$ is a fourth-order polynomial of the party's *normal vote share*, and δ_t is a year fixed effect. To continue with the example above, assume that $w_d = 0.25 \; \forall \, d$ (i.e., all four municipalities within both districts have equal population). Table A3 shows the population-weighted number of mayors elected *as if* randomly (the instrument that we use), the population-weighted number of close elections for the party in the district, and the population-weighted number of actual mayors (the endogenous variable that we instrument):

Table A3: *Instrument*, *close elections*, and *observed mayors* in eight municipalities

	District A			District B		
Party	Instrument	Close election	Observed	Instrument	Close election	Observed
PRI	1/8	3/8	3/4	–1/8	1/8	1/4
PAN	–1/8	3/8	1/4	0	0	1/4
PRD	–1/8	1/8	0	1/8	1/8	2/4

The second stage regression is

$$\text{winner}_{d,t} = \mu + \rho \sum c_{m,t-1} \cdot w_{m,d} + \beta \, \widehat{\#\text{Mayors}}_{d,t} + g\Big(\sum \mathbf{V}_{m,d,t-1}\Big) + \zeta_t + \nu_{d,t},$$

where $\widehat{\#\text{Mayors}}_{d,t}$ is the instrumented treatment and $winner_{d,t}$ is an indicator of whether the copartisan congressional candidate carried the district.

To consider the number of mayors that a party barely won/lost in a district a valid instrument of the number of mayors that a party holds, we need to assess the verisimilitude of a number of assumptions (Sovey and Green 2011). The instrument is *relevant* in that its correlation with *weighted number of mayors* is high (around 0.55–0.60). The instrument is *strong*, as suggested in Table A4 by F-statistics for excluded instruments that amply surpass the F=10 cut-off point commonly used as a rule-of-thumb for weak instruments (Staiger and Stock 1997). The instrument complies with the *monotonicity* assumption, which would be violated if, for example, lower values of mayoralties barely won/lost corresponded to much higher numbers of actual municipalities held by a party because the party manages to subvert electoral results in court or by force (contested mayoral elections exist but are a tiny fraction of all mayoral elections).

The most important assumption, the *exclusion restriction*, is violated by the instrument if the number of close races is not controlled for. Consider an alternative mechanism through which the number of mayors barely won/lost may correlate with the probability of winning the district. First, the number of mayors barely won/lost increases with the number of close races in the district. Second, the number of close races in the district may affect future electoral outcomes directly, for example if parties campaign harder and spend more money in the district in anticipation that elections will be close. This potential violation of the exclusion restriction is in fact noted by Folke (2014). Yet as he argues, this alternative pathway is blocked — and validity of the exclusion restriction restored — through inclusion of the proportion of *close elections* in the district as an additional control. Thus, while including the instrument as a regressor *on its own* would violate the exclusion restriction, the instrument satisfies the exclusion restriction once we control for the proportion of close elections in a district. Note as well that we include controls for district vote share, and in some specifications also for fixed state- and election-effects. These are necessary if the number of close elections is somewhat driven by vote share (parties with larger vote shares are more competitive, and therefore likely to be close to the discontinuity more often), by state characteristics (the voters in a particular state may be evenly divided among parties, therefore generating a larger number of close elections), or by electoral race (federal elections in a particular year may have led to a larger number of even contests).

In keeping with an the analysis of PRI-controlled municipalities to avoid lack of independence across observations with copartisan and opposition governors, Table A4 displays two-stage least squares estimates of the effect of the proportion of copartisan PRI mayors in a district on the probability that the PRI congressional candidate will win the district based on two sets of observations: for candidates with PRI governors and without PRI governors). In both cases, the instrument is the (population-weighted) proportion of municipalities in a district that a party won or lost by a margin of five percentage points or less. In the first-stage specification we regress the (population-weighted) *observed* number of copartisan mayors on the instrument, the (population-weighted) proportion of close mayoral races in the district, a fourth-order polynomial of the party's *normal vote* share in the district, the *margin of victory* in the previous election, the municipality's *poverty* level, and a set of election-year dummies. Models 2 and 4 add state fixed effects. Across all models the instrument is a statistically significant predictor of the (population-weighted) proportion of copartisan mayors in the municipality. Furthermore, F-statistics for the excluded instrument amply surpass the cut-off point suggested by Staiger and Stock (1997), alleviating concerns that the instrument may be weak.

In the second-stage model we estimate the effect of the *instrumented* proportion of copartisan mayors on the probability that the party will win the district, based on a linear probability model.[35] In line with expectations, comparing the intercept estimates in Models 1–2 against those in Models 3–4 confirms that, when there are no copartisan mayors in the district, the probability that a congressional candidate will carry a district in a state with a copartisan governor is larger than the probability that she will win the election if the state is led by a non-copartisan governor. We also find that the (instrumented) proportion of copartisan mayors in the district has a stronger effect on the probability of winning the district when there is an opposition governor. Table A5 shows that the results are essentially identical when we do *not* weight the proportion of copartisan mayors by the municipality's population, and Table A6 shows that the results remain in place — though the estimates become substantially noisier — when using a 2.5 pp. instead of a 5 pp. margin to define close elections.

[35]We did not use a probit/logit specification for the second stage for lack of a "canned" routine to estimate standard errors that account for estimation uncertainty in the first stage.

Table A4: 2SLS estimates: effect of *proportion of copartisan mayors* on a congressional candidate's *probability of victory* in Mexico, 2000-2012 (PRI only)

	Copartisan governor		*Opposition governor*	
	(1)	(2)	(3)	(4)
(a) Second stage regression (outcome: *copartisan victory*)				
Proportion of copartisan mayors	-0.11	-0.03	0.63	0.52
(instrumented)	(0.14)	(0.15)	(0.13)*	(0.12)*
Margin of victory (lagged)	0.32	-0.11	0.08	0.15
	(0.19)	(0.24)	(0.46)	(0.51)
Poverty	0.11	0.19	0.13	0.05
	(0.03)*	(0.04)*	(0.03)*	(0.04)
Proportion of close elections	-0.05	-0.01	0.22	0.27
	(0.18)	(0.17)	(0.17)	(0.19)
Intercept	0.76	0.47	0.08	0.08
	(0.12)*	(0.15)*	(0.06)	(0.08)
RMSE	0.37	0.36	0.34	0.33
(b) First stage regression (outcome: *proportion of copartisan mayors*)				
Proportion barely won/lost	0.96	0.96	1.05	0.97
	(0.04)*	(0.05)*	(0.05)*	(0.13)*
Margin of victory (lagged)	-0.29	-0.46	-0.37	-0.43
	(0.11)*	(0.10)*	(0.30)	(0.16)*
Poverty	-0.04	-0.04	0.01	-0.05
	(0.01)*	(0.03)	(0.02)	(0.08)
Proportion of close elections	-0.47	-0.39	0.36	0.31
	(0.12)*	(0.14)*	(0.15)*	(0.11)*
Intercept	0.77	1.10	0.30	0.32
	(0.03)*	(0.05)*	(0.05)*	(0.05)*
RMSE	0.20	0.18	0.26	0.21
F-statistic (excl. instrument)	548	436	510	59
Previous vote share	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
State effects	No	Yes	No	Yes
num. obs.	336	336	146	146

* $p<0.05$; † $p<0.10$; Instrumental variable estimations of models in Table 2b.

2SLS estimates, standard errors clustered by state in parentheses.

IV is the population-weighted *proportion barely won/lost*, based on a 5 percentage point margin.

Table A5: 2SLS estimates: effect of *proportion of copartisan mayors* on a congressional candidate's *probability of victory* in Mexico, 2000-2012 (unweighted averages, PRI only)

	Copartisan governor		*Opposition governor*	
	(1)	(2)	(3)	(4)
(a) Second stage regression (outcome: *copartisan victory*)				
Proportion of copartisan mayors	-0.03	0.04	0.70	0.56
(instrumented)	(0.14)	(0.15)	(0.15)*	(0.18)*
Margin of victory (lagged)	0.33	-0.09	0.09	0.16
	(0.19)†	(0.22)	(0.47)	(0.49)
Poverty	0.12	0.19	0.12	0.05
	(0.03)*	(0.04)*	(0.03)*	(0.05)
Proportion of close elections	-0.04	-0.01	0.15	0.24
	(0.19)	(0.18)	(0.13)	(0.17)
Intercept	0.70	0.39	0.07	0.06
	(0.11)*	(0.15)*	(0.07)	(0.10)
RMSE	0.37	0.36	0.34	0.33
(b) First stage regression (outcome: *proportion of copartisan mayors*)				
Proportion barely won/lost	0.99	0.99	1.08	0.97
	(0.05)*	(0.05)*	(0.06)*	(0.14)*
Margin of victory (lagged)	-0.40	-0.55	-0.41	-0.45
	(0.13)*	(0.11)*	(0.31)	(0.19)*
Poverty	-0.04	-0.05	0.02	-0.04
	(0.01)*	(0.03)	(0.02)	(0.08)
Proportion of close elections	-0.49	-0.41	0.29	0.22
	(0.12)*	(0.14)*	(0.15)†	(0.10)*
Intercept	0.78	1.11	0.30	0.32
	(0.03)*	(0.05)*	(0.05)*	(0.06)*
RMSE	0.21	0.19	0.27	0.22
F-statistic (excl. instrument)	346	369	366	50
Previous vote share	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
State effects	No	Yes	No	Yes
num. obs.	336	336	146	146

* p<0.05; † p<0.10; Instrumental variable estimations of models in Table 2b.

2SLS estimates, standard errors clustered by state in parentheses.

IV is the (unweighted) *proportion barely won/lost*, based on a 5 percentage point margin.

Table A6: 2SLS estimates: effect of *proportion of copartisan mayors* on a congressional candidate's *probability of victory* in Mexico, 2000-2012 (2.5pp. margin, PRI only)

	Copartisan governor		*Opposition governor*	
	(1)	(2)	(3)	(4)
(a) Second stage regression (outcome: *copartisan victory*)				
Proportion of copartisan mayors	-0.34	-0.26	0.72	0.52
(instrumented)	(0.12)*	(0.13)*	(0.24)*	(0.26)*
Margin of victory (lagged)	0.27	-0.20	0.13	0.15
	(0.22)	(0.27)	(0.48)	(0.52)
Poverty	0.10	0.17	0.13	0.06
	(0.03)*	(0.04)*	(0.03)*	(0.05)
Proportion of close elections	-0.09	0.06	0.27	0.29
	(0.20)	(0.19)	(0.14)*	(0.15)†
Intercept	0.92	0.72	0.06	0.11
	(0.10)*	(0.17)*	(0.09)	(0.13)
RMSE	0.39	0.37	0.35	0.33
(b) First stage regression (outcome: *proportion of copartisan mayors*)				
Proportion barely won/lost	1.05	1.03	1.15	1.06
	(0.05)*	(0.06)*	(0.07)*	(0.17)*
Margin of victory (lagged)	-0.32	-0.43	-0.38	-0.68
	(0.18)†	(0.22)†	(0.32)	(0.22)*
Poverty	-0.05	-0.05	0.01	-0.05
	(0.01)*	(0.03)†	(0.03)	(0.08)
Proportion of close elections	-0.42	-0.38	0.24	0.11
	(0.11)*	(0.13)*	(0.17)	(0.10)
Intercept	0.76	1.16	0.33	0.43
	(0.02)*	(0.05)*	(0.05)*	(0.05)*
RMSE	0.22	0.21	0.29	0.24
F-statistic (excl. instrument)	380	276	268	37
Previous vote share	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
State effects	No	Yes	No	Yes
num. obs.	336	336	146	146

* $p<0.05$; † $p<0.10$; Instrumental variable estimations of models in Table 2b.

2SLS estimates, standard errors clustered by state in parentheses.

IV is the (populatin-weighted) *proportion barely won/lost*, based on a 2.5 percentage point margin.

C Robustness

Table A7: Probit estimates: *proportion of copartisan mayors* on a congressional candidate's *probability of victory* in Mexico, 2000-2012

	Copartisan governor		*Opposition governor*	
	(1)	(2)	(3)	(4)
(a) Full Sample				
Proportion of copartisan mayors	0.81	0.50	0.89	1.03
	(0.31)*	(0.32)	(0.41)*	(0.44)*
Margin of victory (lagged)	0.36	-0.85	1.92	2.16
	(0.70)	(1.08)	(0.94)*	(0.82)*
Poverty	0.33	0.68	-0.08	-0.38
	(0.13)*	(0.22)*	(0.12)	(0.10)*
Intercept	0.12	0.57	-0.78	-0.89
	(0.19)	(0.25)*	(0.18)*	(0.19)*
Null deviance	531.8	531.8	552.1	552.1
Residual deviance	424.3	357.3	424.5	378.7
AIC	448.3	417.3	448.5	438.7
num. obs.	453	453	461	461
(b) PRI-only Sample				
Proportion of copartisan mayors	0.04	0.01	2.82	2.99
	(0.31)	(0.32)	(0.41)*	(0.44)*
Margin of victory (lagged)	1.59	-0.56	2.57	6.08
	(0.70)*	(1.08)	(0.94)*	(0.82)*
Poverty	0.52	0.96	0.69	0.37
	(0.13)*	(0.22)*	(0.12)*	(0.10)*
Intercept	0.57	0.22	-2.62	-3.26
	(0.19)*	(0.25)	(0.18)*	(0.19)*
Null deviance	377.9	377.9	200.2	200.2
Residual deviance	268.9	228.7	86.0	67.5
AIC	292.9	284.7	110.0	111.5
num. obs.	336	336	146	146
Previous vote share	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
State effects	No	Yes	No	Yes

* $p<0.05$; † $p<0.10$; Probit replications of models in Table 2.

ML estimates, standard errors clustered by state in parentheses.

Table A8: OLS estimates: The outcome is the congressional candidate's *vote share*.

	Copartisan governor		*Opposition governor*	
	(1)	(2)	(3)	(4)
(a) Full Sample				
Proportion of copartisan mayors	0.03	0.02	0.01	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Margin of victory (lagged)	-0.14	-0.13	-0.07	-0.00
	(0.05)*	(0.05)*	(0.06)	(0.06)
Poverty	0.02	0.05	0.01	-0.01
	(0.01)*	(0.01)*	(0.01)	(0.01)
Intercept	0.40	0.41	0.32	0.33
	(0.02)*	(0.02)*	(0.01)*	(0.01)*
RMSE	0.08	0.07	0.08	0.08
num. obs.	453	453	461	461
(b) PRI-only Sample				
Proportion of copartisan mayors	-0.02	-0.01	0.03	-0.00
	(0.02)	(0.02)	(0.02)	(0.01)
Margin of victory (lagged)	0.03	-0.02	-0.18	-0.08
	(0.04)	(0.03)	(0.07)*	(0.06)
Poverty	0.02	0.05	0.05	0.03
	(0.01)*	(0.01)*	(0.01)*	(0.01)*
Intercept	0.44	0.41	0.33	0.37
	(0.02)*	(0.03)*	(0.01)*	(0.01)*
RMSE	0.07	0.06	0.07	0.06
num. obs.	336	336	146	146
Previous vote share	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
State effects	No	Yes	No	Yes

* $p<0.05$. Specifications replicate the models in Table 2, but employing *vote share* as the outcome. OLS estimates (standard errors clustered by state in parentheses).

Table A9: OLS estimates: The outcome is the congressional candidate's *margin of victory*.

	Copartisan governor		*Opposition governor*	
	(1)	(2)	(3)	(4)
(a) Full Sample				
Proportion of copartisan mayors	0.09	0.05	0.08	0.06
	(0.03)*	(0.02)*	(0.04)*	(0.03)†
Margin of victory (lagged)	0.05	-0.04	0.11	0.04
	(0.06)	(0.09)	(0.09)	(0.09)
Poverty	0.03	0.07	-0.01	-0.04
	(0.01)*	(0.02)*	(0.02)	(0.02)*
Intercept	0.01	0.05	-0.10	-0.10
	(0.02)	(0.02)*	(0.02)*	(0.02)*
RMSE	0.14	0.12	0.14	0.13
num. obs.	453	453	461	461
(b) PRI-only Sample				
Proportion of copartisan mayors	0.01	0.00	0.12	0.03
	(0.03)	(0.02)	(0.04)*	(0.04)
Margin of victory (lagged)	0.29	0.13	-0.08	-0.09
	(0.05)*	(0.07)†	(0.12)	(0.15)
Poverty	0.03	0.07	0.06	0.01
	(0.01)*	(0.02)*	(0.01)*	(0.01)
Intercept	0.05	0.08	-0.09	-0.06
	(0.02)*	(0.05)†	(0.02)*	(0.02)*
RMSE	0.12	0.11	0.12	0.11
num. obs.	336	336	146	146
Previous vote share	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
State effects	No	Yes	No	Yes

* $p<0.05$. Specifications replicate the models in Table 2, but employing *margin of victory* as the outcome. OLS estimates (standard errors clustered by state in parentheses).

Table A10: OLS estimates: controlling for the effective number of municipalities

	Copartisan governor		*Opposition governor*	
	(1)	(2)	(3)	(4)
(a) Full Sample				
Proportion of copartisan mayors	0.25	0.17	0.28	0.29
	(0.10)*	(0.09)†	(0.13)*	(0.13)*
Effective number of municipalities	0.01	-0.00	-0.01	-0.00
	(0.00)*	(0.01)	(0.00)*	(0.00)
Margin of victory (lagged)	0.03	-0.25	0.54	0.43
	(0.19)	(0.26)	(0.23)*	(0.20)*
Poverty	0.06	0.17	0.01	-0.09
	(0.03)*	(0.04)*	(0.03)	(0.03)*
Intercept	0.49	0.65	0.29	0.24
	(0.07)*	(0.07)*	(0.06)*	(0.05)*
RMSE	0.40	0.38	0.39	0.38
num. obs.	453	453	461	461
(b) PRI-only Sample				
Proportion of copartisan mayors	0.03	0.05	0.52	0.42
	(0.08)	(0.08)	(0.10)*	(0.13)*
Effective number of municipalities	0.01	0.01	0.01	0.00
	(0.01)	(0.00)	(0.01)	(0.01)
Margin of victory (lagged)	0.29	-0.13	-0.04	0.10
	(0.18)	(0.23)	(0.41)	(0.53)
Poverty	0.10	0.17	0.11	0.04
	(0.03)*	(0.04)*	(0.05)*	(0.07)
Intercept	0.62	0.34	0.12	0.14
	(0.05)*	(0.11)*	(0.07)	(0.08)†
RMSE	0.37	0.36	0.34	0.33
num. obs.	336	336	146	146
Previous vote share	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
State effects	No	Yes	No	Yes

* $p<0.05$. Specifications replicate the models in Table 2.

OLS estimates (standard errors clustered by state in parentheses).

Table A11: RD estimates of *municipal incumbency* on the municipal-level performance of Mexican congressional candidates, 2000-2012

	LATE	SE	bwd.	N^-	N^+
Local linear regression (outcome: *copartisan victory*)					
Copartisan governor	-0.09	0.03*	0.21	1355	1649
Non-copartisan governor	-0.11	0.05*	0.11	996	742
Local linear regression (outcome: *turnout*)					
Copartisan governor	-0.01	0.01	0.17	1225	1472
Non-copartisan governor	0.01	0.01	0.15	1322	900

* p<0.05. The outcome variables correspond to *federal* elections held at $t+1$, but are measured at the level of the municipality. The running variable is the *margin of victory* in the municipal election held at t. Bias-corrected estimates are based on a local linear regression fitted separately at both sides of the threshold and employing a triangular kernel. The bandwidth is calculated according to the automatic selection procedure proposed by Calonico, Cattaneo and Titiunik (2014).