Online Appendix to "Rank Effects in Bargaining: Evidence from Government Formation"

Thomas Fujiwara
Princeton, CIFAR, and NBER

Carlos Sanz
Bank of Spain

Appendix A: Additional Results on Spanish Mayoral Appointments

Are mayoral appointments just symbolic? One possible interpretation of the results is that mayoral appointments are mostly symbolic. For example, parties could bargain over rent allocations and policy decisions based on their seat distributions and ability to form majority coalitions, and once those are decided, simply appoint the most voted party the mayor symbolically. There are five reasons this is interpretation is unlikely. First, as previously discussed, mayors are dominant figures that exercise substantial personal discretion over policy once appointed. Second, it is not clear why the results would arise if mayoral appointments were just symbolic (e.g., why parties in a coalition would not appoint mayors for half a term each). Third, if the appointment is symbolic, it is unclear why voters would punish parties that deviate from it, as discussed in Section 6. Fourth, the stakes of appointing a mayor presumably become larger as the size of the municipality grows, both because it controls a larger budget and because larger municipalities must, by law, provide additional public services. Figure A12, however, indicates that RDD "jumps" are larger in municipalities with larger councils, which are also those with more inhabitants (Table A2).[1] Fifth, it does not appear that mayors share power with other parties.

To further shed light on whether the mayoral appointments are just "symbolic," we can study the allocation of deputy mayors. While we do not observe cabinet allocations across municipalities, we do observe the appointment of deputy mayors (*tenientes de alcalde*), which are the second most visible position in municipal government. There may be more than one deputy mayor in a municipality, in which case they are ranked. In cases of absence or illness of the mayor, the (first) deputy mayor takes over her duties.[2]

Formally, the mayor has the discretion to choose which of the elected councilors are appointed deputy mayors. However, it is possible that deputy mayor positions and the mayoral appointment are bargained over by parties. Hence, one could expect that the "effect of most voted" on mayoral appointments is offset by a *negative* effect on appointment of deputy mayors (i.e., most voted are more likely to appoint mayors and the second most voted then appoint most deputy mayors).

[1] There are, of course, other characteristics that are correlated with municipality size that could affect effect sizes. For example, larger municipalities are more likely to have a local media outlet that makes council decisions salient to voters.

[2] If the first mayor is also not available, the second deputy mayor takes charge, and so forth. The mayor has discretion on the number of appointed deputy mayors. In municipalities with more than 5,000 inhabitants, the number of deputy mayors must be between one and one third of the council size. In those with less than 5,000, appointing a deputy mayor is optional. The average municipality in our sample of ties in seats has 2.6 deputy mayors (s.d.=1.6).

Figure A21 repeats the exercise of Figure 2a, using instead the share of deputy mayors appointed by the party and a dummy for the party appointing *all* deputy mayors. It also provides placebo tests based on lagged outcomes. Table A8 provides the corresponding estimates.[3] The estimated effects are all *positive* and sizable. Hence, being the most voted party increases the chance of appointing the mayor and the share (or having all) of the deputy mayors. This is the opposite of what the hypothesis that the two most voted parties obtain equal amount of power would predict.

Which types of coalitions drive the effect? While we cannot directly observe which parties are supporting the mayor, we can indirectly study whether coalitions by the most voted and second most voted, or between the most voted and third-placed party, drives the effects. Although we cannot observe the supporting coalitions, we leverage previously discussed evidence that left-wing PSOE and the right-wing PP are unlikely to support each other's governments. Figure A22 replicates Figure 2a, but separately plotting the cases where i) both the PP and PSOE are the two most voted parties, ii) only one of them was amongst the two most voted parties, and iii) neither are amongst the two most voted parties. The effects are similar (and not statistically distinct) in all three cases. This suggests the effect is not driven by coalitions where the second most voted (or third placed) party supports the first most voted.

Is the most voted effect different for powerful parties? Figure A22 suggests similar effects on a party, regardless of whether it is one of the two dominant parties is Spanish politics or not. To further probe possible differential effects given how powerful a party is, Figure A23 repeats the exercise of Figure 2a plotting separately the cases where the party is in power (i.e., holding the main executive position) at the national (or, alternatively, regional) government at the time of the municipal election. The effects are similar regardless of whether the party is in power or not. We return to this issue on Section 5, in light of our theoretical framework.

Appendix B: Further Information on the Theoretical Framework

Proof of Proposition 1. Voters' strategies are a best response since (as discussed previously) a voter that observes a private signal of value i (but before observing election results) expects i to be the most likely state and prefers party of type i to appoint the mayor. This remains true even after the voter conditions its decision on the possibility it may be pivotal (e.g., one party will be one vote away from a majority). Since all offers are approved in equilibrium, all parties find it optimal to offer to appoint the mayor themselves in every period. Since all parties accept any proposal, any deviation by an individual party cannot change the bargaining outcome (since a majority is accepting). Note this self-fulfilling feature is not needed for this result. A similar equilibrium can be sustained with party A rejecting any proposal that does not make it the mayor. Parties B and C will still be indifferent between accepting or not any offer that does not appoint them the mayor (since A is the status quo).

Proof of Proposition 2. Voters' strategies are a best response since (as discussed previously) a voter that observes a private signal of value i (but before observing election results) expects i to be the

[3]The available data on deputy mayors list their party affiliation, but no their rank. Hence we cannot focus on the effect of appointing, for example, the first deputy mayor that takes office when the mayor is absent. Data on deputy mayors is not available for the entire sample, and hence sample sizes in Table A8 are smaller than in Table 3.

most likely state and prefers party of type i to appoint the mayor. This remains true even after the voter conditions its decision on the possibility it may be pivotal (e.g., one party will be one vote away from a majority or being most voted). For the second and third most voted, a deviation to proposing appointing the mayor itself will be accepted and lead to a payoff of one in that period. However, it will trigger the voters' punishment strategy and make it never be elected again. The value of not deviating and continuing to be re-elected is thus the perpetuity of the probability of being the most voted party (e.g., $\frac{\Pr[p_t^A > \max(p_t^B, p_t^C)]}{(1-\delta)}$ for party A), which is larger than one for all parties given conditions (i)-(iii). Hence, proposing the most voted party appoints the mayor is a best response. Strategies are also best responses regarding accepting offers: the second most voted party is indifferent between accepting or not a proposal that makes the first or third most voted party the mayor. A similar logic applies to the third most voted.

Example of Voter Belief Updating in the Theoretical Framework. To illustrate the workings of the model, this section provides an example using a specific distribution of possible states of the world. Recall that $\boldsymbol{p_t} = [p_t^A, p_t^B, p_t^C]$ denotes the probabilities of state $s_t \in \{A, B, C\}$ occurring. In particular, assume that $G(\boldsymbol{p}_t)$ is such that its possible realizations are i) $\boldsymbol{p}_t = [0.45, 0.35, 0.20]$, ii) $\boldsymbol{p}_t = [0.35, 0.45, 0.20]$, and iii) $\boldsymbol{p}_t = [0.20, 0.20, 0.60]$. Each realization can occur with probability equal to 1/3.

Hence, when a period starts, voters have priors that each of the states of the world are equally likely. After observing a signal σ equal to A, a voter updates and then believes that the probability that realization (i) occurred is $\frac{0.45 \cdot (1/3) \cdot (0.45+0.35+0.2)}{1/3} = 0.45$. She similarly believes that the probability that (ii) occurred is 0.35 and that (iii) occurred is 0.2.

Hence, observing signal A makes her update that the probability of each state occurring: $\Pr(s_t = A|\sigma_t = A) = 0.45^2 + 0.35^2 + 0.2^2 = 0.365$; $\Pr(s_t = B|\sigma_t = A) = 0.45 \cdot 0.35 + 0.35 \cdot 0.45 + 0.2^2 = 0.355$; and $\Pr(s_t = C|\sigma_t = A) = 0.45 \cdot 0.2 + 0.35 \cdot 0.2 + 0.2 \cdot 0.6 = 0.280$. Similarly, observing a signal B will make her believe that state B has a 0.365 probability of occurring (while probability of A and C are 0.355 and 0.280, respectively). A similar calculation yields the updated beliefs after a voter observes signal signal C: $\Pr(s_t = A|\sigma_t = C) = \Pr(s_t = B|\sigma_t = C) = 0.280$ and $\Pr(s_t = C|\sigma_t = C) = 0.440$.

Note that this distribution satisfies the $\Pr(s_t = i|\sigma_t = i) > \Pr(s_t = j|\sigma_t = i)$ for all $i \neq j$ condition. So a voter that observes signal i prefers party of type i to be the mayor. However, if all voters vote according to their signals, the actual vote shares will match one of the $\boldsymbol{p}$ realizations - e.g., if realization (i) occurs, the vote shares of parties A, B, and C will be 0.45, 0.35, and 0.20, respectively. After observing such vote shares, citizens would then update accordingly: e.g., expect that the probability that the state is A is 0.45. This implies that then *all* voters will prefer party A to appoint the mayor, but at this point representation in the legislature is already determined. Given that party A does not have a majority of the votes, it might be possible for B or C to appoint the mayor. This illustrates the main conflict between voters and parties in the model. Vote shares aggregate diffuse information from the voters, which informs which party they prefer would appoint the mayor. However, after the election takes place, the decision on which party appoint the mayors may not necessarily heed to the preferences of voters.

Since $G(\boldsymbol{p}_t)$ is assumed to be i.i.d. and serially uncorrelated, when a new period starts all voters

beliefs about the state of the world return to the same prior (so past election results and mayoral appointments do not inform their rule). Lastly, in this particular example, a near tie between two parties for the most voted position is not possible. However, if realizations (i) and (ii) of the $G(\boldsymbol{p}_t)$ were instead $[0.40+\epsilon, 0.40-\epsilon, 0.20]$ and $[0.40-\epsilon, 0.40+\epsilon, 0.20]$, with $\epsilon \to 0$, we have a case where parties A and B almost tie and the updating works similarly.

Appendix C: Alternative Bargaining Procedures

The theoretical framework assumes a specific bargaining procedure for mayoral appointments. Beyond tractability and simplicity, our particular choice of assumptions is also made to better match the one round of voting present in the Spanish context. This appendix discusses how the results in Section 4 are robust to a different bargaining procedure, which allows for both infinite rounds of bargaining and for rents from office to be divisible across parties.

This alternative procedure is inspired on Baron and Ferejohn (1989). As before, if one party obtains a majority, it can choose the allocation of rents. If no party has a majority, then one is randomly recognized to propose a division of the rents. Recognition probabilities are the same for all three parties. The non-recognized parties can accept or not this proposal. If at least one (non-recognized) party accepts, the recognized party appoints the mayor and the proposed division is realized.

If no party accepts, another identical round of bargaining begins, with another independent draw of the proposing party. Note that we do not need to specify a status quo appointment in this game, and technically the bargaining can last forever if offers are never accepted. Additionally, we abstract from discounting across bargaining rounds (so not to confuse with discounting across periods), however it is straightforward to incorporate them.

If no party has a majority, this bargaining game has an equilibrium with stationary (history independent) and symmetric strategies with the proposer offering $x_t = 1/3$ to one randomly chosen party and $x_t = 2/3$ for itself, with the first proposal being accepted.[4]

Proposition 1 can be adapted to, when no party obtains a majority, having all parties propose keeping $2/3$ of the rents and offering $1/3$ to another (randomly chosen) party. All parties accept such proposal. Note this implies that all parties have equal probability of appointing the mayor. If one party has a majority, then it appoints the mayor with certainty.

Proposition 2 can be similarly adapted. Note that we now equate "appointing the mayor" with "having a proposal accepted." The new equilibrium strategy for a most voted party is: i) if recognized, offer to keep all the rents to itself; ii) if not recognized, to reject all offers. The equilibrium strategy for second and third most voted parties is: i) if recognized, offer to keep all rents to itself; ii) if not recognized, accept any offer that assigns it non-zero rents. If the offer assigns it zero rents, accept if it is from the most voted party and reject if it is from the second and third most voted.

These are clearly best responses to the most voted party. The strategies for second and third most voted parties are best responses given that they are indifferent between accepting or rejecting an offer

[4] A proposer keeps y and offers one randomly drawn other party $1-y$. For the other party to accept, its payof must be $1-y > V$, where V is the continuation value of this legislative bargaining game. The proposer thus optimally makes this inequality bind, so the proposal is accepted. Hence, the continuation value equals $V = \frac{1}{3}y + \frac{2}{3}(1-y) = \frac{1}{3}$ and $y = \frac{2}{3}$.

that assigns zero rents. A deviation where they make an offer that is accepted (off the equilibrium path) cannot be a best response. Such deviation yields at most a payoff of one, since the party is never re-elected again, which is less than the continuation value of being re-elected given conditions (i)-(iii).

This adapted version of Proposition 2 also leads to the most voted party appointing the mayor every period. Interestingly, the party appointing the mayor in the equilibrium described in Proposition 2 obtains more rents ($x = 1$) than the one described in Proposition 1 ($x = 2/3$).

Appendix D: Do Municipalities that Appoint Most Voted Mayors Have Less Corruption?

This appendix addresses the model's prediction that municipalities following the norm have mayors of higher (state-specific) quality and thus higher voter welfare. In particular, we report an association between a municipality's tendency to appoint most voted mayors and lower government corruption.

Interpretation. A formal interpretation of these results in light of the model can be obtained by assuming that different states of the world represent that one party (A, B, or C) is fielding the only honest candidate (or one party has no opportunity to engage in corruption) in that period. If voters prefer honesty (or less corruption) and have imperfect information about which party is more honest (observing only individual signals), the setup of the model extends to the most voted candidate being more honest and municipalities under the equilibrium from Proposition 1 (no norm) having more corruption than those under the equilibrium from Proposition 2 (norm).

The assumption that voters prefer less corruption seems plausible in the context of Spanish municipalities. Instances of municipality-level corruption lower trust in government (Solé-Ollé and Sorribas-Navarro, forthcoming) and incumbent's votes (Costas-Pérez, Solé-Ollé, and Sorribas-Navarro (2012)).

An important caveat is that voters preferences over mayoral appointments (and having a most-voted mayor) may be driven mostly by considerations separate from corruption (e.g., how mayors allocate spending). However, it is not possible to observe voter welfare or to infer the quality of policymaking from observable policies (e.g., budgets or spending) or outcomes that governments have little control over (e.g., income or population growth). Thus, our focus on corruption is also partly driven by these data availability concerns.

Another caveat is that our model provides little guidance over which municipalities will adopt the equilibrium following the norm. Our model thus does not suggest an identification strategy for the effects of the norm on outcomes such as corruption. Hence, the results in this appendix are based on correlations and, although they appear robust to controlling for many relevant factors, it is difficult to rule out the possibility of other confounding factors or reverse causality.

Context and data. Government corruption in Spain is usually linked to *municipal* regulation of land use (in a typical case, local officials take bribes in exchange for amendments to land use plans and building permits).

We use a measure of corruption based on newspaper reports from Solé-Ollé and Sorribas-Navarro (forthcoming) that covers all Spanish municipalities in the 1991-2015 period. It contains a dummy ($corruption_{it}$) indicating whether a corruption case was uncovered in municipality i during the electoral

term starting at year t. A corruption case occurs in 5.7% of the observations in our sample, and 22.5% of municipalities experienced at least one corruption case during 1991-2015, indicating that corruption is widespread.[5]

Whether a municipality is acting consistently with the norm from Proposition 2 is not directly observable, but we can construct a proxy variable ($norm_i$) in the following manner. For each municipality in the sample, we calculate the number of times in which the two most voted parties tied in seats and their vote shares were less than 1 p.p. apart. If this number is zero, we code $norm_i$ as missing, otherwise, we code $norm_i$ as the share of times after such close elections that the most voted party appointed the mayor. The intuition, based on our previous analysis, is that a strong signal of whether or not a municipality follows the norm is only available when observing parties almost tying in the number of votes.

Results. To estimate the effect of the norm on corruption, we regress $corruption_{it}$ on $norm_i$, and study the robustness of the results to a variety of fixed effects and controls. While the outcome varies at the municipality-electoral term level, $norm_i$ does not vary across time within a municipality.[6]

Column (1) in Table A9 presents the estimate from this regression. To interpret the magnitudes, consider a municipality following the equilibrium without the norm. It would have a value of $norm_i$ of approximately 1/2 and a 5.7% probability of observing a corruption during a 4-year electoral term. A municipality following the norm in all periods (as in the equilibrium in Proposition 2) would have $norm_i = 1$ and a 4.4% probability of observing a corruption scandal. This difference is not only statistically significant but also economically substantial. Columns (2)-(5) sequentially add a number of controls (described in the Table A9's notes). We highlight that including province-year effects does not affect our results. Intuitively, the results hold when comparing two municipalities from the same province in the same year, which is remarkable since most confounding effects are likely to systematically vary at this level.

Lastly, we reiterate that the results reported in this appendix should be interpreted as a correlation between municipalities acting consistently with a "most voted party appoints the mayor" prescription and lower instances of corruption. Fully ruling out the possibility of omitted variables or reverse causality driving the result would require an identification strategy leveraging an understanding of why some municipalities act as if following the norm that arises in our model, which is beyond the scope of this study.

[5] Our corruption data was originally created by Fundación Alternativas, which hired a journalist in each Spanish province to compile news items referring to its municipal corruption scandals by looking at the municipal, regional, and national press. Solé-Ollé and Sorribas-Navarro (forthcoming) expanded the dataset for the 1999-2007 period by conducting internet searches of newspapers. The same authors have expanded the data to the 2007-2015 period by searching all national and many regional newspapers in the Factiva archive, relying on a machine learning algorithm to identify the actual cases of corruption occurring by municipality. The data is harmonized to be comparable throughout the 1991-2015 period. Further informaton on the data can be found in Fundación Alternativas (2007) and Solé-Ollé and Sorribas-Navarro (forthcoming), which also present a general overview of municipal corruption cases in Spain.

[6] This implies that we cannot include municipality fixed effects. The sample is an unbalanced panel of 2,390 observations from 450 unique municipalities. $norm_i$ is equal to zero 44% of the times and equal to one 52% of the cases (with the few remaining cases being equal to 1/2 or 2/3). Note that the fact that we only observe one (or a few) close elections for each municipality only generates classical measurement error in $norm_i$ and thus can only generate attenuation bias.

Appendix E: The Role of the Status Quo Rule

To the best of our knowledge, the only differential institutional treatment of parties by rank of their votes in Spanish municipal elections is the status quo described in Section 4. If no candidate receives a majority of votes in the council election, the party with the most votes appoint the mayor. While, at first pass, this appears to be likely to explain our results, there are four reasons we believe the status-quo rule cannot be the main driver of our results.

First, and perhaps most importantly, there is no similar status quo rule, or any other institutional advantage, that is given to the second most voted party. Hence, the status quo rule cannot play a role in explaining the second most voted versus third most voted effects described in Section 4.3 and thus cannot account for the entirety of our evidence. Second, it is not clear why the existence of this status quo would make voters "punish" second most voted parties that deviate from the norm or why it would interact with the vote share of the third most voted party (Subsection 6.2) or the size of municipality or council (Figure A13).

Third, note that parties' coordination failures or mistakes when casting votes for mayor are unlikely to generate our effects via the status-quo rule. As noted in Subsection 4.1, a majority of council members can easily replace the mayor at their will at any point of the term. Hence, even if by mistake in casting votes a majority was not obtained and the first-placed party appointed a mayor that displeased a majority, that could be undone quickly.

Fourth, also as discussed in Subsection 4.1, the situation of the vast majority of councils in our sample fits a three-player majority game: any two of the three most voted parties can form a majority coalition. In such cases it is particularly unclear why the status quo should matter. To formalize this argument, we outline below a voting game that approximates these conditions and the Spanish institutions for selecting a mayor. It shows that, given sensible equilibrium refinements (i.e., parties not taking weakly dominated actions or allowing two parties to coordinate in their deviation), the status-quo rule is irrelevant in defining which party elects the mayor.

Lastly, it should be noted that there is little available data on whether the status quo rule was applied or not (e.g., whether a mayor was appointed through a majority of votes or not). However, we believe there would be little meaningful information in this data. For example, suppose that, in a three party legislature, each party votes for their own leader, and the mayor is appointed by the status quo rule. For both the second and third most voted parties, unilaterally deviating from this strategy and voting for the most voted mayors does not affect the final outcome (who is the mayor) but it does change whether the status quo rule is applied or not. Given there is indifference in equilibrium, there is little information to be learned about parties incentives and beliefs from this choice. Moreover, the voting game outlined below illustrates that this indifference argument applies in all the equilibria where the most voted party appoints the mayor.

The role of the status quo in a voting game. To illustrate why the status quo rule is unlikely to play an important role in explaining the empirical results, we analyze a game matching the rules and incentives that parties face in our sample of Spanish municipalities. We focus on the case of a legislature with three parties, in which any two can form a majority. This case matches 90% of our

sample, as discussed in Subsection 4.1.

Consider a game with three parties (A, B, and C), indexed by i. As in Spanish municipal councils, each party has only one candidate for mayor, which we also label A, B, and C. Party preferences over the mayor are $u_A(A) > u_A(B) > u_A(C)$, $u_B(B) > u_B(C) > u_B(A)$, and $u_C(C) > u_C(B) > u_C(A)$. This describes a situation in which two parties (B and C) are ideologically aligned. Each party prefers to appoint the mayor itself. For parties B and C, their second option is the aligned party, and their least preferred option is C. While we assume party A prefers B over C, this is not crucial to the results. The strategy space is $\{a, b, c, \phi\}$. Parties can vote for any of the parties or abstain. Matching a situation where A and B are tied in seats and C has the same or fewer seats than A and B, if any two of the three parties vote for the same party i, then party i appoints the mayor.

For concreteness, we can think of A as the PP, B as the PSOE, and C as the IU. The leftist PSOE and IU prefer one of them to appoint the mayor over the right-wing PP. This case is depicted in Figure 4. Focusing when the PP is the reference party, assuming A is the most voted approximates the blue circles to the right of the cutoff, and the case where A is the second most voted the blue circles to the left of the cutoff. The question we address is whether the observed jump can be explained by A (the PP) changing to being the status quo as it crosses the cutoff.

Assume that A has a status quo status: it obtains the mayor if no party obtain two or more of the votes. In this case there are 14 Nash equilibria in pure strategies in the described game. Letting (s_A, s_B, s_C) denote equilibrium strategies, these are (a, a, a), (a, ϕ, a), (a, a, ϕ), (a, ϕ, ϕ), (a, b, b), (a, c, c), (b, b, b), (c, c, c), (ϕ, a, a), (ϕ, ϕ, a), (ϕ, a, ϕ), (ϕ, ϕ, ϕ), (ϕ, b, b), and (ϕ, c, c). Out of these 14 Nash equilibria, A appoints the mayor in eight, B in three, and C in three.

This multiplicity of equilibria is standard in voting games. However, the eight equilibria where A appoints the mayor are not robust to either i) allowing a coordinated deviation by two parties (e.g., strong Nash equilibrium or coalition-proofness) or ii) focusing on cases where parties do not play a weakly dominated strategy (e.g., trembling hand perfection or dominance solvability).

First, none of the eight equilibria where A appoints the mayor are robust to allowing both B and C to make a coordinated deviation. A appointing the mayor is the worst outcome for parties B and C. If the jointly deviate to either both voting for B or both for C, they can increase their utility. Note that while Nash equilibria only considers unilateral deviations, coordinated deviations seems like a natural case in a council with only three parties represented and where they can communicate before voting.

Second, note that actions b and c are weakly dominated (by A and ϕ) for party A.[7] Additionally, a and ϕ are dominated by b (for party B) and c (for party C).[8] Hence, there are only four Nash equilibria where a party is not playing a weakly dominated strategy: (a, b, b), (a, c, c), (ϕ, b, b), and (ϕ, c, c). In none of them, party A appoints the mayor, even though it is the status quo. Assuming parties choose a weakly dominated action in a small voting game is unattractive. For example, trembling hand

[7] If the other two parties are not casting the same vote, A and ϕ can guarantee the best outcome for A. If only one other party is playing a or ϕ, those actions are strictly better than b and c for A. If the other two parties are both voting b or c, A is indifferent between all actions.

[8] b guarantees the best outcome for B as long as one other party is playing b. If only one other party is playing b, b is strictly better than any other action for B. If no other party is voting for b, B is indifferent between b, a, and ϕ. A similar argument applies to party C.

perfection eliminates all equilibrium with weakly dominated strategies: as long as party i believes that, even with a very small probability, one other party will vote for i, it will not be a best response to follow a weakly dominated strategy.

To illustrate the irrelevance of the status quo status further, one could reanalyze the game but now making party B the status quo. Following a similar argument, it can be shown there is no equilibrium where A appoints the mayor and players do not choose a weakly dominated strategy. Hence, A's ability to appoint the mayor is unaffected by whether it is the status quo or not.

Appendix F: Further Alternative Explanations

Politicians' outcome bias. A possible alternative explanation is that politicians that receive the most votes become inherently more motivated and exert more effort into forming a government. We believe there are two reasons why this is unlikely to explain the entirety of our results. First, we estimate sizable effects, which would suggest motivation has perhaps unreasonably large impacts on political outcomes. Second, it is not clear how politician motivation can generate other results such as voters punishing those that deviate from the norm and the heterogeneity by strength of third place.

Agreement among parties. Another possible explanation for our results is that parties create a (perhaps implicit) agreement that the most voted party should form the government. This is a distinct mechanism from the one suggested by the theoretical framework only if the reason for such agreement is not that voters would enforce the norm. In other words, the theoretical framework shows how a norm that arises from voters' strategies determine parties' behavior.

It is not straightforward why causality would run in the other direction. Even if parties decided on their own to enforce the "most voted party appoints the mayor" norm, it is not clear why voters would punish a party that deviates from it. Similarly, it is not clear why such agreements would be more common when the third placed party obtained more votes. Additionally, it is not evident why parties would find this norm desirable. One possibility is that if bargaining after every election is costly, the norm would be in their interest. However, the costs of bargaining seem small compared to the importance of a mayoral appointment. If that is the case, there would be strong incentives for second or third most voted parties to systematically renege on this agreement, eventually making its effect disappear.

One, albeit indirect, test of this mechanism is that the effect of being most voted should be stronger in municipalities with more frequent cases of ties in number of seats.[9] Panel A of Figure A20 repeats our main RDD plot (Figure 2a) separating the sample into cases from municipalities that experienced multiple cases of the two most voted parties tying in seats, and those that only experienced one case, during our sample period.[10] The effects are similar in both subsamples. The same applies when looking at cases with even more frequent ties (two or more) in Panel B. There is no evidence that municipalities where ties in seats occur frequently are more likely to present the norm we study.

[9] Presumably, previous experience with costly bargaining would be the reason for norm to arise in future cases.

[10] Note that a municipality must experience at least one case to enter the sample. This definition separates the sample into two subsamples of approximately same size.

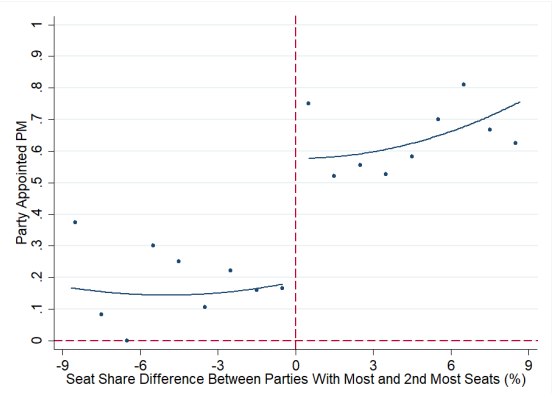
Lastly, note that discussion above regards parties making agreements at the municipal level. In principle, one could envision parties would making a nationwide (or province-wide) agreement and forcing its local politicians to follow it. Many of the issues raised above would also apply to this explanation. For example, it is not clear, why would parties find this particular agreement useful (versus any other that split mayoral allocations evenly across first and second most voted) or why voters would punish parties that deviate from the norm. Moreover, there are two pieces of evidence pointing out that the norm we study operates at a local level. First, Figure A22 indicates that our main result is of similar magnitude regardless of whether the two most voted parties are (both, one of, or neither) the two larger national parties (PP and PSOE) or not. Presumably, incentives for a large-scale agreement would not be the same for the parties with nationwide coverage and other parties. Second, the punishment of second most voted parties that appoint mayors occur at the municipal level (when a party in a given municipality deviates from the norm). It is not clear why an agreement at the national or regional level would lead voters to punish parties for local deviations.

The central government prefers winners. Another mechanism in the similar spirit would be based on assuming that the central government prefers most voted mayors (e.g., it is more likely to award them funds). As before, it is not clear why a rational and fully informed central government would have such preferences and the argument requires that the central government presents some outcome bias or an inherent preference for "winners" (and for second most voted over third most voted parties). Similarly to the discussion above, the issues related to "assuming an agent prefers the most voted to appoint mayors" also apply here.

We highlight, however, three pieces of evidence that are easier to reconcile with our model than with an explanation based on the central government preferring most voted mayors. First, it is not clear why the heterogeneity by strength of the third-place would arise under this alternative explanation. Second, as discussed previously, the concept of "it is more democratic to appoint the most voted" appears to be incorporated into the opinions of voters and the discourse of local and national politicians. If the advantage of most voted parties was entirely due to instrumental reasons related to the ability of obtaining more resources from the central government, it is not clear why that would be the case (especially when it comes for national politics, which is the highest level of government itself). Third, the case studies and cross-country evidence discussed in Section 2 relates to national governments too.

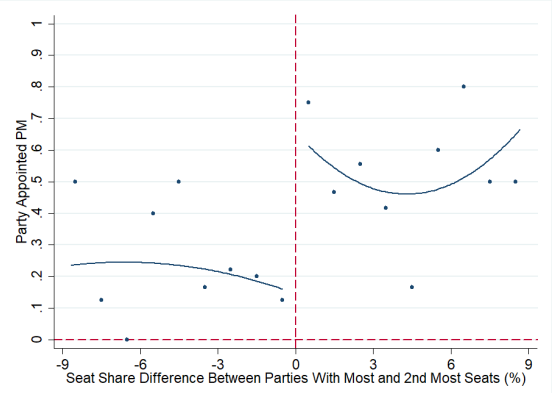
Appendix Figures and Tables

Figure A1: Effect of Having the Most Seats, Excluding Bulgaria - National Parliaments Data



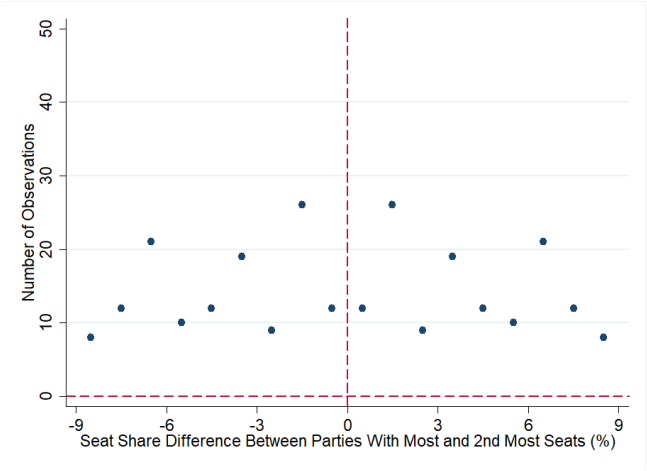
The unit of observation is a party-country-year. Sample is restricted to the two parties with the most seats in the parliament. The running variable (horizontal axis) is the difference in *seat* shares between the two parties with the most seats: positive with the most seats and negative for the party with the second most number of seats. Circles represent the local averages of a dummy indicating whether the party appoints the prime minister (Panel A) or if the party appointed the prime minister in the previous $(t-1)$ term (Panel B). Averages are calculated within 1 p.p.-wide bins of seat share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data. Bulgaria is excluded from the sample.

Figure A2: Effect of Having the Most Seats, Diermeier and Merlo (2004)'s Sample - National Parliaments Data



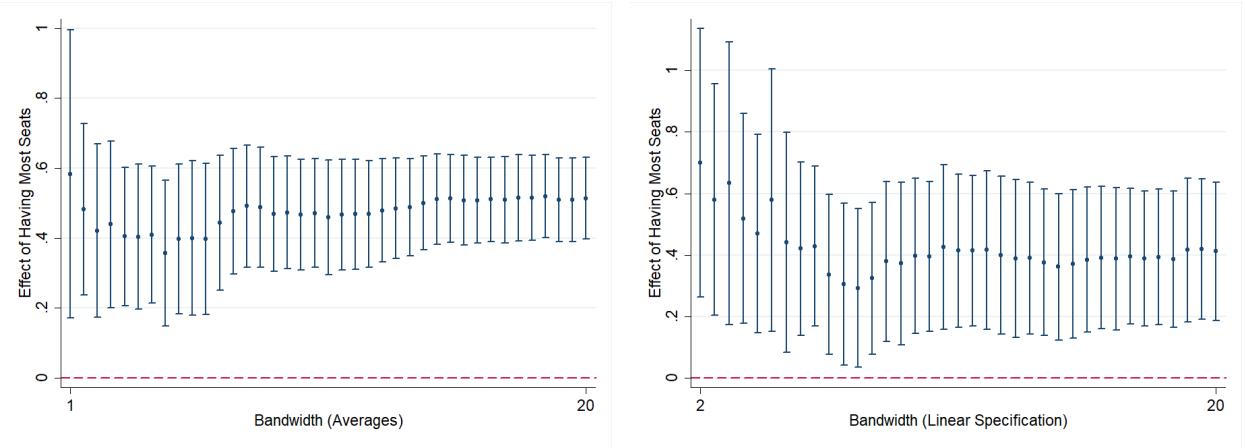
The unit of observation is a party-country-year. Sample is restricted to the two parties with the most seats in the parliament. The running variable (horizontal axis) is the difference in *seat* shares between the two parties with the most seats: positive with the most seats and negative for the party with the second most number of seats. Circles represent the local averages of a dummy indicating whether the party appoints the prime minister. Averages are calculated within 1 p.p.-wide bins of seat share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data. The sample is restricted to the 11 countries studied by Diermeier and Merlo (2004): Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Netherlands, Norway, and Sweden.

Figure A3: Histogram - National Parliaments Data



The unit of observation is a country-year-party. Each circle in represents the number of observations in the respective circle on Figure 6 of the main text. Sample is restricted to the the two most voted parties. Circles represents the number of observations in each 1 p.p.-wide bin of seat share difference.

Figure A4: Robustness to Bandwidth Choice - National Parliaments Data

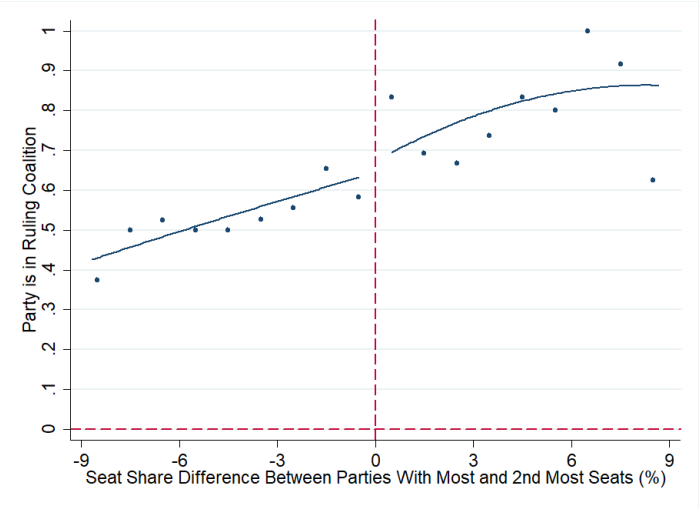


(a) Effect of 1st vs. 2nd (Specification: Means)

(b) Effect of 1st vs. 2nd (Specification: Linear)

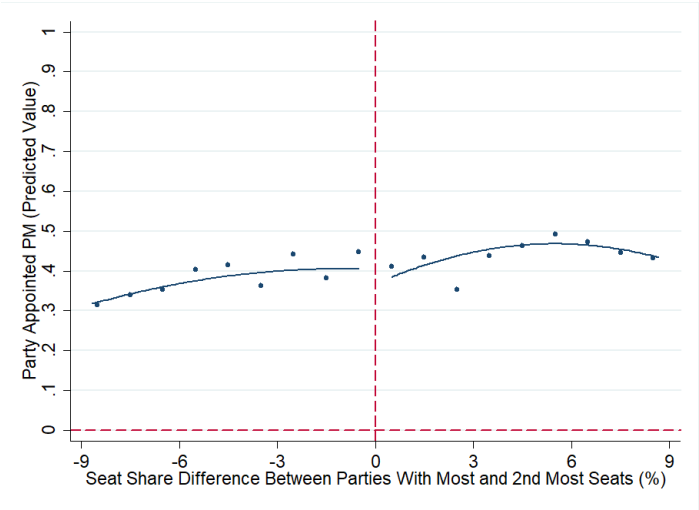
Circles represent estimated effects, using different bandwidth choices (horizontal axis). Whiskers represent the 95% confidence interval based on standard errors clustered at the country level.

Figure A5: Effect of Having Most Seats on Being in Ruling Coalition - National Parliaments Data



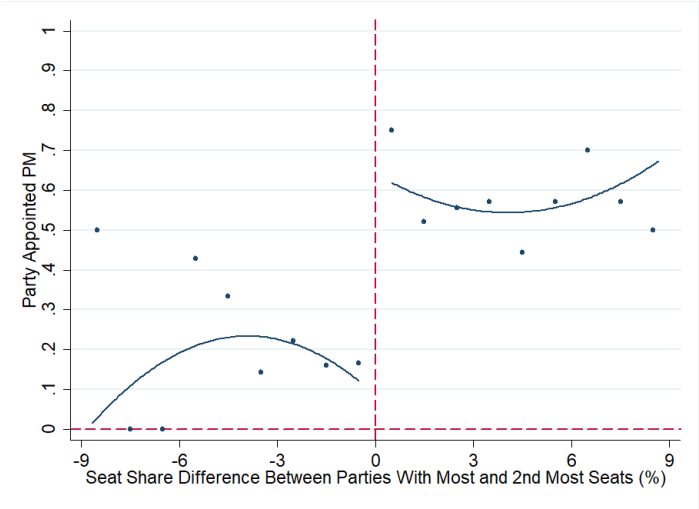
The unit of observation is a country-year-party. Sample is restricted to the two parties with the most seats in the parliament. The running variable (horizontal axis) is the difference in seat shares between the two most voted parties: positive for the party with the most seats and negative for the party with second most seats. Circles represent the local averages of a dummy indicating whether the party is part of the ruling coalition (represented in the cabinet). Averages are calculated within 1 p.p.-wide bins of seat share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A6: Covariate Balance: National Parliaments Data



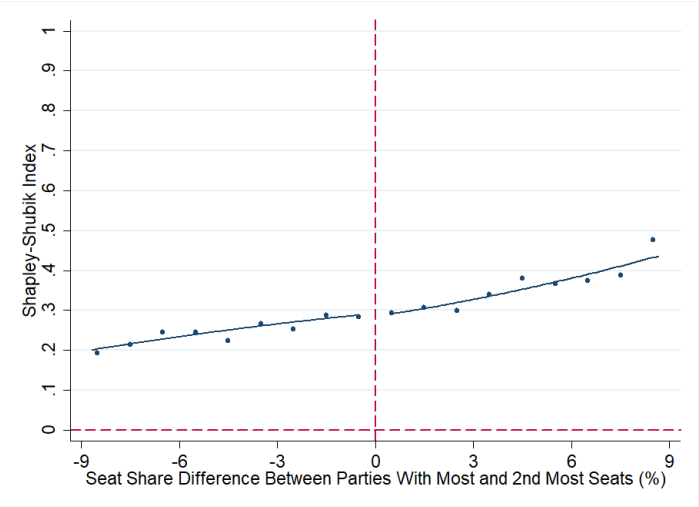
The unit of observation is a country-year-party. Sample is restricted to the two parties with the most seats in the parliament. The running variable (horizontal axis) is the difference in *seat* shares between the two most voted parties: positive for the party with the most seats and negative for the party with second most seats. Circles represent the local averages of the *predicted probability* of the party appointing the prime minister. Predictions are based on regressing a dummy indicating whether the party appointed the prime minister on a set of party family/ideology type (see text for details). Averages are calculated within 1 p.p.-wide bins of seat share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A7: Effect of Having Most Seats in Non-"Dominant" Cases - National Parliaments Data



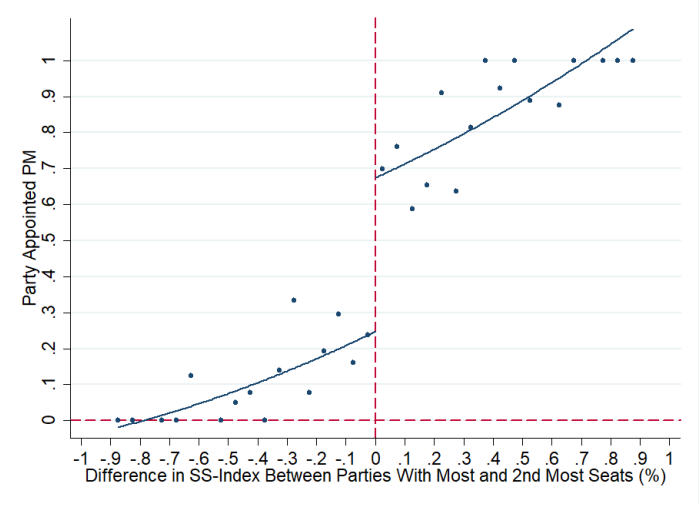
The unit of observation is a country-year-party. Sample is restricted to the two parties with the most seats in parliaments with non-"dominant" seat compositions (it excludes 51% of elections where the first and third placed parties can form a majority, while the second and third placed cannot). The running variable (horizontal axis) is the difference in seat shares between the two most voted parties: positive for the party with the most seats and negative for the party with second most seats. Circles represent the local averages of a dummy indicating whether the party appoints the prime minister. Averages are calculated within 1 p.p.-wide bins of seat share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A8: Effect of Having Most Seats on Shapley-Shubik Index- National Parliaments Data



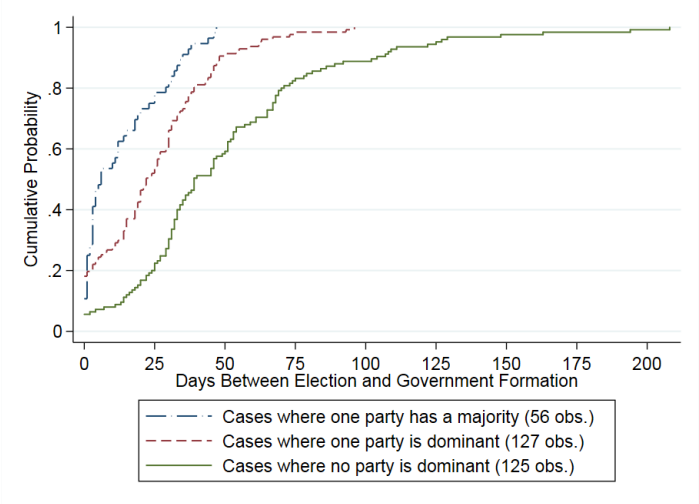
The unit of observation is a party-country-year. Sample is restricted to the two parties with the most seats in the parliament. The running variable (horizontal axis) is the difference in *seat* shares between the two parties with the most seats: positive with the most seats and negative for the party with the second most number of seats. Circles represent the local averages of the Shapley-Shubik index. Averages are calculated within 1 p.p.-wide bins of seat share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A9: Effect of Having the Most Seats, Shapley-Shubik Index as Running Variable - National Parliaments Data



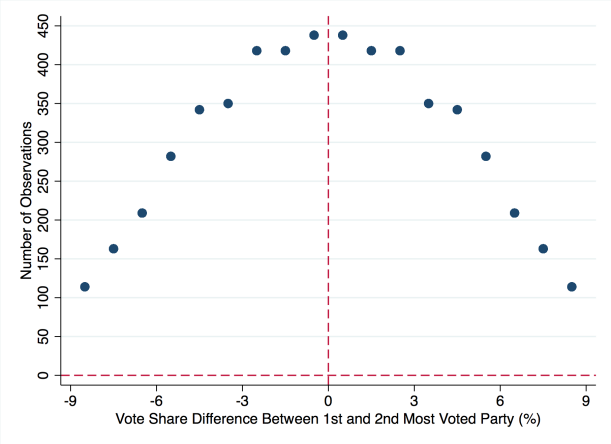
The unit of observation is a party-country-year. Sample is restricted to the two parties with the most seats in the parliament. The running variable (horizontal axis) is the difference in the Shapley-Shubik index between the two parties with the most seats: positive with the most seats and negative for the party with the second most number of seats. Circles represent the local averages of a dummy indicating whether the party appoints the prime minister. Averages are calculated within 1 p.p.-wide bins (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A10: CDF of Bargaining Delays in Government Formation - National Parliaments Data

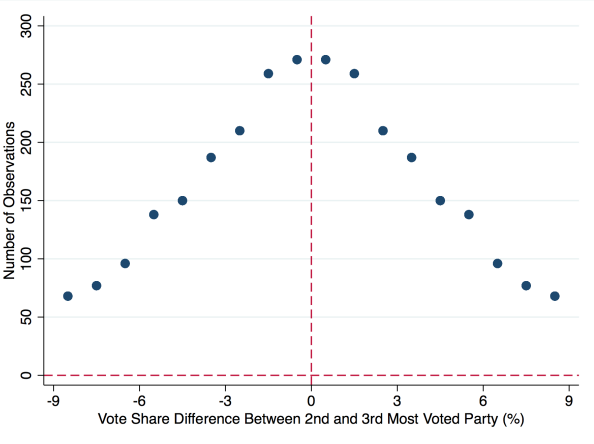


The unit of observation is a country-year. The figure plots the cumulative distribution function (CDF) of the time between election and government formation for the cases where: i) one party in the legislature has a majority of votes, ii) the party with most seats is "dominant" (i.e., can form a majority with the party with the third most seats, while the second most voted cannot do the same), iii) no party is dominant.

Figure A11: Histograms: Distribution of Running Variable



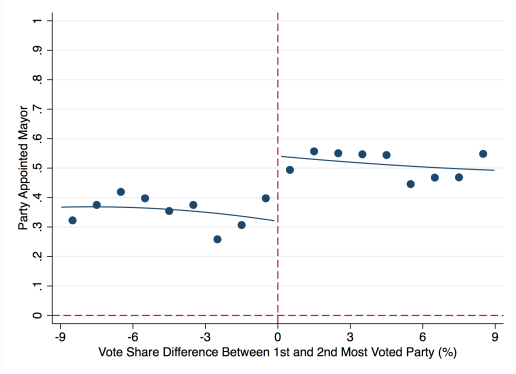
(a) 1st-vs.-2nd Histogram



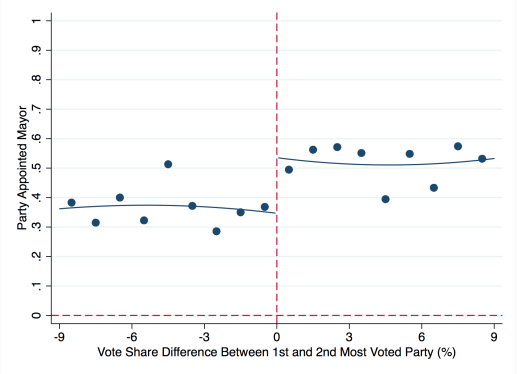
(b) 2nd-vs.-3rd Histogram

The unit of observation is a party-municipality-year. Each circle in Panel A (Panel B) represents the number of observations in the respective circle on Figure 1 (Figure 2) of the main text. Panel A (Panel B) restricts the sample to the two most voted (second and third most voted) parties in elections in which they tied in seats. Panel B further restricts the sample to elections where the most voted party did not obtain a majority of seats. The running variable (horizontal axis) is the difference in vote shares between the first and second (Panel A) or second and third (Panel B) most voted parties. Circle represents the number of observations in each 1 p.p.-wide bin of vote share difference.

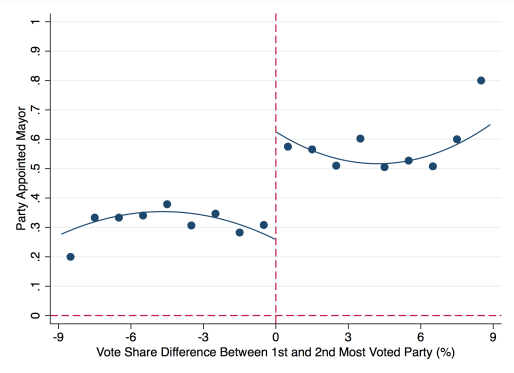
Figure A12: Effect of Being Most Voted, Heterogeneity by Council Size



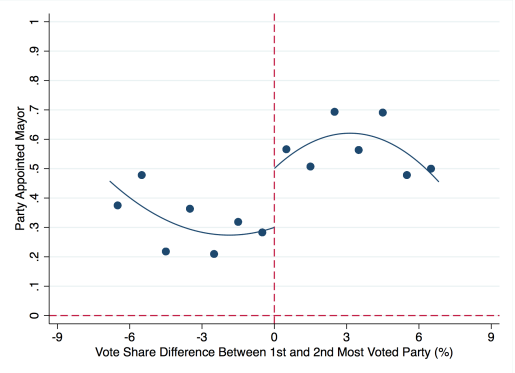
(a) 7-Member Councils



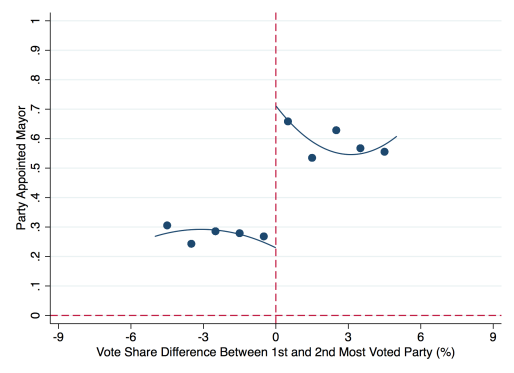
(b) 9-Member Councils



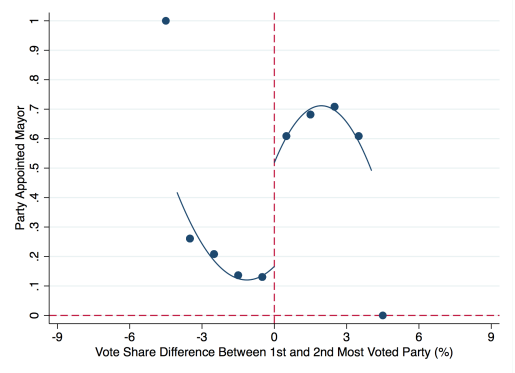
(c) 11-Member Councils



(d) 13-Member Councils



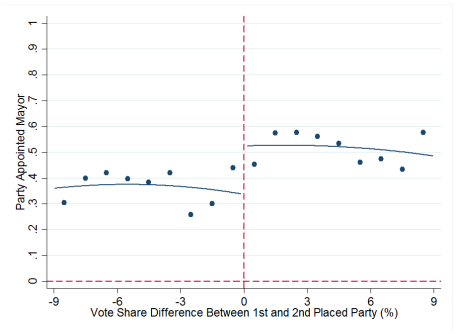
(e) 17-Member Councils



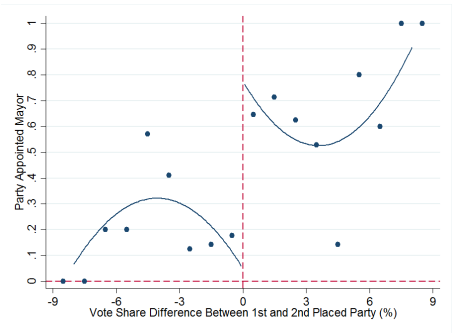
(f) 21-Member Councils

The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties in elections in which they tied in seats. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties: positive for the most voted party and negative for the second most voted. Circles represent the local averages of a dummy indicating whether the party appointed the mayor. Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data. Each panel restricts the sample to elections with a specific council size. We report plots for all council sizes with a sample of at least 150 observations (75 elections).

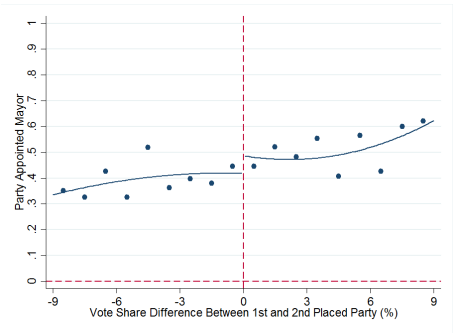
Figure A13: Effect of First-Place, by Council Type



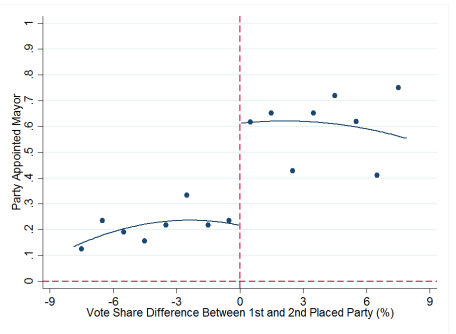
(a) 3-3-1 Councils



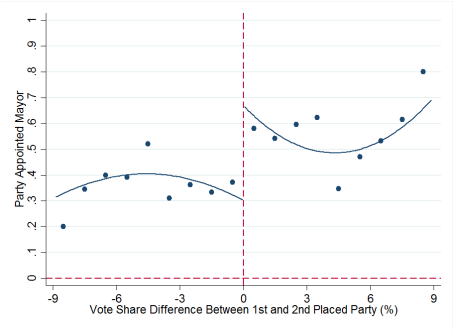
(b) 3-3-3 Councils



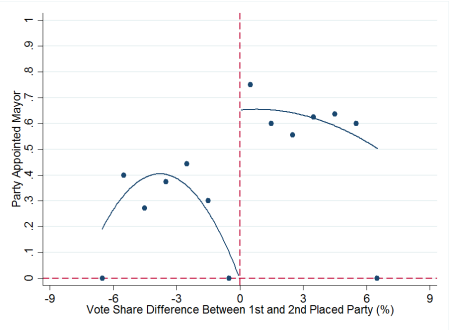
(c) 4-4-1 Councils



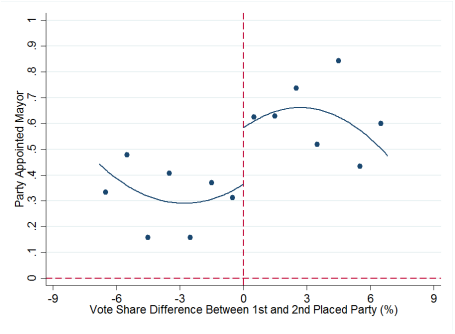
(d) 4-4-3 Councils



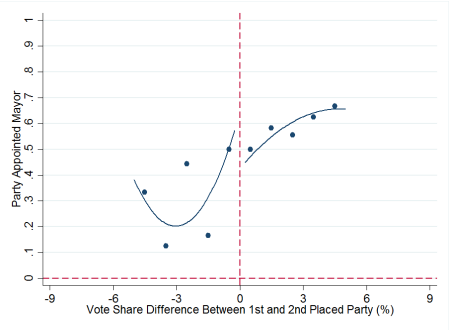
(e) 5-5-1 Councils



(f) 5-5-3 Councils



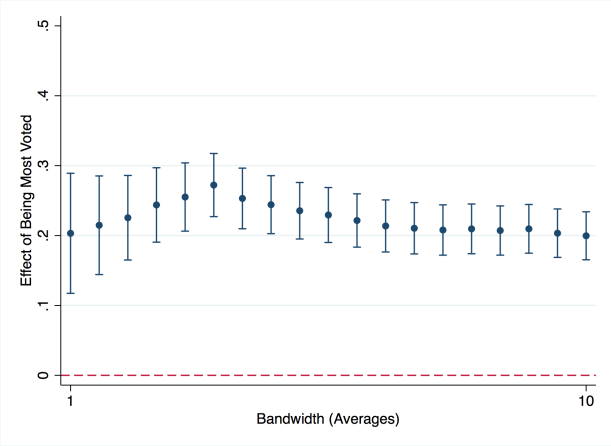
(g) 6-6-1 Councils



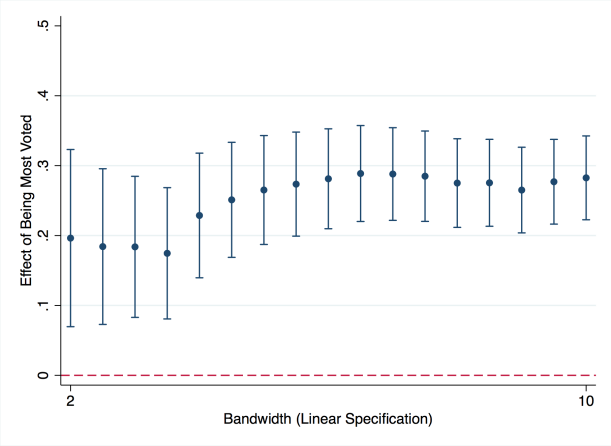
(h) 8-8-1 Councils

Notes are same to those on Figure A12, except each panel restricts the sample to councils with a particular seat configuration. We report plots for all configurations with a sample of at least 90 observations (45 elections).

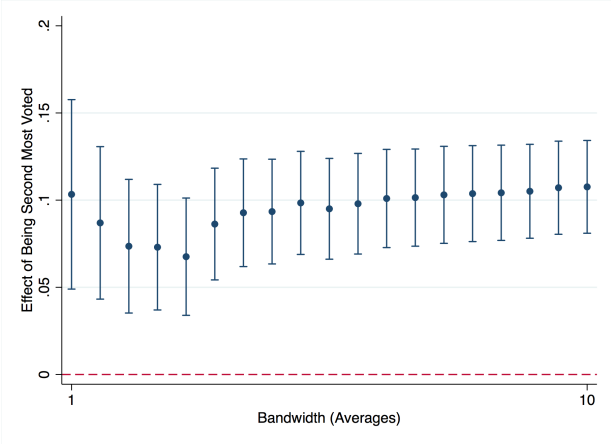
Figure A14: Robustness to Bandwidth Choice.



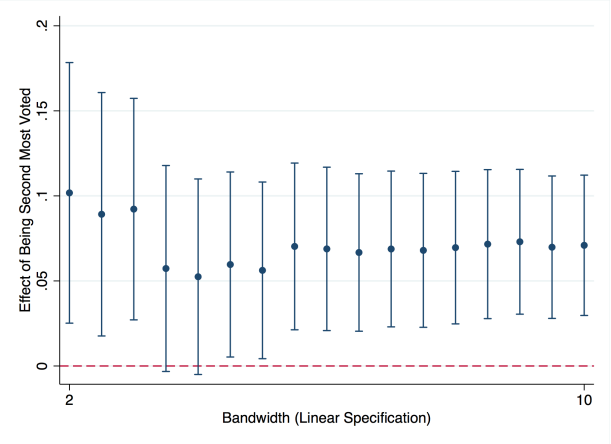
(a) Effect of 1st vs. 2nd (Specification: Means)



(b) Effect of 1st vs. 2nd (Specification: Linear)



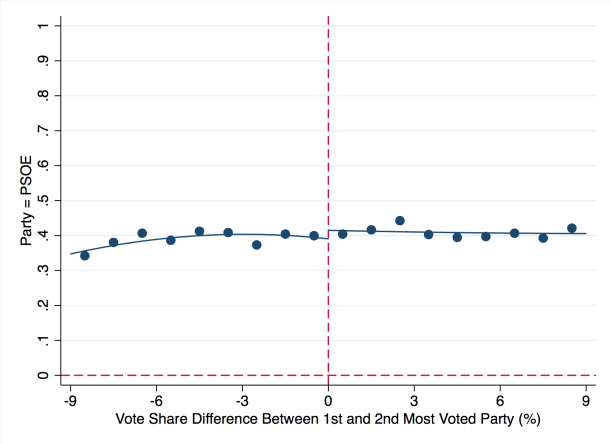
(c) Effect of 2nd vs. 3rd (Specification: Means)



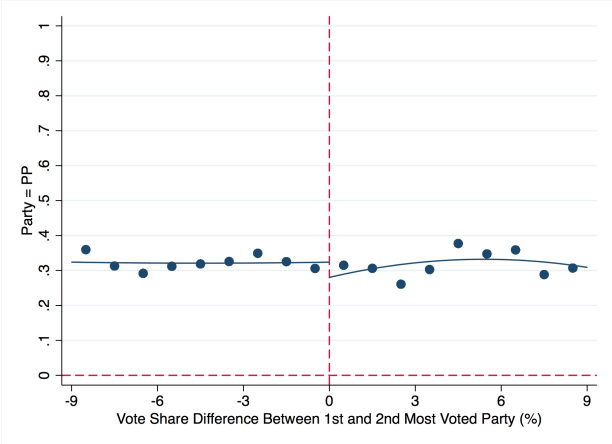
(d) Effect of 2nd vs. 3rd (Specification: Linear)

Circles represent estimated effects, using different bandwidth choices (horizontal axis). Whiskers represent the 95% confidence interval based on standard errors clustered at the municipality level.

Figure A15: Covariate Balance (1st vs. 2nd): Placebo "Effect" on Party Identity



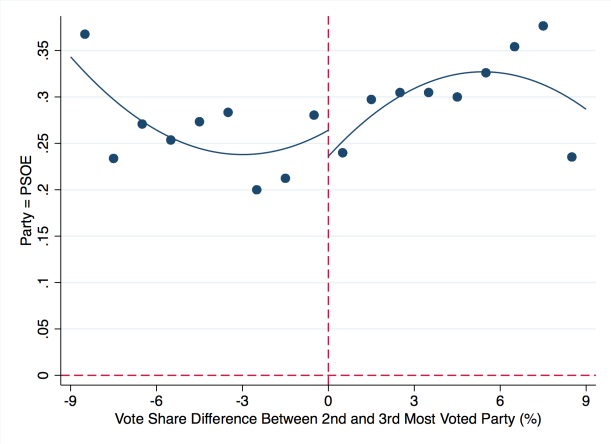
(a) Indicator for party being the PSOE



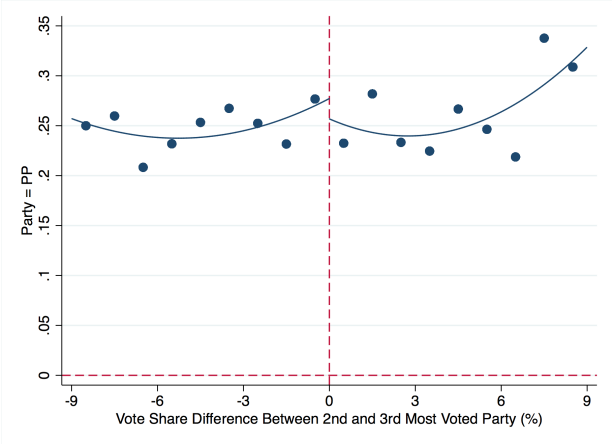
(b) Indicator for party being the PP

The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties in elections in which they tied in seats. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties: positive for the most voted party and negative for the second most voted. Circles represent the local averages of a dummy indicating whether the observation's party is the *Partido Socialista Obrero Español* (Panel A) or *Partido Popular* (Panel B). Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A16: Covariate Balance (2nd vs. 3rd): Placebo "Effect" on Party Identity



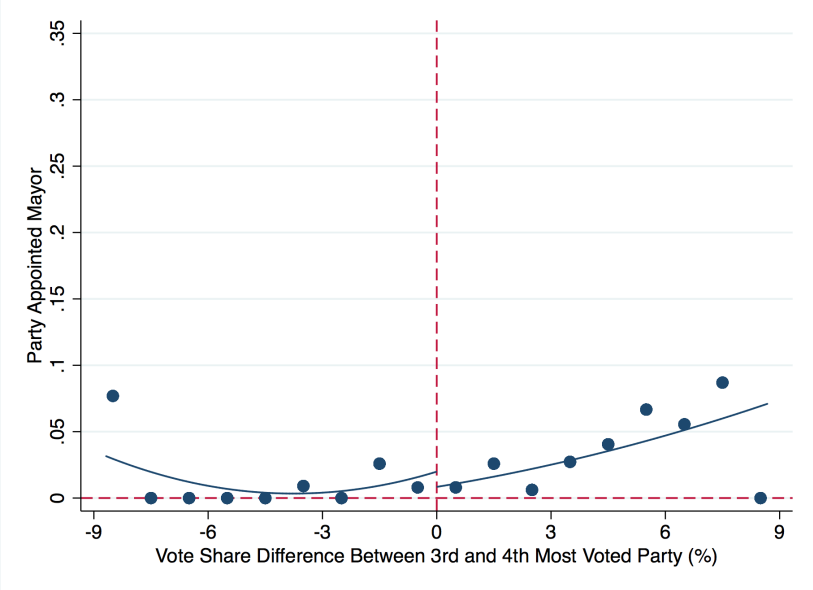
(a) Observation belongs to the PSOE



(b) Observation belongs to the PP

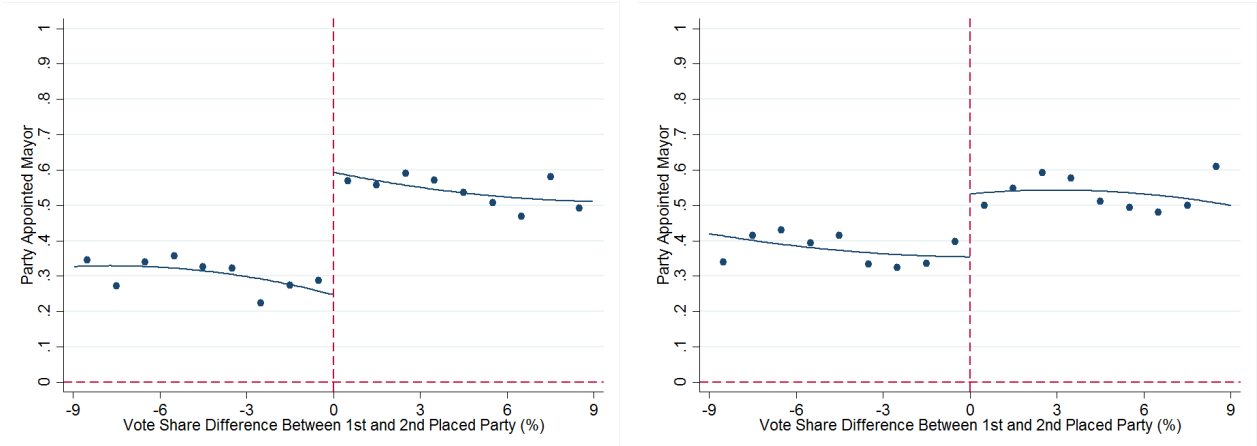
The unit of observation is a party-municipality-year. Sample is restricted to the second and third most voted parties in elections in which they tied in seats and the most voted party did not obtain a majority of seats. The running variable (horizontal axis) is the difference in vote shares between the second and third most voted parties: positive for the second most voted party and negative for the third most voted. Circles represent the local averages of a dummy indicating whether the observation's party is the *Partido Socialista Obrero Español* (Panel A) or *Partido Popular* (Panel B). Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A17: Effect of Being Third Most Voted: Third versus Fourth Place



The unit of observation is a party-municipality-year. Sample is restricted to the third and fourth most voted parties in elections in which they tied in seats and the most voted party did not obtain a majority of seats. The running variable (horizontal axis) is the difference in vote shares between the third and fourth most voted parties: positive for the third most voted party and negative for the fourth most voted. Circles represent the local averages of a dummy indicating whether the party appoints the mayor. Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A18: Effect Heterogeneity by Third-Placed Party Vote Share

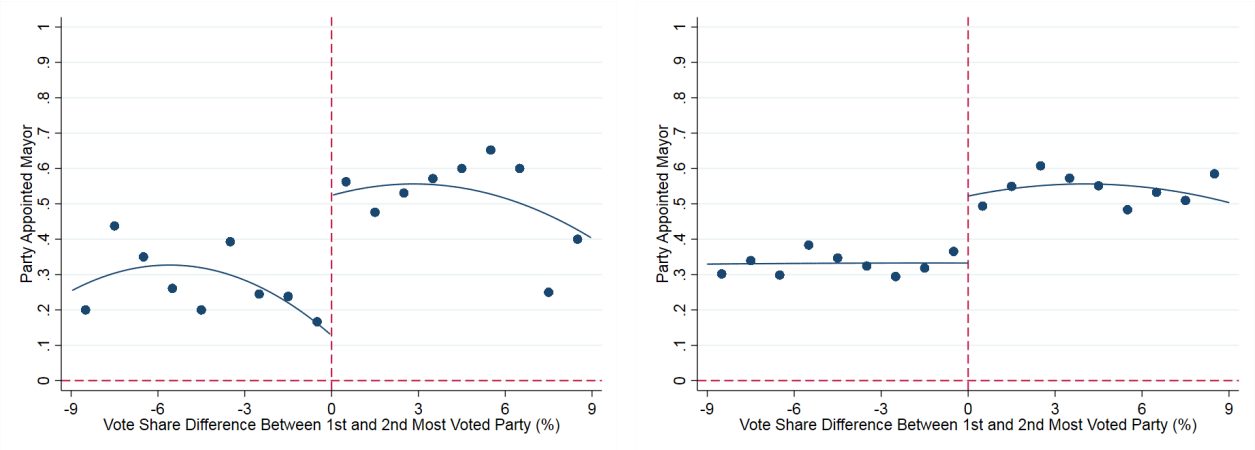


(a) Third placed party vote share above median

(b) Third placed party vote share below median

The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties in elections in which they tied in seats. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties: positive for the most voted party and negative for the second most voted. Circles represent the local averages of a dummy indicating whether the party appoints the mayor. Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data. In Panel A (Panel B), sample is further restricted to elections where the third most voted party has vote share above (below) the median of the sample used in column (1) in Table 1: 16.5%.

Figure A19: Effect Heterogeneity by Whether Third-Placed Party was (or will be) the Most Voted

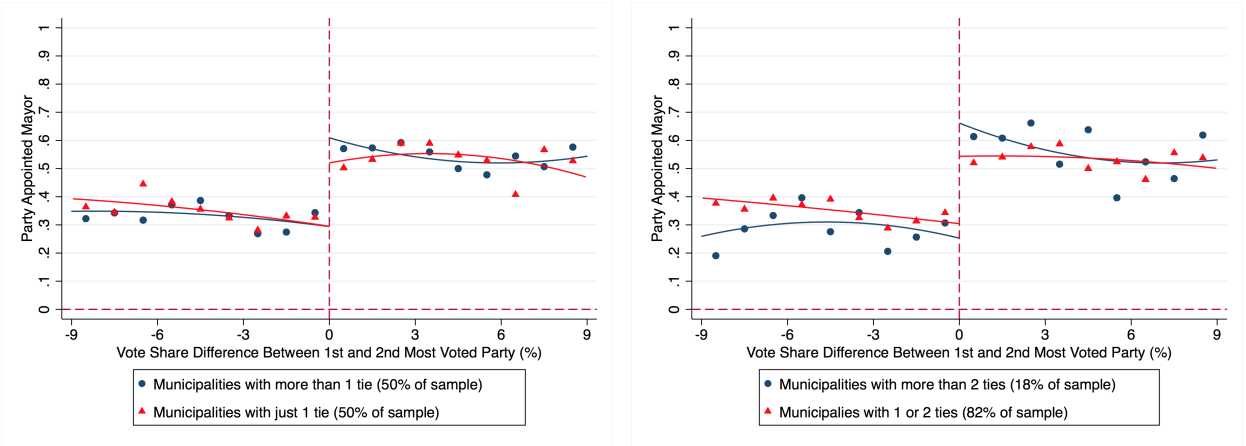


(a) Third placed party vote is the most-voted in at least one of he next (or last) three elections

(b) Third placed party vote is not the most-voted in any of the next (or last) three elections

The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties in elections in which they tied in seats. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties: positive for the most voted party and negative for the second most voted. Circles represent the local averages of a dummy indicating whether the party appoints the mayor. Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data. In Panel A (Panel B), sample is further restricted to elections where the third-placed party has ever (never) been the most voted party at any of the three previous or subsequent elections.

Figure A20: Effect Heterogeneity by Frequency of Ties in Municipality

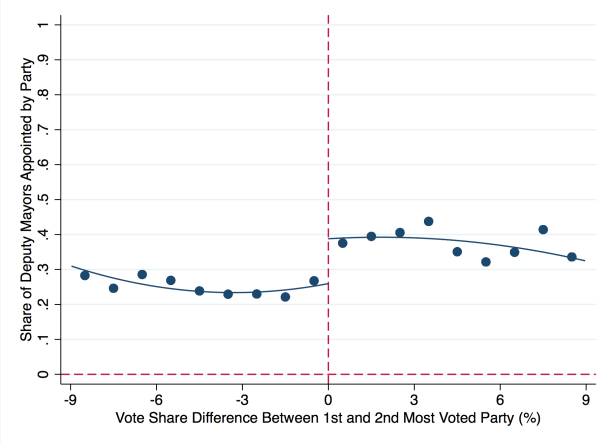


(a) Effect of Most Voted, by Frequency of Ties

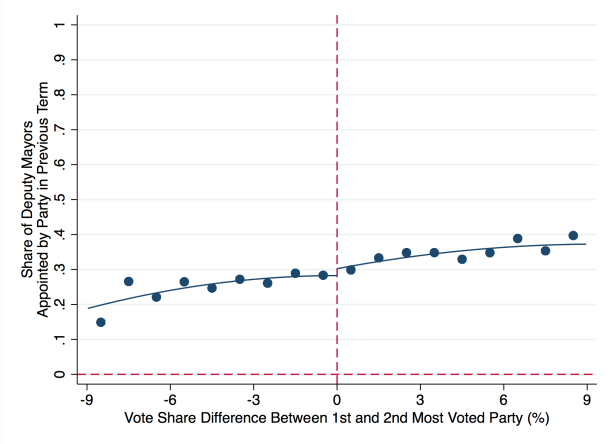
(b) Effect of Most Voted, by Frequency of Ties

The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties in elections in which they tied in seats. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties: positive for the most voted party and negative for the second most voted. Markers represent the local averages of a dummy indicating whether the party appoints the mayor. Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data. In Panel A, red triangles restrict the sample to municipalities with only one occurrence of the two most voted parties tying in seats in the sample period. Blue circles restrict the sample to municipalities where more than one tie in seats occurred. In Panel B, red triangles restrict the sample to municipalities where more than two ties occurred in the sample period, while the blue circles restrict it to cases where one or two ties occurred. See text for further details.

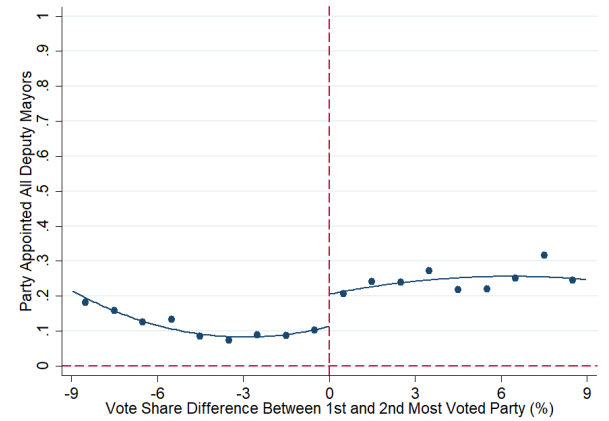
Figure A21: Effect of Being Most Voted on Deputy Mayors' Allocation



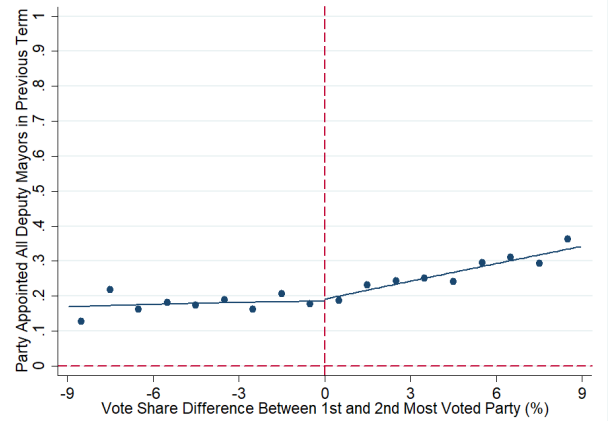
(a) Effect of Most Voted on Share of Deputy Mayors



(b) Placebo Test: "Effect" of Most Voted on Lagged Share of Deputy Mayors



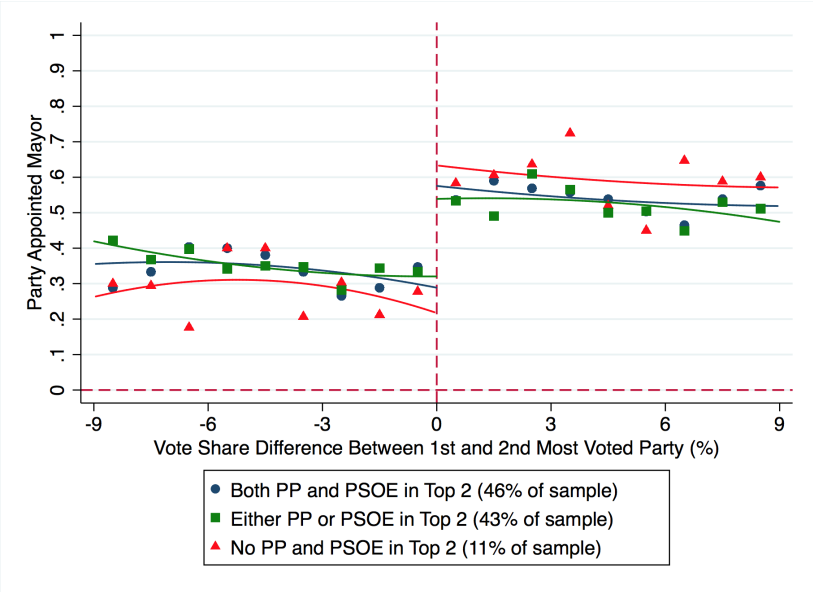
(c) Effect of Most Voted on Indicator for Appointing all Deputy Mayors



(d) Placebo Test: "Effect" of Most Voted on Indicator for Appointing all Deputy Mayors

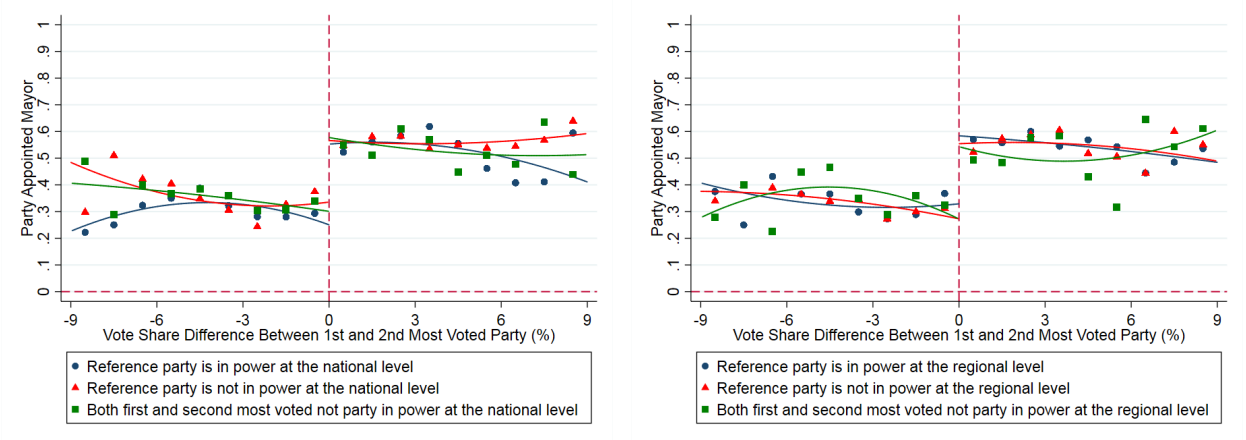
The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties in elections in which they tied in seats. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties: positive for the most voted party and negative for the second most voted. Circles in Panel A and B represent the local averages of the share of deputy mayors belonging to the party (Panel A) or that belonged to the party in the previous $(t-1)$ term (Panel B). Circles in Panel C and D represent the local averages of an indicator for all deputy mayors belonging to the party (Panel C) or all having belonged to the party in the previous $(t-1)$ term (Panel D). Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data.

Figure A22: Effect of Being Most Voted: Heterogeneity by Party Identity



The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties in elections in which they tied in seats. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties: positive for the most voted party and negative for the second most voted. Markers represent the local averages of a dummy indicating whether the party appoints the mayor. Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data. The three separate plots are for the cases where both, either, or neither the *Partido Popular (PP)* and/or the *Partido Socialista Obrero Español (PSOE)* are amongst the two most voted parties.

Figure A23: Is the Most Voted Effect Different for Powerful Parties?



(a) National (b) Regional

The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties in elections in which they tied in seats. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties: positive for the most voted party and negative for the second most voted. Markers represent the local averages of a dummy indicating whether the party appoints the mayor. Averages are calculated within 1 p.p.-wide bins of vote share difference (horizontal axis). Continuous lines are a quadratic fit over the original (unbinned) data. In panel A (B), the three separate plots are for the cases where the reference was in power at the national (regional) level, was not in power at that level, and in which both parties were not in power at that level.

Table A1: Additional Results on Effect of Having Most Seats on Appointing Prime Minister

Dependent Variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: Excludes Bulgaria from Sample					
Party Appointed	0.205	0.318**	0.583**	0.392***	0.397**
Prime Minister		(0.131)	(0.210)	(0.124)	(0.147)
N		220	24	496	496
Panel B: Sample Restricted Countries in Diermeier and Merlo (2004)'s Sample					
Party Appointed	0.214	0.300**	0.625**	0.250	0.414**
Prime Minister		(0.127)	(0.205)	(0.138)	(0.155)
N		132	16	340	340
Panel C: Shapley-Shubik Index as Running Variable					
Party Appointed	0.181	0.431***	0.560***	0.373***	0.411***
Prime Minister		(0.138)	(0.0498)	(0.125)	(0.143)
N		254	504	504	504
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Standard errors clustered at the country level in parentheses. The unit of observation is a party-country-year. The sample is restricted to the two parties with the most seats in the parliament. Each figure in columns (1)-(4) reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. 2nd-Place Mean is the estimated value of the dependent variable for the party with the 2nd-most seats that tied with the party with most seats (using the specification in column 1). Optimal bandwidths are based on Imbens and Kalyanaraman (2012), being equal to 7.38%, 7.43%, and 0.183, for the three panels, respectively.

Table A2: Distribution of Council Sizes

		Number of Municipality-Elections		
Population	Number of Seats	Total	Tie in seats (1st/2nd)	Tie in seats (2nd/3rd)
251 to 1,000	7	15097	822	695
1,001 to 2,000	9	6773	662	214
2,001 to 5,000	11	7064	707	310
5,001 to 10,000	13	3674	365	183
10,001 to 20,000	17	2260	192	89
20,001 to 50,000	21	1369	93	47
50,001 to 100,000	25	469	34	16
100,000+	-	416	23	12

Source: Ley 7/1985, *Reguladora de las bases del régimen local*, article 179.
For municipalities with more than 100,000 inhabitants, one more seat is added for every additional 100,000 inhabitants or fraction thereof, adding one more if needed for odd number of seats.

Table A3: Effect of Being First (Instead of Second) Most Voted: Alternative Specifications

Dependent Variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: Main Outcome (mayor serves at least 3/4 of term)					
Party Appointed	0.353	0.185***	0.203***	0.295***	0.241***
Mayor		(0.058)	(0.044)	(0.036)	(0.046)
N		2028	876	5796	5796
Panel B: Outcome is appointing mayor for entire term					
Party Appointed	0.323	0.199***	0.217***	0.306***	0.254***
Mayor		(0.059)	(0.042)	(0.036)	(0.045)
N		1876	876	5796	5796
Panel C: Outcome is appointing mayor for longer than other parties					
Party Appointed	0.374	0.205***	0.221***	0.310***	0.268***
Mayor		(0.061)	(0.045)	(0.037)	(0.047)
N		1998	876	5796	5796
Panel D: Outcome is appointing initial mayor					
Party Appointed	0.360	0.242***	0.249***	0.343***	0.290***
Mayor		(0.062)	(0.045)	(0.037)	(0.047)
N		1892	876	5796	5796
Panel E: Main outcome, sample restricted to cases where "two parties out of top-3 needed for majority"					
Party Appointed	0.370	0.163***	0.200***	0.294***	0.238***
Mayor		(0.062)	(0.047)	(0.038)	(0.048)
N		1898	790	5472	5472
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Standard errors clustered at the municipality level in parentheses. The unit of observation is a party-municipality-year. The sample is restricted to the two most voted parties in elections in which they tied in seats. Each figure in columns (1)-(4) reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. 2nd-Place Mean is the estimated value of the dependent variable for a 2nd most voted party that tied with the 1st most voted party (using the specification in column 1). Optimal bandwidths are based on Imbens and Kalyanaraman (2012), being equal to 2.32%, 2.13%, 2.29%, 2.16%, and 2.38% for the five panels, respectively.

Table A4: Effect of Being First (Instead of Second) Most Voted: Including 1979 and 2015 Elections

Dependent Variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: Main Outcome (mayor serves at least 3/4 of term), with 1979 data					
Party Appointed	0.343	0.209***	0.204***	0.290***	0.242***
	(0.0621)	(0.0417)	(0.0347)	(0.0437)	
N		1882	980	6474	6474
Panel B: Outcome is appointing mayor for entire term, with 1979 data					
Party Appointed	0.330	0.205***	0.218***	0.300***	0.254***
	(0.0551)	(0.0405)	(0.0339)	(0.0427)	
N		2206	980	6474	6474
Panel C: Outcome is appointing mayor for longer than other parties, with 1979 data					
Party Appointed	0.366	0.223***	0.220***	0.304***	0.267***
	(0.0610)	(0.0427)	(0.0351)	(0.0447)	
N		2050	980	6474	6474
Panel D: Outcome is appointing initial mayor, with 1979 data					
Party Appointed	0.360	0.244***	0.245***	0.335***	0.288***
	(0.0565)	(0.0424)	(0.0352)	(0.0445)	
N		2282	980	6474	6474
Panel E: Outcome is appointing initial mayor, with 1979 and 2015 data					
Party Appointed	0.374	0.222***	0.233***	0.310***	0.267***
	(0.0543)	(0.0409)	(0.0341)	(0.0431)	
N		2464	1074	7004	7004
Panel F: Main outcome, sample restricted to cases where "two parties out of top-3 needed for majority", with 1979 data					
Party Appointed	0.357	0.193***	0.205***	0.295***	0.244***
	(0.0647)	(0.0443)	(0.0364)	(0.0459)	
N		1790	888	6110	6110
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Standard errors clustered at the municipality level in parentheses. The unit of observation is a party-municipality-year. The sample is restricted to the two most voted parties in elections in which they tied in seats. Each figure in columns (1)-(4) reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. 2nd-Place Mean is the estimated value of the dependent variable for a 2nd most voted party that tied with the 1st most voted party (using the specification in column 1). Optimal bandwidths are based on Imbens and Kalyanaraman (2012), being equal to 2.32%, 2.13%, 2.29%, 2.16%, and 2.38% for the five panels, respectively.

Table A5: Effect of Being Most Voted: Cases with a Left-Wing Majority

Dependent Variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: Effect for PSOE (conditional IU being third most voted)					
PSOE Appointed	0.543	0.267*	0.248**	0.417***	0.264*
Mayor		(0.153)	(0.118)	(0.109)	(0.145)
N		155	64	423	423
Panel B: Effect for PP (conditional IU being third most voted)					
PP Appointed	0.102	0.242*	0.231**	0.313***	0.246*
Mayor		(0.146)	(0.110)	(0.101)	(0.139)
N		155	64	423	423
p-value: test of equal effects		0.7826	0.8097	0.1469	0.8412
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Standard errors clustered at the municipality level in parentheses. The unit of observation is a party-municipality-year. The sample is restricted to elections in which the two most-voted parties tie in seats and the third-placed party is the *Izquierda Unida (IU)*. Panel A uses only observations regarding the *Partido Socialista Obrero Español (PSOE)* in elections where the *Partido Popular (PP)* is the other "top two" party. Panel B uses only observations regarding the *PP* in elections where the *PSOE* is the other top-two party. Each figure in columns (1)-(4) reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. 2nd-Place Mean is the estimated value of the dependent variable for the 2nd most voted party. The optimal bandwidth is calculated based on the entire sample and is 2.32% (Imbens and Kalyanaraman 2012).

Table A6: Comparing Magnitude of Effects:
Effect of Being Most Voted, by Legislature Type

Dependent Variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: First and second most voted tied in seats					
Party Appointed	0.353	0.185***	0.203***	0.295***	0.241***
Mayor		(0.059)	(0.044)	(0.037)	(0.046)
N		2028	876	5796	5796
Panel B: Most voted has one more seat than second most voted, but no more "real" bargaining power					
Party Appointed	0.259	0.305***	0.254***	0.431***	0.352***
Mayor		(0.069)	(0.077)	(0.046)	(0.059)
N		1424	252	5862	5862
Panel C: Most voted has one more seat than second most voted and more "real" bargaining power					
Party Appointed	0.120	0.667***	0.700***	0.618***	0.650***
Mayor		(0.049)	(0.073)	(0.036)	(0.045)
N		1648	160	6382	6382
Panel D: Most voted has a majority of seats					
Party Appointed	0.006	0.978***	0.982***	0.977***	0.976***
Mayor		(0.003)	(0.009)	(0.003)	(0.004)
N		26806	788	56204	56204
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Standard errors clustered at the municipality level in parentheses. The unit of observation is a party-municipality-year. The sample is restricted to the two most voted parties. Each panel focus on a different case of seat composition in the legislature. Each figure in columns (1)-(4) reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. 2nd-Place Mean is the estimated value of the dependent variable for a 2nd most voted party that tied with the most voted party (using the specification in column 1). Optimal bandwidths are based on Imbens and Kalyanaraman (2012), being equal to 2.32%, 4.48%, 7.03%, and 23.18% for the four panels variables, respectively.

Table A7: Effect of Being Most Voted on Appointing the Mayor, by Probability of Third-Placed Becoming the Most Voted

Dependent Variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: Third most voted party vote is the most voted at some election (t-3, t+3)					
Party Appointed	0.170	0.494***	0.396***	0.351***	0.453***
		(0.152)	(0.111)	(0.117)	(0.153)
N		218	96	550	550
Panel B: Third most voted party vote is never the most voted (t-3, t+3)					
Party Appointed	0.400	0.0541	0.128*	0.170**	0.111
		(0.0976)	(0.0738)	(0.0689)	(0.0946)
N		804	312	2298	2298
p-value: test of equal effects		0.0137	0.0411	0.1818	0.0543
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Standard errors clustered at the municipality level in parentheses. The unit of observation is a party-municipality-year. The sample is restricted to the two most voted parties in elections in which they tied in seats. In Panel A (Panel B), sample is further restricted to elections where the third-placed party has ever (never) been the most voted party at any of the three previous or subsequent elections. Each figure in columns (1)-(4) reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. 2nd-Place Mean is the estimated value of the dependent variable for a 2nd most voted party that tied with the 1st most voted party (using the specification in column 1). The optimal bandwidth is calculated based on the entire sample and is 2.32% (Imbens and Kalyanaraman 2012).

Table A8: Effect of Being Most Voted on Deputy Mayors' Allocation

Dependent Variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: Outcome is share of deputy mayors					
Party Share of Deputy	0.278	0.094**	0.108***	0.183***	0.125***
Mayors		(0.045)	(0.034)	(0.029)	(0.036)
N		1732	758	4930	4930
Panel B: Outcome is share of deputy mayors in previous period (placebo test)					
Party Share of Deputy	0.294	0.007	0.015	0.008	0.026
Mayors, $t-1$		(0.037)	(0.035)	(0.029)	(0.036)
N		2544	758	4930	4930
Panel C: Outcome is indicator for appointing all deputy mayors					
Party Appointed All	0.110	0.079**	0.103***	0.151***	0.091***
Deputy Mayors		(0.036)	(0.029)	(0.026)	(0.030)
N		1814	758	4930	4930
Panel D: Outcome is indicator for appointing all deputy mayors in previous period (placebo test)					
Party Appointed All	0.191	-0.009	0.010	-0.003	0.017
Deputy Mayors, $t-1$		(0.035)	(0.030)	(0.027)	(0.034)
N		2444	758	4930	4930
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Standard errors clustered at the municipality level in parentheses. The unit of observation is a party-municipality-year. The sample is restricted to the two most voted parties in elections in which they tied in seats. Each figure in columns (1)-(4) reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. 2nd-Place Mean is the estimated value of the dependent variable for a 2nd most voted party that tied with the most voted party (using the specification in column 1). Optimal bandwidths are based on Imbens and Kalyanaraman (2012), being equal to 2.28%, 3.47%, 2.37%, and 3.30% for the four panels, respectively.

Table A9: Municipalities That Appoint Most-Voted Mayors Have Fewer Instances of Corruption

	Dependent Variable is Corruption Indicator				
	(1)	(2)	(3)	(4)	(5)
$Norm_i$	-0.0253**	-0.0264**	-0.0259**	-0.0253**	-0.0209**
	(0.0117)	(0.0109)	(0.0115)	(0.0118)	(0.0103)
N	2390	2390	2390	2390	2390
Dep. Variable Mean for $norm_i = 0$ cases	0.0693	0.0693	0.0693	0.0693	0.0693
Province FE		Y			
Year FE		Y			
Province × Year FE			Y	Y	Y
Mayor's Party FE				Y	Y
Additional Controls:					Y

Standard errors clustered at the municipality level in parentheses. The unit of observation is a municipality-year. Each column shows the results of a regression of an indicator of whether a corruption case was uncovered in municipality i during the electoral term starting at t against $norm_i$, the share of times that, after an election where the two most voted parties were only 1 p.p. vote share away and tied in seats, the most voted party appointed the mayor. Column (2) adds province and year fixe effects. Column (3) adds a set of dummies for each province-year interaction. Column (4) adds a set of dummies indicating the party of the appointed mayor. Column (5) adds the additional controls: the logarithm of the municipality's population, voter turnout (as a share of registered voters), and a set of four dummies, indicating whether the observation fits in one of the four categories described in Figure 5 (e.g., whether one party has a majority or the top two parties are tied in seats).