Review of Economic Studies (2020) **87**, 1261–1295 doi:10.1093/restud/rdz004

Advance access publication 29 January 2019

Rank Effects in Bargaining: Evidence from Government Formation

THOMAS FUJIWARA
Princeton, CIFAR, and NBER
CARLOS SANZ
Bank of Spain

First version received June 2018*; Editorial decision November* 2018*; Accepted January* 2019 (*Eds.*)

Theories of multilateral bargaining and coalition formation applied to legislatures predict that parties' seat shares determine their bargaining power. We present findings that are difficult to reconcile with this prediction, but consistent with a norm prescribing that "the most voted party should form the government". We first present case studies from several countries and regression discontinuity design-based evidence from twenty-eight national European parliaments. We then focus on 2,898 Spanish municipal elections in which two parties *tie* in the number of seats. We find that the party with slightly more general election votes is substantially more likely to appoint the mayor. Since tied parties should (on average) have equal bargaining power, this identifies the effect of being labeled the most voted. This effect is comparable to that of obtaining an additional seat, and is also present when a right-wing party is the most voted and the second and third most voted parties are allied left-wing parties who can form a combined majority. A model where elections both aggregate information and discipline incumbents can rationalize our results and yields additional predictions we take to the data, such as voters punishing second most voted parties that appoint mayors.

Key words: Norms, Bargaining, Government formation, Rank effects.

JEL Codes: D70, D72, D90.

1. INTRODUCTION

Multilateral bargaining and coalition formation play a role in many economic and political environments, such as mergers, trade negotiations, and conflict. Theories on the topic focus on how coalition payoffs interact with bargaining procedures to determine outcomes. Less attention is devoted to the role of social norms, informal rules, and conventions about what is an "appropriate" or "justified" outcome. According to a recent survey, "the impact of such norms or processes on equilibrium coalition structures is nontrivial, interesting, and largely unexplored" (Ray and Vohra, 2014).

The editor in charge of this paper was Nicola Gennaioli.

This article studies bargaining and coalition formation in legislatures, which are key determinants of political outcomes and public policy. We provide evidence that bargaining outcomes are consistent with a norm prescribing that the most voted party should hold the executive in a parliamentary system of government (*e.g.* the prime minister should be a member of the most voted party).

Politicians commonly appeal to this prescription as "democratic" or "respecting the will of the people". For example, Canadian prime minister Stephen Harper stated in an interview that "my position has always been, if we win the most seats, I will expect to form the government. And if we don't, I won't". When the interviewer pressed him by asking "So, even as the current government, if you're just a couple of seats behind, you wouldn't try to figure out a way to…", Harper cuts in to say "No. […] I would not serve as Prime Minister […] we ask people to make a choice of a government, and so I think that the party that wins the most seats should form the government" (Wells, 2015).

Although Mr. Harper's states that a plurality (*the most seats*) legitimizes a government, formal Canadian institutions are majoritarian: a prime minister requires the support of more than half of parliament. Furthermore, Mr. Harper's motivations are not directly attributable to his self interest, as his party was not expected to win the most seats. Politicians and voters in other contexts make arguments similar to Mr. Harper's. The next section provides examples from New Zealand, U.K., Italy, and Spain.

This example illustrates a broader issue that Maskin and Sen (2016) refer to as the "serious confusion when a plurality win is marketed as a majority victory" and their argument that "understanding of the critical difference between a plurality and a majority could improve politics around the world". Maskin and Sen's (2016) point goes beyond government formation, extending to issues as the legitimacy of the BJP's hindu-nationalism in India, the effects of the Muslim Brotherhood on Egypt's democratic institutions, and Donald Trump's nomination as presidential candidate in the USA.

To study the informal prescription that the "party with most seats should form government", we first leverage a regression discontinuity design (RDD) embedded in government formation episodes in twenty-eight national European parliaments in the postwar period to show that the *party with the most seats is more than twice as likely to form a government* than the party with the second most number of seats. Next, we focus on Spanish municipalities, which provide a sharper identification strategy. Each municipality elects a council by (closed list) proportional representation under the D'Hondt apportionment rule in a single-district election.[1] In its first meeting, the council selects, by *majority* rule, one of its members to hold the powerful position of mayor. We focus on 2,898 elections where the *two most voted parties tied in number seats*, exploiting a RDD comparing the probability of appointing the mayor between the first and second most voted parties that are few votes apart. Existing theories predict both parties have equal bargaining power. Our results indicate that *the party with slightly more votes is roughly 20 p.p. more likely to appoint the mayor.*

As an example, consider a council with eleven seats where parties *A*, *B*, and *C* obtained, respectively, 42.0%, 41.9%, and 16.1% of the votes. Parties *A* and *B* thus receive five seats each, and party *C*, one seat. In theory, *A* and *B* should have equal bargaining power and be equally likely to form a government with majority support. Thus, the probability *A* and *B* appoints the mayor should be the same. Our findings indicate that this probability is 55% for party *A* and 35% for party *B*.

1. Proportional representation allocates seats to parties in proportion to their votes. Exact proportionality is not possible given integer constraints and D'Hondt rule (described in Section 4) allocates seats respecting such constraints.

Spanish electoral rules do not award an *obvious* advantage to the most voted. The only formal rule treating parties differentially by vote ranks is when defining the "status quo". If a *majority* of council members cannot agree on a mayor, the leader of the most voted party is appointed.[2] While at first pass it may seem to play an important role, there are several reasons this rule is unlikely to explain our results. The most compelling is that we find a similar effect of being second (instead of third) most voted (when both parties are tied in seats). There is no rule favouring the second most voted party and hence the status quo rule cannot fully explain our evidence.[3]

Intuitively, the RDD isolates a comparison between two groups of parties that should have, on average, equal ex-ante bargaining power. However, one group can be thought of as being randomly assigned the "most voted" label—in a setting where being the most voted should be irrelevant. Comparing most voted parties to lower ranked ones in general (*i.e.* including cases where parties' vote shares differ substantially) is potentially confounded by the different number of seats and other factors that lead one party to outperform others. However, these are held constant in our RDD.

Additionally, by focusing on parties that almost tied in votes, our results indicate that agents respond to *rankings* that provide no additional information conditional on the publicly available variable that fully determines the ranking (votes). This adds to evidence of rank-based decision-making in politics (Anagol and Fujiwara, 2016; Folke *et al*., 2016; Meriläinen and Tukiainen, 2018; Pons and Tricaud, 2018) and is consistent with models where agents make choices over-weighting salient attributes (Bordalo *et al*., 2013). Our results thus speak to how narrow victories are interpreted (or marketed) by some as "the will of the people". For example, only 51.9% of votes in a British 2016 referendum were for leaving the European Union and only 50.6% of the votes in a 1995 Quebec referendum were against independence from Canada. These results went on to be interpreted (by some) as "what the people has chosen". After becoming the U.S. president with 50.7% of the vote, George W. Bush justified his agenda by saying "the people made it clear what they wanted" and "I earned capital in the campaign, political capital, and now I intend to spend it" (Stevenson, 2004).

Our results are difficult to reconcile with existing theories of legislative bargaining and government formation. Their starting point is the number of seats held by voting blocks (political parties). Given legislative procedures, seat distributions fully determines which coalitions can be formed and hence parties' bargaining power. Such theories thus take parties' *seats* as the primitives and ignore the role of the votes that lead to these seat allocations: seat distributions should be sufficient to study bargaining outcomes.[4] Seat allocations are also the focus of empirical work. For example, Gamson's Law is the empirical regularity that cabinet positions are distributed proportionally to coalition parties' contribution of *seats*. Structural models of government formation also take seat allocations as the starting point. Indeed, standard datasets

2. Only party leaders can become mayor. Thus, in our example only three members of the eleven-member council can be mayor. Section 4 describes these electoral rules in detail.

3. Sections 4 and 7 discuss additional reasons why the status quo rule cannot explain our results. Mayors require continuous support from a majority during the term. In our example, parties *B* and *C* could appoint a *B* mayor anytime they agree to do so, making it less likely the status quo binds. We also provide evidence consistent with voters punishing second most voted parties that appoints mayors. It is unclear why the status quo rule would generate this result.

4. This applies to both non-cooperative and cooperative theories (*e.g.* core or bargaining sets). Ray and Vohra (2014) survey coalition formation and Laver (1998) surveys government formation in particular. Examples of models of legislative bargaining and government formation are Baron and Ferejohn (1989), Austen-Smith and Banks (1990), Laver and Shepsle (1990), Baron (1991, 1993), Merlo and Wilson (1995), Morelli (1999), and Snyder Jr *et al*. (2005).

used in this literature usually contain information only on parties' seat allocations (and not their general election votes).[5]

Moreover, in a substantial number of cases, parties act in a manner consistent with the prescription even when it goes against programmatic affinity between parties. The effect of being the most voted is of similar magnitude when we restrict attention to cases where the most voted party is the main right-wing party (*Partido Popular*—PP), while the second and third most voted parties are, respectively, the main left-wing party (*Partido Socialista Obrero Español*—PSOE) and its common leftist ally (*Izquierda Unida*—IU). This implies that, even though the two left-wing parties have a combined majority that could appoint the mayor, the most-voted right-wing party frequently appoints the mayor instead.[6]

To gauge the magnitude of the effect of being most-voted, we compare it to the effect of obtaining one additional seat, which can also be identified in a RDD. Having a plurality (but not a majority) of seats has an effect only slightly larger than being the most voted (but tied in number of seats). This suggests that the importance of the novel effect we document relative to previously studied determinants of bargaining outcomes.

While these results can appear as *prima facie* irrational or puzzling, we present a simple model that can rationalize them. Elections in our model have two roles: information aggregation and incumbent disciplining. Elections aggregate disperse information about an uncertain state of the world. Thus, *after* an election, voters update beliefs about which party they prefer to appoint the mayor. However, parties' representation in a council is set at this point, and bargaining over mayoral appointments can be based on factors that ignore voters' interests. This creates a conflict between voters and parties and a role for disciplining. The model has multiple equilibria, which can be interpreted as different norms (self-enforcing rules of behaviour) that voters can adopt.[7] A norm that matches our results and where voters punish second most voted parties that appoint the mayor, constitutes an equilibrium (that also maximizes voters' expected welfare, suggesting an instrumental value for the norm).

We provide three pieces of evidence consistent with the model. First, we test whether voters' behaviour is consistent with punishing parties that deviate from the norm. We use a triple-difference strategy that leverages variation across time, whether a party was barely the first or second most voted (but tied in number of seats), and whether it appointed the mayor. Second most voted parties that appoint the mayor lose votes in the next election, compared to most voted parties that appoint the mayor. Second, the model predicts a specific pattern of heterogeneous effects: the effect of being most voted is stronger when the vote share of the third-placed party is larger. Some alternative explanations for our results (*e.g.* the status quo rule or the effect arising as an agreement among parties in repeated bargaining) do not naturally lead to these. Third, we provide evidence suggesting that the norm we study affects policymaking: Spanish municipalities that appoint most-voted mayors have fewer instances of government corruption and, in European countries, appointing the prime minister from the party with most seats is associated with shorter bargaining delays and longer government duration.

This article is related to four strands of the literature. First, as previously discussed, our result is difficult to reconcile with existing theories of bargaining and coalition formation and suggests the importance of a relatively unexplored factor in bargaining outcomes.

5. This is the case of the data from National European parliaments used in this article (the *European Representative Democracy Data Archive*). Gamson's Law is discussed further in Laver (1998) and Carroll and Cox (2007). Examples of structural estimation of government formation are Merlo (1997) and Diermeier *et al*. (2003).

6. We also document that programmatical affinities are, on their own, predictive of outcomes. In the overall sample, when the PSOE and IU hold a combined majority of seats, it is likely that one of the parties appoints the mayor.

7. Persson *et al*. (1997) also interpret different equilibria in voting models as different norms.

Second, the results are relevant for comparative politics and the design of electoral systems. Existing models suggest that our results have policy consequences. In particular, the results we document add first-past-the-post considerations to proportional systems. For example, Lizzeri and Persico (2001) compare how public good provision differs under different electoral systems. In their model, parties under proportional representation maximize vote shares (which translates proportionally to power), while parties under plurality rule maximize the probability of being most voted (a winner-takes-all contest). At first pass, the Spanish context fits the proportional representation case. However, our results suggest this characterization misses that parties should not simply maximize vote shares but also aim at being the most voted. In other words, our results suggest that incentives under proportional representation may be more similar to those under plurality rule than previously acknowledged, affecting the policy and welfare consequences of electoral rule design. Relatedly, some countries (*e.g.* France, Greece, Italy, and Portugal) award a seats premium to the most voted party in their proportional representation systems. The results we document can generate equivalent *de facto* premia even without such explicit rules.

Third, our results speak to models where voters see being a "winner" as having value in itself (Callander, 2007, Callander and Wilson, 2008, and Agranov *et al*., 2018) and to the (previously discussed) evidence that rank-based decision-making affects political outcomes. Fourth, our results speak to the literature on "nominal" versus "real" bargaining power.[8]

The next section provides examples from the political discourse in multiple contexts and its implications to broader phenomena. Section 3 provides the evidence from European national parliaments and Section 4 provides the evidence from Spanish local elections. Section 5 describes the theoretical framework and Section 6 provides tests of its additional predictions. Section 7 discusses alternative explanations. Section 8 concludes.

2. MOTIVATING EXAMPLES AND GENERAL ISSUES

This section discusses further examples of how the prescription that "the party with the most seats should form government" has been incorporated into the discourse of voters and politicians and its relation to a broader pattern where pluralities are reinterpreted as majorities (Maskin and Sen, 2016).

2.1. *Examples from New Zealand, the United Kingdom, Italy, and Spain*

Since a 1994 electoral reform, no party in New Zealand obtained the majority needed to form a government, but for seven consecutive elections, the most represented party formed the government. The first exception occurred in 2017, when the second-placed Labour party formed a government. This was perceived as due to a "maverick" third-placed party "kingmaker" and met with confusion by voters. As put by the governor-general, "the leader of the party with the largest number of seats in Parliament has always been able to form a government. While some voters think that will always be the case, it may not".[9]

Similar issues appear in British politics. During campaigning for the 2015 general election, the leader of the Liberal-Democrat Party (Nicholas Clegg) stated that "the party that gets the

8. See Warwick and Druckman (2001) and Frechette *et al*. (2005) and the discussion on Section 4.4.

9. The New Zealand example is also illustrative since, in the seven elections between 1994 and 2017, two parties (Labour and National) alternated in power. Government formation in New Zealand has an "*absence of formal procedures*". The description above and quotes are from Daly, 2017. The comment section in Daly, 2017 provides many examples of voters' opinions (*e.g.* "clearly the highest polling party has the largest mandate to negotiate with a potential coalition partner").

most votes and most seats, in other words the party that gets the biggest mandate from the British people, even if it does not get a slam-dunk majority, it seems to me right to give that party the space and the time to try and settle a government" (Perraudin, 2015) and Scottish Labour Party leader Jim Murphy stated that "the biggest party gets to form the government" (McKinney, 2015).

After the March 2018 Italian elections, the Five Star Movement was the most voted party and obtained the most seats in parliament, but fell short of a majority. The party argued that the most voted party appointing the prime minister would be "more democratic". As one of its candidates put it, "given that the Five Star Movement received almost a third of all votes and is by far the single most popular political force in Italy, any other choice would be undemocratic" (The Guardian, 2018).

Voters also agree with these statements. A nationally representative poll found that 55% of Spaniards agree that "it is more democratic that the most voted party forms the government, even if that party does not have an absolute majority of the votes" (El País, 2015).[10] Moreover, in multiple instances, leaders of both major national parties (PP and PSOE) have made campaign promises to not form government if their party was not the most voted in both national and local elections (Europa Press, 2007). These statements also apply to local politics: former PP leader and prime-minister Mariano Rajoy stated that "we have always supported that the mayor is the person who received the most votes. It follows from a common-sense democratic rule" (El País, 2018). Subsection 4.2 discusses a related case study from Olivenza's 2011 election.[11]

2.2. *Majorities versus pluralities*

An interesting common aspect in the four contexts discussed above (and the Canadian example in the introduction) is that they appeal to notions of what is "democratic" or the "people's choice", even though they relate to cases with a plurality, but not a majority, of votes (or seats). A puzzling aspect is that it could also be argued that a majority of voters did not vote for the party being labelled "the people's choice". Moreover, the Canadian case raises another interesting point: the sitting prime minister agrees that his incumbency status is less important than "winning the most seats" when it comes to appointing a prime minister.

These examples (as well as our empirical results) illustrate Maskin and Sen's (2016) argument on the confusion (or reinterpretation) of pluralities as majorities discussed in the introduction. A 2004 campaign speech by PSOE leader José Luis R. Zapatero illustrates this connection. Zapatero promises to try to form a government only if his party obtains a "*sufficient majority*". He clarifies that "a sufficient majority is being the most voted party in Spain, it is having more votes than the second place, because this would mean that most Spaniards approve our proposals" (El País, 2004). The quote conflates majority with plurality in a circular fashion and is striking since his party was then expected to be the "second place".[12]

2.3. *Example of a broader pattern: Donald Trump's nomination*

An example of how this issue applies to contexts beyond government formation is Trump's nomination as a presidential candidate. The Republican Party's rules state that a *majority* of

10. One caveat is that the question does not clarify what disagreeing implies. Given the "more democratic" wording and "even without an absolute majority" qualifier, a logical alternative is that disagreeing implies "it is equally democratic for a first, second, or third most voted party to appoint the form the government if none has a majority".

11. Similar pledges have been made at regional governments, for example in Andalusia (ABC, 2004).

12. The context of the speech was indeed a discussion on whether his party (PSOE) would attempt to form a government even if the PP was the most voted (as expected in the polls). Moreover, as a politician engaged in a discussion over government formation, it is unlikely that Zapatero is unaware of the distinction between majority and plurality.

delegates is needed to nominate a candidate. However, early in the 2016 primaries, many expected Trump to achieve a plurality, but not a majority, of delegate votes. Commentators (and Trump himself) declared that the candidate with the most delegate votes should be the nominee. Moreover, 62% of Republican voters agreed with the statement that with "no delegate majority, the GOP nominee should be the one with the most votes".[13] Moreover, Silver (2016) argued that a reason for Trump's eventual success at securing a majority was that "Republican voters were swayed by Trump's arguments that the candidate with the most votes and delegates should be the nominee". If Silver's argument is correct, a rule that was designed to be majoritarian was reinterpreted as voters as pluralitarian, along the lines of Maskin and Sen (2016).

3. EVIDENCE FROM EUROPEAN NATIONAL PARLIAMENTS

This section provides two results regarding government formation in European parliaments. First, most represented parties that have "just one more seat" than the second most represented are more likely to appoint prime ministers. Second, cases where the most represented party appoints the prime minister are associated with shorter delays in government formation and longer government durations.

Data and context. Our data cover 308 episodes of government formation following an election in twenty-eight European countries in the 1944–2010 period.[14] It contains the party affiliation of the appointed executive (prime minister) and the number of seats of each party in the lower house. In most cases, there is no formal institutional advantage for the party with the most seats in forming government.[15] It is possible that some countries have informal rules, such as the head of government asking the most represented party to be the first formateur (the party which makes a proposal put to a vote). To the extent these are not coded in laws, they can be understood as norms. Moreover, Diermeier and Merlo (2004) show that formateurs are not chosen by seat rank in a sample of European parliaments.[16]

3.1. *Effect of having the most seats on appointing prime ministers*

3.1.1. Empirical strategy. We document this result by leveraging a RDD in which the sample is restricted to (1) only include two parties with the most seats in the legislature and (2) exclude cases where one party has a majority of seats. Our final sample includes 504 parties from 252 elections.[17]

13. The other option in the survey was "GOP nominee should be the best party standard-bearer", which 33% of respondents agreed with. The survey occurred in April 2016 (Murray, 2016 and Flegenheimer, 2016).

14. The dataset is the The European Representative Democracy Data Archive (Andersson and Ersson, 2014) and includes Austria, Belgium, Bulgaria, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherland, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the U.K.

15. The only exception is Bulgaria, which in 1993 stipulated that the most voted party should be the first formateur. Repeating our analysis excluding Bulgaria leads to similar results (Supplementary Appendix Figure A1 and Table A1).

16. Diermeier and Merlo (2004) show that formateur selection is better explained by selection probabilities being proportional to seat shares than following seat ranks. Their result is not inconsistent with ours since they study formateur selection (and not PM appointments) and their maximum likelihood estimation aims at fitting all the data, while we focus on behavior near a RDD cutoff. While Diermeier and Merlo's (2004) sample includes only eleven countries, restricting our analysis to only the countries in their sample does not affect our results (Supplementary Appendix Figure A2 and Table A1).

17. The data contain three cases in which the two most represented parties tie in number of seats: the Netherlands in 1952, and Belgium and Estonia in 2003. In the Dutch and Belgian cases, the most voted party appointed the prime minister. In the Estonian, the second most voted party did. These cases are excluded from our estimating sample since a party must be labelled having the "most" and "second most" seats to enter equation (3.1).

Define s_{it} as the seat share of the party with the most seats minus the seat share of the party with the second most seats in country i at election year t. We define the running variable for party p as:

$$x_{pit} = \begin{cases} s_{it} & \text{if } p \text{ has the most seats} \\ -s_{it} & \text{if } p \text{ has the second most seats.} \end{cases}$$

Therefore, if $x_{pit} > 0$, then party p has the most seats (otherwise, has the second most seats). Our outcome of interest is y_{pit}, a dummy indicating whether party p appointed the relevant executive member for the entire term following the election at t. We refer to this outcome as "appointing the prime minister (PM)" throughout. The effect of having the most seats is given by $\lim_{x_{pit}\downarrow 0} \mathrm{E}[y_{pit}|x_{pct}] - \lim_{x_{pit}\uparrow 0} \mathrm{E}[y_{pit}|x_{pit}]$, which can be estimated by a local polynomial regression:

$$y_{pit} = \theta_0 + \theta_1 \cdot 1\{x_{pit} > 0\} + g_0(x_{pit}) + g_1(x_{pit}) \cdot 1\{x_{pit} > 0\} + \epsilon_{pit} \tag{3.1}$$

using observations within a given bandwidth around the threshold. g_0 and g_1 are polynomials estimated separately on each side of the cutoff. Thus, θ_1 captures the effect of having the most seats (instead of having the second most seats) in a parliament where the two parties with the most seats almost tied in seats. Baseline estimates use a linear specification ($g_0 = g_1 = x_{pit}$), as suggested by Lee and Lemieux (2010), and the Imbens and Kalyanaraman (2012) optimal bandwidth. We discuss the robustness to different bandwidths and polynomial orders. Standard errors are clustered at the country level.

There are three noteworthy aspects of this setup. First, variables that do not vary across parties within an country-year will be distributed symmetrically around the RDD threshold. For each election, both a first and second placed party enter the sample in symmetric fashion (one has $x_{pit} = a$ and the other $x_{pit} = -a$). The identifying variation thus comes from comparing parties within an election.[18] Second, conditioning which observations enter the sample by a variable that varies only at the election or country level will not affect the RDD's "internal validity" (balance in covariates around the cutoff). For example, restricting the sample to only the cases where a party does not have a majority of seats cannot create imbalance around the cutoff since "one party having a majority" varies at the election level. Third, while each election enters the sample twice, this "double-counting" of elections does not artificially affect standard errors clustered at the country level.

3.1.2. Main result. Figure 1a plots the probability that a party appoints the PM against the seat share difference between the two parties with the most seats (parties placed third and lower are excluded from the sample). The 504 observations are aggregated into bins of 1-p.p. width of the running variable (x_{pit}) and local averages for each bin are plotted. The solid lines are from a quadratic polynomial based on the original (unbinned) data and fitted separately on each side of the cutoff. The "jump" at the cutoff suggests that a party with "one more seat" is almost 40 p.p. more likely to appoint the PM. Moreover, the relationship between the outcome and running variable is relatively flat on the left of the cutoff. This suggests that as second-placed parties increase their number of seats (relative to the first-placed), they are not more likely to appoint the PM. However, the additional seat that "flips" a party into being the most represented has a sizable impact.

18. Figure A3 presents the distribution of observations, demonstrating this symmetry.

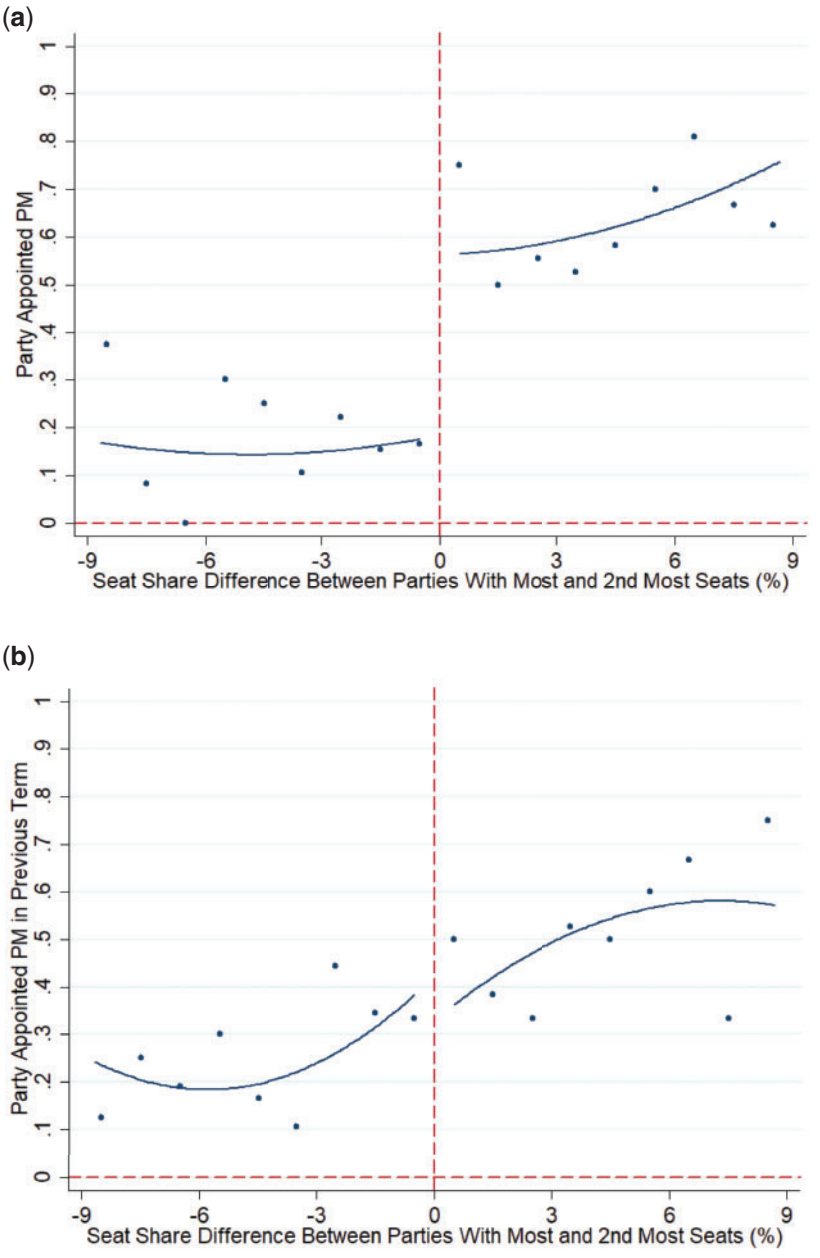

FIGURE 1

(a) Effect of having the most seats—national parliaments data; (b) Placebo test: "effect" of most seats on lagged outcome.

Notes: 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.

TABLE 1
National parliaments data: effect of having most seats on appointing Prime Minister

Dependent variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: main outcome					
Party appointed	0.202	0.303**	0.583**	0.387***	0.387**
Prime Minister		(0.130)	(0.210)	(0.122)	(0.145)
N		224	24	504	504
Panel B: lagged outcome (Placebo Test)					
Party appointed	0.414	−0.0316	0.167	0.0852	0.136
Prime Minister, $t-1$		(0.163)	(0.332)	(0.122)	(0.129)
N		152	24	504	504
Panel C: covariate balance (outcome predicted from party ideology)					
Party Appointed	0.417	−0.021	−0.035	−0.015	−0.009
PM (Predicted)		(0.028)	(0.078)	(0.026)	(0.033)
N		222	24	504	504
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Notes: 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 7.39%, 4.82%, and 7.36%, for the three dependent variables, respectively. See text for the construction of the outcome on Panel C.

Panel A of Table 1 shows the equivalent regression results. A party with the second most seats that almost ties with the one with most seats has a 20.2% chance of appointing the PM (the "2nd-Place Mean", which is the estimated θ_0 from equation (3.1)). Column (1) indicates that the party with most seats (but almost tied) is 30.3 p.p. more likely to appoint the PM (the estimated θ_1). Both figures are based on using only the 224 observations within the optimal bandwidth. Column (2) compares the average outcome for parties that are only 1% of seat share apart and finds an even larger effect. Both linear specifications are robust to the choice of bandwidth (Figure A4). Using the full sample and a quadratic or cubic polynomial yields similar results, shown in columns (3) and (4).

Interestingly, this effect is not driven by an increased probability of being in the ruling coalition (having cabinet positions). Figure A5 replicates the exercise of Figure 1a using a dummy equal to one if the party is in the ruling coalition as the outcome. We do not see a jump at the cutoff.[19]

3.1.3. Covariate balance. Figure 1b repeats the exact same exercise from Figure 1a, but with the lagged outcome on the *y*-axis. It thus plots whether the party appointed the PM in the *previous* term against their current seat share difference. The absence of a jump at the cutoff indicates that parties with close seat shares are equally likely to have appointed the incumbent PM. Panel B of Table 1 present the analogous regression results.[20]

3.1.4. Does one seat make a difference in the ability to form majorities? The effects presented so far can perhaps be explained by the party with most seats having a *numerical*

19. Unfortunately, our data does not contain information on the *share* of cabinet positions each party obtained.

20. Party identity is also balanced around the RDD cutoff. Andersson and Ersson (2014) classify the parties in our data into twelve ideological families (*e.g.* social-democratic, liberal). We regress our outcome on a set of dummies indicating each family. We use the predicted value from this regression as an outcome (Panel C of Table 1 and Figure A6). If a particular "family" was more likely to be on a particular side of the cutoff, we would observe a non-zero effect.

advantage in forming coalitions. For example, it is possible that the first placed party can form a majority with only the third-placed party, while the second placed cannot. To investigate if such differences drive our results, Figure A7 repeats the exercise of Figure 1a, but using a subsample of 254 observations that excludes all cases where the party with the most seats could form a majority coalition with the third most represented party, while the second most represented cannot. The effects are strikingly similar to those from the main sample. This is consistent with the effect of "having the most seats" in itself, as opposed to higher bargaining power associated with the ability to form different coalitions.

To further explore this issue, we calculated the Shapley–Shubik power index for each of the 504 parties in the sample. Figure A8 repeats the exercise of Figure 1a using the index as the outcome. Parties in the 1-p.p. bin just left of the cutoff have an average index of 0.284, while those just right have essentially the same average (0.293). This suggests that the previous results are not driven by the differences in bargaining power captured by the index. Moreover, Figure A9 repeats the exercise of Figure 1a using the Shapley–Shubik index as the running variable, and the size of the "jump at the cutoff" is similar to the one in Figure 1a. Panel C of Table A1 presents the corresponding estimates.[21]

Ideally, one could restrict the sample to cases where the top two parties have the same Shapley–Shubik index. However, this subsample would include only forty-six elections, of which only six would involve the two most represented parties being 1% of seat share apart. This highlights the value of studying Spanish municipalities, where there are 2,898 cases with the two most voted parties tying in number of seats (implying they have the exact same Shapley–Shubik index).

3.2. *Bargaining delay and government duration*

3.2.1. Delays in government formation. Delays in government formation can be costly, with previous literature associating delays in government formation with financial market volatility.[22] They are also frequent. In our sample, it takes on average 38 days for a government to be formed when no party has a majority (median delay is 33 days). In 19% of cases, the delay is longer than 60 days. Delays are longer when the most represented party does not have clear dominance. Excluding the cases where the most represented party can form a majority coalition with the third most represented party (while the second most represented cannot) increases the average delay in the sample to 50 days and the share of cases with delays longer than 60 days to 32%.

3.2.2. The party with most seats appointing the PM is associated with shorter delays. Panel A of Table 2 presents results from regressing the (log) number of days taken to form a government against a dummy indicating whether the party with the most seats appointed the

21. The index is described in Shapley and Shubik (1954). Intuitively, the index measures the fraction of sequences in which the party is pivotal in forming a coalition if parties entered in coalitions sequentially. It varies from zero (the party is not pivotal in any sequence) to one (the party has a majority). We obtain similar results replicating the analysis using the Banzhaf III (1964) index (which measures the fraction of possible majority coalitions in which the party would be pivotal) and the Coleman–Shapley index (which combines features of the two previous indexes and is described in Casajus and Huettner (2018)). The correlation between these indexes in our data is always above 0.99.

22. For example, it took 194 days for the representatives elected in the June 2007 Belgian election to form a government. In these six months, a government with caretaker status was "unable to take policy decisions" and a crisis ensued. A similar case happened after the 2006 Czech Republic election, which led to a seven-month period without a government and left "legislation and important reforms in a state of limbo". The quotes are from Golder (2010), which also surveys previous literature on the causes and consequences of bargaining delays in government formation. Figure A10 plots the cumulative distribution function of delays in government formation in our data for each of the subcases discussed.

TABLE 2
Appointment of PMs from the party with the most seats is associated with shorter delays in government formation and longer governments

	(1)	(2)	(3)	(4)
Panel A: dep. var. is log(days between election and government formation)				
Party with most seats appointed PM	−0.348**	−0.356**	−0.269*	−0.277*
	(0.173)	(0.150)	(0.160)	(0.165)
N	308	308	308	308
Panel B: dep. var. is log(length of government in days)				
Party with most seats appointed PM	0.177***	0.210***	0.183**	0.173**
	(0.0674)	(0.0683)	(0.0708)	(0.0722)
N	301	301	301	301
Country FE:		Y	Y	Y
"Legislature Class" FE:			Y	Y
Party ideological family FE:				Y

Notes: Robust standard errors in parentheses. The unit of observation is a country-year. Each figure in columns reports a separate regression of the specified dependent variable against a a dummy indicating whether the party with the most seats appointed the prime minister. Column (2) adds a set of country fixed effects. Column (3) adds a set of five dummies indicating each the classes of seat distributions in a legislature described by Laver and Benoit (2015), which capture the relative ability of parties to form coalitions (*e.g.* a dummy for whether one party has a majority and another for whether any two of the three most represented parties can form a coalition). Column (4) adds a set of 12 dummies indicating the ideological family of the most represented party (*e.g.* the party being liberal, communist, green). See text for further details.

PM. The result in column (1) indicates that, when the PM is from the most represented party, the time taken to form a government is almost 35% shorter. This result is robust to controlling for country fixed effects (column 2). One possible confounding factor is that cases where the most represented party is stronger (*e.g.* has a majority or a large plurality of seats) are associated with both shorter delays and the eventual appointment. To address this issue, column (3) controls for a set of dummies capturing Laver and Benoit's (2015) five classes of the implied bargaining power given the seat distribution. Finally, column (4) adds to this specification a set of dummies capturing the "ideological family" (*e.g.* conservative, liberal, green) of the most represented party in the sample. The coefficient remains negative and sizable.

The results suggest that the phenomenon we study may affect welfare. Another (more speculative) interpretation is that it arises exactly to shorten such delays. It is important to caveat, however, that results in Table 2 are based on correlations and, although robust to controls, it is difficult to rule out the possibility of other confounding factors.

3.2.3. Term lengths. Panel B of Table 2 repeats the exercise described above using the length of the government (number of days in power) as the outcome.[23] If PMs from the party with the most seats are perceived as a "more democratic" choice, they may have an advantage in staying in power too. This is what the results suggest, as PMs from parties with the most seats stay in power 18–21% longer. These results are also robust to the previously discussed set of controls, but the caveats regarding the result being based in correlations also apply.

4. EVIDENCE FROM SPANISH MUNICIPAL GOVERNMENTS

As discussed in Subsection 3.1, the results from European national parliaments have the drawbacks of a small sample size and the inability to fully control for potential differences in bargaining

23. We approximate term length by the days between the election that led to that government and the next election.

power associated with additional seats. Given Spain's numerous municipalities (and small council sizes), we can observe many elections where the two most voted parties tie in seats.[24]

4.1. *Context, data, and empirical strategy*

4.1.1. Electoral rule and government formation procedure. Spanish national law regulates how the formation of municipal governments. First, voters elect a municipal council in a general election. Second, the members of the council elect one of its own to be the mayor. General elections occur simultaneously in all municipalities every four years. Councils (*concejos*) are elected by proportional representation in single-district (at large) elections. Council size is a function of municipal population (Table A2). Each party presents a ranked list of candidates. In the election, each voter picks one of the party-lists. The conversion from party votes to seats follows D'Hondt rule.[25]

In their first meeting after the election, councilors choose a mayor (*alcalde*). Leaders of each party (the candidates ranked first in the lists) are eligible to be mayor. Each councilor votes for one of the leaders or abstain. If a leader has more than 50% of votes, she is appointed mayor. If no leader obtains a majority, a status-quo rule dictates that the leader of the most voted party in the general election is appointed mayor. Section 7 discusses why this status-quo rule is unlikely to drive our results.[26]

The mayor can be replaced at any moment in the term by a censure motion (*moción de censura*): a proposal to both remove the current mayor and appoint another councilor as mayor. This requires approval by a majority of the council. Only one censure motion can occur per term. Another possibility is a motion of no confidence (*cuestión de confianza*), which is proposed by the mayor in certain cases requiring approval of the council (*e.g.* budgets). The number of votes required for the mayor to lose the motion depends on the context (*e.g.* in a budget vote, the mayor loses if nays outnumber yeas). If a mayor loses the motion, the council elects a new mayor following rules used for electing the mayor for the first time. There are no "off-schedule" general elections: citizens only vote every four years.[27]

This combination of rules implies mayors not only need to obtain the support a majority to first get elected, but also must keep the support of that majority throughout the term, as it is

24. To the best of our knowledge, there is no other context with a comparably large number of ties in seats. Contexts where the executive is directly elected are not suitable for our analysis. Other countries with parliamentary local governments have a smaller number of municipalities, making it unlikely that ties are commonly observed (*e.g.* Finland has 311 municipalities and Sweden has 290, while our Spanish sample covers more than 5,900 municipalities).

25. This rule is also known as Jefferson rule and used to assign U.S. House districts to American states according do their population. It works as follows. The total votes cast for each party in the electoral district is divided, first by one, then by two, then three, up to the total number of seats to be allocated. Let p be the number of parties and s the number of seats. Then a grid of numbers with p rows and s columns is created, where the entry in the ith row and jth column is the number of votes won by the ith party, divided by j. The s winning entries are the s highest numbers in the whole grid (each party is given as many seats as there are winning entries in its row). Political parties must also obtain at least 5% of the votes to receive seats. This system refers to municipalities with more than 250 inhabitants. Other municipalities use an open-list system and are excluded from our sample. Sanz (2017) and Gonzalez-Eiras and Sanz (2018) estimate the effects of the change in electoral rules at the 250-person cutoff on turnout and women's representation, respectively.

26. Note, in particular, that we find effects of being second most voted versus third most voted, and there is no rule that awards an advantage to the second most voted party. If no candidate receives a majority and two or more parties obtained the exact same number of votes in the *general election*, then a lottery is run among the tied parties (in 1979 and 1983, such ties in *general election* cases were decided by appointing the oldest party leader). Exact ties in votes in the *general election* are very uncommon, and its few occurrences are deleted from our sample.

27. If a mayor resigns, is convicted of of a crime, or dies, a new mayor is chosen by the same procedure. When appointing a new mayor after a removal, the candidate from the party of the removed mayor is the next person in the party-list.

straightforward for a different majority to appoint a new mayor. In practice, in 97% of cases a mayor from the same party stays in office for the entire term. This number is 89% in the cases where the top two parties tie in number of seats, and 88% when they do so under a small (below 1 p.p.) vote share difference.

4.1.2. Importance of municipal governments. Municipal governments manage approximately 15% of the Spanish public expenditure (6% of GDP). Spanish law dictates which services must be provided by municipal governments. They collect taxes on residential properties, businesses, vehicles, as well as fees and user charges. Total tax and fee revenue by municipalites 4% of Spanish GDP. Hence, municipal-level policymaking is consequential to voters.[28]

4.1.3. The role of mayors. Mayors are the "the center of gravity of political life in the municipality" who "by law holds the most important executive functions and exercises leadership in municipal politics" (Vallés and Brugué, 2001). They chair council meetings and appoint cabinet members and staff. They have substantial control over the level allocation of expenditures (by preparing budgets). Spanish municipal governments exemplify a cases of strong executive power (Sweeting, 2009) and have been described as "municipal presidentialism" (Magre-Ferran and Bertrana-Horta, 2005).

4.1.4. Data. Our sample comprises all municipal elections in Spain since restoration of democracy. Elections have occurred in four-year intervals since 1979. The source is the *Ministerio del Interior* (Ministry of Internal Affairs). Our sample is based on the councils elected in the 1983–2011 elections. The sample covers 37,122 elections from 5,993 different municipalities. In all, 2,898 elections have the first and second most voted parties tied in number of seats.[29]

We observe the party affiliation of mayors. Unfortunately, we do not observe her supporting coalition. Neither the identity of the members or parties that voted for a given mayor, or the vote count of the election for mayor within the council, are recorded by the Ministry. Information about the allocation of cabinet positions within municipalities is also unavailable.[30] Given that mayors may not necessarily serve the entire four-year term, we define our main outcome as a dummy equal to one if the mayor that spent three quarters of the term in power during the term belongs to that party.[31]

4.1.5. Characteristics of elections close to the cutoff. There are 438 elections where the two most voted parties tie in seats and their vote share difference is below 1 p.p.. In these cases, 90% have councils such that a majority requires support from any two of the three most

28. Bagues and Campa (2017) describe the role of municipal governments in further detail. All municipal governments must provide lighting, graveyards, refuse collection, street cleaning, and water and sewer. Larger municipalities must provide social services and education. Some small municipalities choose to provide other services (*e.g.* childcare).

29. The 2015 election is not included since the term is still in progress (and mayoral appointments may change). The 1979 election is not included since we use lagged values in placebo tests. Including the 1979 and 2015 data does not affect our results (Table A4).

30. We do observe the party affiliation of deputy mayors, as described in Subsection 4.2.

31. We condition on spending at least three quarters of the term to exclude cases in which two parties share the term in two halves (which occur rarely) and so that our definition is not sensitive to cases where mayors spend a very short (weeks) in office. However, the effects are similar if we define the outcome as being mayor for more than any other party, being the first mayor to be appointed, or serving the full term (89% of cases).

voted parties. This includes both cases where only three parties received representation and cases where the fourth placed party cannot be pivotal for a majority (*e.g.* a 11-seat legislature with a 4-4-2-1 vote division). Therefore, the vast majority of the councils in our sample can be thought of as essentially three-party councils. The coalition formation game that approximates this context is thus one where any two out of three players can form a coalition that allocates payoffs: a "three-player majority game".

In these 438 cases, councils are relatively small (79% have 13 or fewer legislators) and have the first and second most voted parties "almost tying", on average, with 36.5% of votes and 39% of seats, with the third most voted obtaining vote (seat) share of 17.7% (16.7%).

4.1.6. Empirical strategy. We implement a RDD in a sample is restricted to (1) only cases where the first and second most voted parties have the exact same number of seats, and (2) only include the first and second most voted party. This sample has 5,796 observations (from 2,898 elections). The implementation follows closely the one described in Subsection 3.1. Specifically, we estimate equation (3.1), but now the running variable x_{pit} for party p in municipality i at time t is:

$$x_{pit} = \begin{cases} v_{it} & \text{if } p \text{ is the most voted} \\ -v_{it} & \text{if } p \text{ is second most voted} \end{cases}$$

where v_{it} is the vote share difference between the first and second most voted parties. Thus, if $x_{pit} > 0$, party p has the most votes. The same three points discussed in Subsection 3.1 on the symmetry of variables that are constant across parties within an municipality-year, conditioning the sample on such variables, and clustering standard errors at the municipal level apply here.[32]

4.2. *Effects of being most voted*

Figure 2a plots the probability that a party appoints the mayor against the vote share difference between the first and second most voted parties (parties placed third and lower are excluded from the sample). The 5,796 observations are aggregated into bins of 1 p.p. width of the running variable and the local averages for each bin are plotted. The solid lines are from a quadratic polynomial based on the original (unbinned) data and fitted separately on each side of the cutoff. The jump at the cutoff is the graphical counterpart to θ_1. Figure 2a indicates that the second most voted party appoints the mayor 33.6% of the time, while the first most voted party does so 53.9% of the time.[33]

A perhaps surprising pattern that an upward slope is not observed in Figure 2a. Conditional on the vote ranks, higher vote margins for the most voted parties are not associated with increased odds of appointing the mayor. Such slopes must be interpreted with caution, as there are both compositional effects and omitted variables that can affect them.[34]

32. Figure A11 presents the distribution of observations, demonstrating the symmetric distribution.

33. In elections where the vote share difference is below 1 p.p., the third-placed party appoints the mayor 3.7% of the time, and the fourth and fifth placed do so 0.5% of the time each. In the remaining 8% of cases, no party appointed a mayor that served for at least three quarters of the four-year term.

34. Such compositional effects and "omitted variables" are constant when comparing cases around the cutoff. A possible compositional effect occurs since smaller councils are less likely to be further away from the cutoff (*e.g.* no two parties with a 8 p.p. vote share difference can have the same number of seats in a 21-member council, but that is possible in a 7-member council). However, Figure A12 replicates Figure 2a for specific council sizes and indicates the lack of slope is not driven by such compositional effect. Figure A13 replicates Figure 1a for specific council seat compositions.

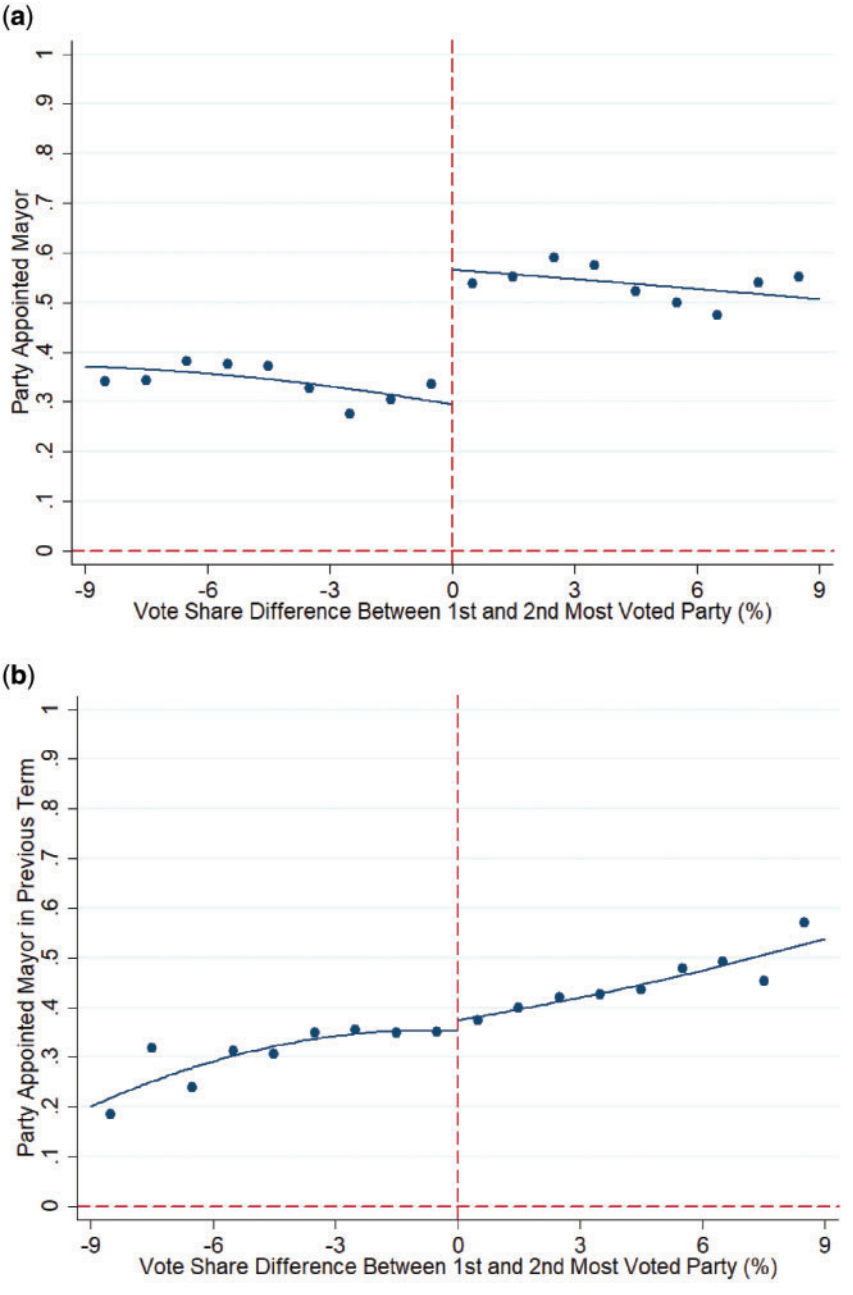

FIGURE 2

(a) Effect of being first (instead of second) most voted; (b) Placebo test: "effect" of most voted on lagged outcome.

Notes: 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 (Panel A) or if the party appointed the mayor in the previous ($t-1$) term (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.

TABLE 3
Effect of being first (instead of second) most voted

Dependent variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: main outcome					
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: lagged outcome (Placebo Test)					
Party appointed	0.358	0.011	0.023	−0.015	0.014
Mayor, $t-1$		(0.046)	(0.040)	(0.034)	(0.043)
N		2714	876	5796	5796
Panel C: covariate balance					
Party is PP	0.310	0.006	0.009	−0.027	−0.036
		(0.041)	(0.039)	(0.034)	(0.042)
N		3088	876	5796	5796
Party is PSOE	0.405	0.0003	0.005	0.018	0.001
		(0.045)	(0.044)	(0.036)	(0.046)
N		3222	876	5796	5796
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Notes: 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 2.32%, 3.19%, 3.75%, and 3.92% for the four dependent variables, respectively.

Panel A of Table 3 shows the equivalent regression results. A second most voted party that almost ties in votes with the most voted (but has the same number of seats) has a 35% chance of appointing the mayor (the "2nd-Place Mean", which is the estimated θ_0 from equation (3.1)). Column (1) indicates that the most voted that almost tied in seats is 19 p.p. more likely to appoint the mayor (θ_1). Both figures are based on using only the 2028 observations from elections where the top-two parties are only 2.32% of the total votes apart (the optimal bandwidth). Column (2) compares the average outcome for the first and second most voted parties that are only 1% of total votes apart and finds a similar effect. Even focusing on the forty-six observations from close cases (bandwidth below 0.1%), the estimated effect is 0.522 (SE = 0.167). Results are robust to the choice of bandwidth (Figure A14). Using the full sample and a quadratic or cubic polynomial yields similar results, shown in columns (3) and (4).[35]

4.2.1. Interpretation. As discussed in the introduction, comparing the two most voted parties that obtained almost the same number of seats identifies the effect of being most voted. When comparing two groups that should have the same bargaining power, differences in outcomes isolate the effect of the most voted label. Note that an effect size of 20 p.p. is both consistent with 20% of the municipalities having a effect equal one and with all municipalities in the sample

35. Table A3 replicates Panel A of Table 3 when defining "appointing the mayor" in different manners. Panel E of Table A3 restricts the sample to the cases where two of the top three most voted parties are needed to form a majority. Table A4 repeats this analysis incorporating the data for 1979 and (when possible) 2015. In all cases, the effects are similar in magnitude to those in Panel A of Table 3.

having an effect that only binds with 20% probability (or a combination of these two extreme cases).

4.2.2. Covariate balance. Figure 2b repeats the exercise from Figure 2a, but with the lagged outcome on the *y*-axis. It plots whether the party appointed the mayor in the *previous* term against their current vote share difference. The absence of a jump at the cutoff indicates that close first and second most voted parties are equally likely to be the incumbent mayor. Figure A15 repeats this exercise for party identity, showing that neither of the two main national parties, the PSOE and the PP, are more likely to finish in first place in a close election. Panels B and C in Table 3 present the analogous regression results. The point estimates are close to zero and statistically insignificant. These results suggest that incumbent mayors or the major parties are unable to manipulate election results. As previously discussed, for any variable that does not vary across parties within an election, there is perfect balance by construction and there cannot be "bunching" of municipalities around the cutoff.[36]

4.2.3. Is the effect present when the second and third most voted are aligned? We now focus on cases where the right-wing PP and the left-wing PSOE are the most voted parties, and the left-wing IU is the third most voted. The left-wing parties (PSOE and IU) thus have a combined majority are able to appoint one of their leaders as mayor, regardless of whether the PP is the most voted.

The red triangles on Figure 3 repeat the exercise of Figure 1a, but restricting the sample to only cases where the observation regards the PSOE in an election where the PP is the other top two most voted party and the IU is the third most voted party. Hence, the jump at the cutoff indicates that, when the PSOE is barely the second most voted, it appoints the mayor 55% of the time. If the PSOE is the most voted, it appoints the mayor 80% of the time. Similarly, the blue circles indicate that when the PP is the second most voted by a close margin, it appoints the mayor only approximately 10% of the time, however, when it the most voted, it appoints the mayor almost 35% of the time. Table A5 provides the corresponding table.[37]

The surprising aspect is that the PSOE is, overall, much more likely to appoint the mayor than the PP when the IU is the third place (red triangles well above blue circles in Figure 3). However, it is still the case that a small difference in vote shares that awards the PP the "most voted label" is enough to substantially increase the chance it appoints the mayor. In other words, the "most voted party appoints the mayor" prescription "bites" even in the cases when aligned second and third most voted parties can form a coalition to appoint a mayor of their own.

4.2.4. An example. The results of Olivenza's 2011 election can illustrate our results. The PP obtained 2912 votes and 7 seats; the PSOE, 2886 votes and 7 seats; and the IU 1376 votes and 3 seats. Given the 17-member council, the PSOE and IU could appoint one of their leaders as mayors. However, the PP appointed the mayor. The IU leader justified their decision of not supporting the PSOE to the media by stating it needed to accept "the decision of the people" and "what democracy has said", even though "it hurts me" that we will have a government "from the right" (Europa Press, 2011). The surprising aspect is that the "decision of the people" is based

36. See Figure A11 which shows how many observations are in each bin of Figure 2.

37. The estimates on Table A5 are obtained by estimating equation (3.1) restricting the sample to cases there the party p is the PSOE (Panel A) or PP (Panel B) and, in both cases, the IU is the third place and the PP and PSOE are the two most voted parties. The subsamples are not defined by which party is most voted and thus produce a balanced RDD.

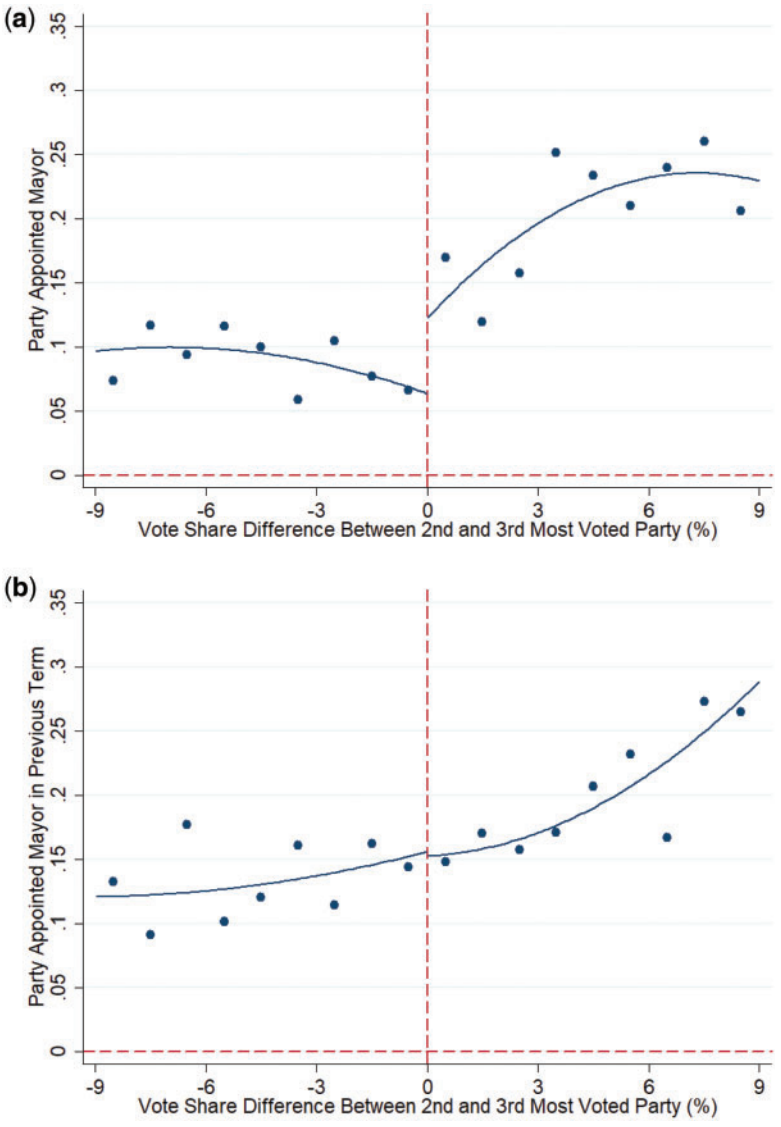

FIGURE 3

(a): Effect of being second (instead of third) most voted; (b): Placebo test: "effect" of second most voted on lagged outcome.

Notes: 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 party appoints the mayor (Panel A) or if the party appointed the mayor in the previous ($t-1$) term (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.

TABLE 4
Effect of being second (instead of third) most voted

Dependent variable	3rd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: main outcome					
Party appointed	0.067	0.092**	0.103***	0.059**	0.073**
Mayor		(0.043)	(0.028)	(0.028)	(0.036)
N		888	542	3132	3132
Panel B: lagged outcome (Placebo test)					
Party appointed	0.149	0.004	0.004	−0.024	0.006
Mayor, $t-1$		(0.037)	(0.034)	(0.034)	(0.045)
N		1844	542	3132	3132
Panel C: covariate balance					
Party is PP	0.312	−0.092	−0.044	−0.025	−0.033
		(0.072)	(0.043)	(0.043)	(0.056)
N		856	542	3132	3132
Party is PSOE	0.285	−0.028	−0.040	0.0043	−0.031
		(0.057)	(0.043)	(0.043)	(0.056)
N		1234	542	3132	3132
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Notes: Standard errors clustered at the municipality level in parentheses. The unit of observation is a party-municipality-year. The 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. 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. 3rd-Place Mean is the estimated value of the dependent variable for a 3rd most voted party that tied with the 2nd most voted party (using the specification in column 1). Optimal bandwidths are based on Imbens and Kalyanaraman (2012), 1.69%, 3.96%, 1.63%, and 2.41% for the four dependent variables, respectively.

on only 26 votes out of more than 7,000. Moreover, in a proportional representation system that does not formally reward the most voted, a majority of voters preferred the left-wing parties, but being most voted seemed to matter beyond that.

Supplementary Appendix A discusses three additional sets of results. First, it discusses why the possible interpretation that mayoral appoints are symbolic is unlikely and provides results on deputy mayor appointments. Second, it shows that the results are likely driven both by coalitions where the most voted party forms a coalition with the second- and third-most voted. Third, it shows that the effect is not stronger for "powerful" parties that are in power at the national or regional level.

4.3. *Effect of being second (instead of third) most voted*

The exercise of the previous subsection can also be applied to estimate if being labelled the second most voted, instead of third most voted, also has an effect on the probability of appointing a mayor. To do so, we redefine the sample such that (1) it only includes elections where the *second* and *third* most voted party obtained the *same number of seats* and the most voted party did not obtain a majority of seats and (2) only includes the second and third most voted parties.[38]

Figure 3a provides a graphical exercise similar to Figure 2a but, since it is based on this newly defined sample, all parties left of the cutoff are the third most voted, while parties to the right are the second most voted. A discontinuity at the cutoff is visible, although it is smaller and noisier than in Figure 2a. No discontinuity is visible in the placebo graph (Figure 3b) which plots lagged outcomes, increasing confidence that the jump observed in Figure 3a is not driven by noise.

38. Our sample contains 1,565 elections (from 1,204 unique municipalities) satisfying these two conditions.

Table 4 provides the regression results. Column (1) indicates that a third most voted party that almost ties with the second most voted has a 6.7% probability of appointing the mayor. That probability is almost 16% for the party finishing in second-place in such an "almost tie". Columns (2)–(4) indicate that effects of similar magnitude are estimated using different specifications.[39] This result suggests that the effect generalizes to lower ranks. Since there is no status quo rule (or any other formal differential treatment) benefiting second versus third placed parties and thus it cannot play a role in the effects on Figure 3a and Table 4. Moreover, it suggests the similar effects of being most voted are also not driven by the status quo rule. Section 7 further discusses this issue.

Finally, a similarly defined third (versus fourth) most voted effect and found effects is (statistically) close to zero. This is perhaps expected, since such lower ranked parties rarely appoint the mayor.[40]

4.4. *Comparison to effects of one additional seat*

While, given the discussion above, any non-zero effect of "being most voted" is perhaps surprising, it is also useful to gauge the magnitude of our effects to that of being awarded one additional seat. A similar RDD approach can be used to estimate the effect of one additional seat on the probability of appointing the mayor. So far, we have restricted our sample to cases where the first and second most voted parties have the same number of seats. However, there are also cases where the first and second most voted almost tie in votes, but the most voted is awarded one more seat. Whether an additional vote leads to one additional seat is defined by the rounding in D'Hondt rule.[41] Hence, we can estimate the effect of receiving one additional seat by restricting the sample to cases where the first and second most voted parties have different number of seats.[42] It is useful to separate the cases where the most voted party obtains one more seat in three different categories.

1. The additional seat creates more *nominal*, but no more *real* bargaining power. An example is a 5-4-2 seat distribution in an 11-seat council. While on party has more seats than others, its ability to form coalitions is not larger (any two parties can form a majority). The Shapley–Shubik or Banzhaf index is the same for both top two parties in this case.
2. The additional seat can generate both more *nominal* and *real* bargaining power. An example is a 5-4-1-1 seat distribution in an 11-seat council. The most voted party only needs one other party to form majority, the second most voted may need two. The Shapley–Shubik or Banzhaf index in this case differs for the top two parties.
3. The additional seat awards a majority of seats to the most voted party (*e.g.* a 6-5 seat distribution in an 11-seat council). In this case, the Shapley–Shubik or Banzhaf index is one for the most voted (and zero for the second most voted).

39. Panels B and C and Figure A16 show the covariate balance and Figure A11 shows the number of observations in each bin of Figure 3a. Figure A14 presents the robustness of the estimates to bandwidth choice.

40. Figure A17 presents a graphical analysis similar to Figures 1a and 2a, redefining the sample accordingly. Note the *y*-axis range matches that of Figure 2a for comparison. This sample covers 996 elections in which the third and fourth most-voted parties tied in seats, and the third (fourth) appointed the mayor in only 25 (11) of those.

41. For example, a 11-seat council where party vote shares are 42.0%, 41.9%, and 16.1% will have a 5-5-1 seat allocation. If the vote shares are the slightly different 41.0%, 40.9%, and 18.1%, the seat allocation is 5-4-2.

42. Note that we condition the sample definition to a variable defined at the municipality-year level and thus maintains the RDD "internal validity" as discussed in Subsections 3.1 and 4.1.

These three cases are directly observable and we can thus define three separate subsamples accordingly.[43] Figure 5 provides the regression discontinuity plot for these different subsamples. The blue circles correspond to the case where the two most voted parties tie in number of seats and are thus exactly the same as in Figure 2a. The red triangles plot the cases where the most voted party has one more seat than the second most voted party, *but no additional bargaining power* (case 1). The effect of just being labelled the most voted is similar in magnitude to obtaining one additional seat. The comparison between the two cases is perhaps surprising. It suggests that the effect of "being most voted" is of similar magnitude as "being most voted and obtaining an extra seat". The corresponding estimates are provided on Panel B of Table A6, which show that the estimated effect is larger, but not substantially so. Based on our baseline specification (column 1), the effect of being labelled the most voted is 60% the size of the effect of being labelled the most voted and obtaining an extra seat.

Similarly, one additional seat that is associated with more real bargaining power does provide substantially more ability to appoint mayors. These are depicted in green squares, representing case 2 above. Panel C of Table A6 provide the corresponding estimates. The specification on column (1) indicates that the effect of simply being labelled the most voted is 28% the size of the effect of the being labelled the most voted and obtaining an additional seat that awards real bargaining power. Finally, receiving a majority of seats (brown diamonds) makes a party almost surely appoint the mayor. Panel D of Table A6 provides the corresponding estimates.

These sets of results suggest that the effect of rank labels is comparable to the effects of additional seats, which have been the previous focus of the literature on legislative bargaining. Previous work (*e.g.* Warwick and Druckman (2001), Frechette *et al.* (2005)) has noted that nominal bargaining affecting outcomes conditional on real bargaining constitutes a puzzle. Our results suggest an additional mechanism that may help explain the puzzle.[44]

5. THEORETICAL FRAMEWORK

This section presents a model to illustrate a specific mechanism that can drive our results and yields predictions that guide the subsequent empirical analyses. Our starting point is a framework of political accountability (Barro, 1973; Ferejohn, 1986; Persson *et al.*, 1997). We add to it legislative bargaining and a role for elections in aggregating diffuse information. *After* an election, voters can infer information about an uncertain state of the world from vote shares. This informs voters on which party they prefer to appoint a mayor. However, parties' representation is set at this point and they may bargain and form coalitions based on factors that ignore voters' interests (*e.g.* rent allocations). This creates a conflict of interest between voters and parties and a reason for disciplining. The model has multiple equilibria, which can be interpreted as different norms that voters may adopt.

The model illustrates how a norm that might appear, *prima facie*, to be irrational behaviour can be sustained in equilibrium by a set of rational agents with standard preferences. An alternative modelling strategy would be to assume voters inherently prefer a mayor from the most-voted party and will punish lower ranked parties that appoint mayors (or reward most-voted mayors).

43. Of the 37,122 elections in our sample, 7.8% have the two most voted tying in seats (the focus of Subsection 4.2), and 7.9%, 8.6%, and 75.7% in cases (1), (2) and (3), respectively.

44. A caveat when comparing the effects in different panels of Table A6 is that municipalities that enter each subsample may differ systematically. For example, elections that constitute "case 3" are likely to have fourth placed parties obtaining larger vote shares than those that constitute "case 2". However, we note that the difference between one additional vote leading to one additional seat is mostly driven idiosyncratic rounding in D'Hondt rule, as exemplified in footnote 41.

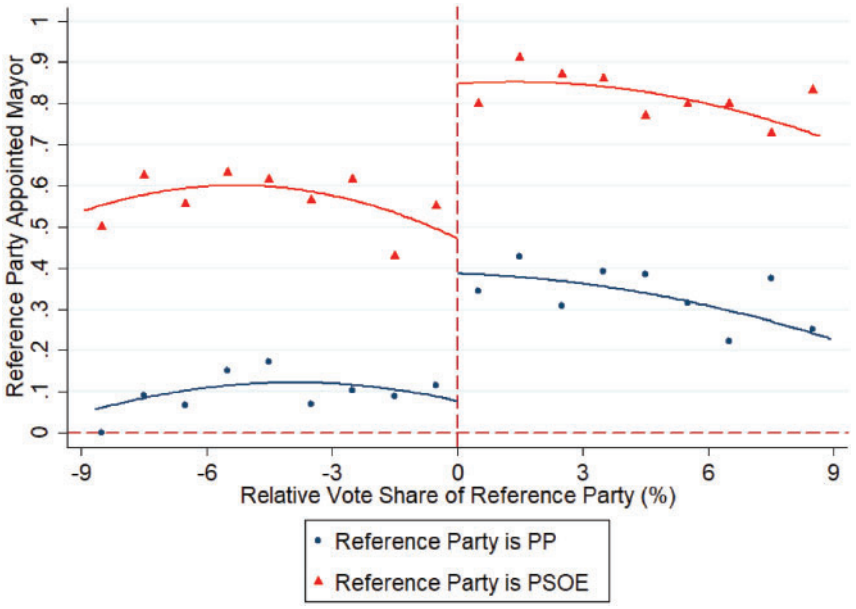

FIGURE 4

Effect of being most voted: cases with left-wing majority.

Notes: The unit of observation is a party-municipality-year. Sample is restricted to elections in which the *Partido Socialista Obrero Español (PSOE)* and the *Partido Popular (PP)* are the two most voted parties and the third most voted party is the *Izquierda Unida (IU)*. The running variable (horizontal axis) is the difference in vote shares between the two most voted parties, taking either the PSOE or the PP as the reference party. Hence red triangles (blue circles) to the left of the vertical line at zero are cases where the PSOE (PP) was the second most voted party and, to the right, the 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.

Doing so will naturally lead to outcomes similar to the equilibrium described in Proposition 2. Such a model could be interpreted as formalizing a norm as a principle of correct action that is internalized in *preferences*. In our model, norms are interpreted as equilibrium outcomes (self-enforcing mutual beliefs).

5.1. *Setup*

A large (odd) number of identical and infinitely lived voters maximize $E\sum_{t=0}^{\infty}\delta^t u_t$, where $0<\delta<1$, E is the expectations operator, and u_t is their utility. Every period, one state s_t of the world is realized. There are three possible states: $s_t \in \{A,B,C\}$. There are also three types of parties (A, B, and C), of which one must appoint the mayor. Voters receive positive utility if the mayor's type matches the state of the world. $u_t=1$ if $m_t=s_t$ and $u_t=0$ if $m_t\neq s_t$, where m_t denotes the party of the mayor. This can be interpreted as different events occurring, each of them being better dealt with by a given mayor, or only one party in each period having an honest or competent leader.

Each party also maximizes an expected utility function, $E\sum_{t=0}^{\infty}\delta^t x_t$, where x_t denote the rents they obtain from office: $x_t=1$ if the party appoints the mayor, and zero otherwise (rents are indivisible).[45] The bargaining procedure is as follows. If one party received a majority of votes in

45. This can be understood as the mayor setting a specific policy to the preferences of the party, or mayors not being able to commit to share the spoils of office with its supporting coalition. Supplementary Appendix C discusses how this can be relaxed.

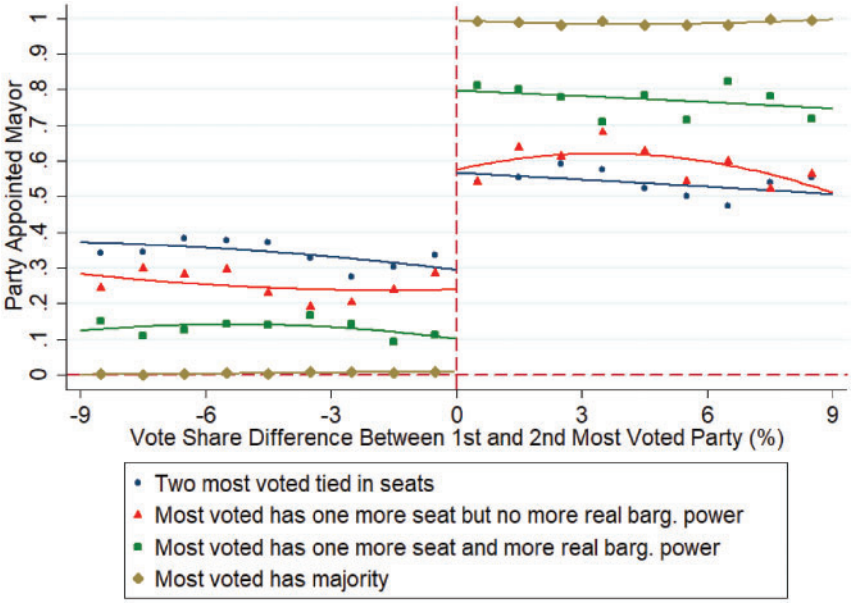

FIGURE 5

Effect of first place by legislature type.

Notes: The unit of observation is a party-municipality-year. Sample is restricted to the two most voted parties. Each plot restricts the sample to a different case of seat composition in the legislature. 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.

the previous election, it can single-handedly choose the mayor. If no party had a majority of votes, then one party is randomly "recognized" (selected to propose which party appoints the mayor). All parties then vote on whether to accept or not this proposal. If one of the two non-recognized parties accepts, the mayoral appointment is realized. If not, party A appoints the mayor. This procedure matches the one round of voting by majority rule feature of Spanish municipalities.[46] The choice of party A as the status quo is without loss of generality and made to illustrate how status-quo rules play no role in our argument. Parties' recognition probabilities are a continuous function of vote shares.[47]

Parties' preferences and the bargaining procedure are thus independent of voters' welfare and the states of the world. This creates a dissonance between voters' and parties' interests. Voters prefer the mayor that matches the state, but the choice of mayor may be determined by unrelated factors. If states were observable, voters would award a majority to the party matching the state.

5.2. *Uncertainty and information structure*

However, states of the world are not directly observed by voters or parties. The probability that state s occurs in period t is p_t^s, with $\boldsymbol{p}_t$ denoting the vector $[p_t^A, p_t^B, p_t^C]$. Moreover, voters

46. It is analogous to parties having votes weighted by their previous election's votes and deciding by majority rule.

47. For example, recognition probabilities being 1/3 for all parties or the same as the vote share in the previous election satisfy this condition. Continuity rules out the most voted party being recognized with certainty, which would make the model unattractive to study our empirical results.

and parties face uncertainty about the vector $\boldsymbol{p}_t$, which is drawn every period from a (common knowledge) distribution $G(\boldsymbol{p}_t)$ that is independent and identically distributed over time. Each voter individually observes a signal σ_t every period. The three possible signals are also $\{A, B, C\}$, with the probability the signal is s_t given by p_t^s, drawn independently for each voter.[48]

Hence, each period a voter updates her beliefs about the state of the world twice. At the start of the period, all voters have the same priors based on the expected value of $G(\boldsymbol{p}_t)$. After she observes her private signal of value i, she forms new beliefs $\Pr(s_t = k|\sigma_t = i)$ for all $k \in \{A, B, C\}$, which informs her vote. After observing the election results, she updates again, based on other voters' strategies. If all citizens vote according to their signals (*e.g.* vote for A if signal is A), she will expect the probability that the state is s to be the vote share of party of type s. However, by this time parties' representation in the legislature is already defined and, when another election occurs, a new vector $\boldsymbol{p}_{t+1}$ and state s_{t+1} will be drawn, making the previous information irrelevant.

We assume that $G(\boldsymbol{p}_t)$ is such that $\Pr(s_t = i|\sigma_t = i) > \Pr(s_t = j|\sigma_t = i)$ for all $i \neq j$. This implies that, after a voter observes a private signal of value i (but before observing election results) she expects i to be the most likely state and prefers party of type i to appoint the mayor. Supplementary Appendix B provides an example of a $G(\boldsymbol{p}_t)$ function and illustrates how voters update in the model.

5.3. *Timing and elections*

The sequence of events is the following. At the start of every period t, nature draws the vector $\boldsymbol{p}_t$. Based on this vector it draws the state of the world and individual signals. Citizens then cast their votes. There are six possible votes: voting for one of the *i*ncumbent parties A^I, B^I, or C^I that were in office in the preceding period or voting for one of *ch*allenger parties A^{Ch}, B^{Ch}, or C^{Ch}. In other words, for each of the three party types, there is always a challenger party of the same type that is identical in all respects to the incumbent. An incumbent that receives zero votes is never re-elected again. Parties then appoint the mayor according to the procedure described above. Payoffs are realized and a new identical period starts over.

5.4. *Discussion of assumptions*

This setup captures a dual role for elections. They can serve as an information aggregation mechanism and as a way to discipline incumbents to behave in consonance with voters' interests. The assumption of a challenger party of each type makes this dual role clearer. While alternative assumptions that would lead the choice to punish one incumbent also reward the other incumbent parties could be more realistic, they would complicate the model and create a conflict between the information aggregation and disciplining roles of elections. Note, however, that many parties, with presence at the national, regional, and municipal level, operate in Spain. If the types of parties are interpreted as ideologies (*e.g.* left, center, right), this can be interpreted as multiple leftist parties (some regional or municipal) that can replace one another.[49] While the model assumes a specific bargaining procedure, Supplementary Appendix B discusses how the results are robust to allowing multiple rounds of bargaining and rents to be divisible across parties.

48. All realizations of $G(\boldsymbol{p})$ satisfy $p^A + p^B + p^C = 1$.

49. We restrict our attention to equilibria where all voters receiving the same signal cast the same vote and where two parties of the same type will not receive votes in an election. We hence abstract from the possibility of two parties of the same type being represented to keep the exposition concise.

5.5. *Equilibria*

We restrict our attention to sequentially rational equilibria in which every voter chooses a pure strategy that conditions her decision on her last observed signal and the incumbent party's behavior in the preceding period. All parties choose pure strategies that condition only on the result of that period's election.

This model has multiple equilibria, similarly to Barro (1973), Ferejohn (1986), and Persson *et al.* (1997). Since an incumbent party is identical to a challenger of the same type, voters find choosing either an incumbent or challenger of the same type *ex post* optimal. Different equilibria where voters condition their choices on incumbent behaviour or not, or condition in different ways, can be interpreted as different norms voters can adopt. Since they are equilibria, they are also self-enforcing (when others follows the norm, each individual also finds it optimal to do so). This interpretation of multiple equilibria as different norms in is discussed in Persson *et al.* (1997). We do not fully characterize the equilibria in this model, but focus on two cases: one equilibrium with the "most voted appoints the mayor" norm and one without it. We begin with the latter.

Proposition 1 *There exists an equilibrium where, every period, a citizen observing signal $\sigma_t = i$ votes for the incumbent party of type i. A party that obtains a majority of votes appoints the mayor. If no party obtains a majority, then each party, if recognized, makes an offer to appoint the mayor itself. All parties, if not recognized, accept any proposal.*

Supplementary Appendix B presents the proofs. In this equilibrium, if no party receives a majority of votes, each party has a chance of appointing the mayor equal to their recognition probability—which must be the same for two parties that tied in votes. Hence, this equilibrium does *not* generate a jump in the RDD studied in Section 4. Those results, however, can be captured by the following equilibrium.

Proposition 2 *If $G(\boldsymbol{p}_t)$ is such that three conditions are satisfied: (i) $\Pr[p_t^A > \max(p_t^B, p_t^C)] > 1-\delta$; (ii) $\Pr[p_t^B > \max(p_t^A, p_t^C)] > 1-\delta$, and (iii) $\Pr[p_t^C > \max(p_t^A, p_t^B)] > 1-\delta$, then there exists an equilibrium where, every period, a citizen observing signal $\sigma_t = i$ votes for the challenger of type i if, in the previous period, i both appointed the mayor and was not the most voted party. If, in the previous period, i did not appoint the mayor or did so after being the most voted, a citizen observing signal $\sigma_t = i$ votes for the incumbent of type i. All parties, if recognized, propose that the most voted party appoints the mayor. The most voted party accepts a proposal in which it appoints the mayor, but rejects all other proposals. The second (third) most voted party rejects a proposal in which it appoints the mayor, but accepts all other proposals.*

The intuition behind Proposition 2 is that a second or third most voted that is recognized compares the utility of appointing the mayor for one period and never being reelected again with the continuation value of being reelected. The latter is the perpetuity of the probability of being the most voted party, which conditions (i)–(iii) guarantee is smaller than the one-period gain from deviating from the norm.

5.6. *Interpretation of the norm*

The equilibrium in Proposition 2 generates the RDD jump described in Section 4. Even when the two most voted parties are one vote apart, the most voted appoints the mayor. This occurs even though voters (rationally) understand that the difference in expected welfare between appointing

the first and second most voted is close to zero. This highlights the interpretation of equilibria as norms: the behaviour of voters and parties is mutually self-enforcing.

While the "most voted appoints the mayor" norm is associated with an equilibrium where agents strategically play best responses, it can also be interpreted as players following a simple heuristic or rule-of-thumb. Voters reelect the party that they perceive as the best one for future conditions, but punish a party that appointed the mayor but was not the most voted. This norm can be enforced simply by the notion that it is "unfair" or "undemocratic" for a party that did not win the most votes to appoint the mayor. Voters demanding that the most voted party appoints the mayor is consequential to welfare in the cases the most voted party has substantially higher vote share than the second most voted. However, this behaviour will appear to be based on ranks and bind behaviour even in cases where the consequences are minimal (parties almost tying).

In the equilibrium described in Proposition 2, the most voted party always appoints the mayor. In practice, second placed parties appoint the mayor with non-trivial frequency. There are two ways to reconcile this with the model. The first is that not all municipalities are in the equilibrium with the norm. Under this interpretation, our effects pin down the share of municipalities following the norm.[50] Another possibility is to incorporate deviations from the norm. These can occur due to trembling hand shocks in proposal strategies that make second placed parties appoint mayors on the equilibrium path and predict that voter punishment for such second placed parties can be observed.

Proposition 2 provides conditions for the norm to occur: for any party, the benefit of deviating from the norm is smaller than the benefit of continuing to follow it and perhaps obtain rents in a future period when it is the most voted party. Whether this condition is satisfied can be approximated empirically in two manners. First, we can observe if the third-placed party has been (or will be) the most voted in a past (or future) election. Second, the condition is more likely to be met when the third most voted party has a larger vote share.[51] The model predicts that the norm applies uniformly to all parties. This is consistent with the evidence we discuss in Section 4 and Supplementary Appendix A: effects are similar regardless of whether the most voted party is aligned with higher levels of government or the third-placed party.

5.7. *Welfare*

The equilibrium in Proposition 2 maximizes expected voter welfare. It guarantees the party most likely to match the state (given *aggregate* information) of the world appoints the mayor in every period. This is not the case in the equilibrium described in Proposition 1. While the model does not directly address equilibrium selection (why some places adopt the norm), its optimality can provide an explanation for its prevalence.[52]

50. Voters also prefer the second most voted party to appoint the mayor instead of the third most voted and thus the model could be extended to generate a norm that awards an advantage to the second most voted party over the third most voted. Such extension can be derived from an exogenous impediment to the most voted party appointing the mayor (*e.g.* a fixed probability the most voted party is not allowed to do so in a period). In such cases, norms where voters impose that the second most voted take preference over the third most voted can also be an equilibrium.

51. Intuitively, if vote shares are 45.5%, 44.5%, and 10%, the third party is unlikely to be the most voted in future elections. Compare that to an election where the vote shares are 35.5%, 34.5%, and 30%, where it is more likely that the third-placed party will be the most voted in a future election.

52. Proposition 2, however, specifies the necessary conditions for the equilibrium with the norm we study, and hence clarifies why it might not occur in some contexts. It is possible to construct an equilibrium where the second (or third) most voted must appoint the mayor. However, voters' expected welfare would be lower in such equilibrium.

6. EMPIRICAL IMPLICATIONS OF THE THEORETICAL FRAMEWORK

The section above described three additional testable predictions from the model. First, voters punish second most voted parties that appoint mayors. Second, the effect of being most voted is larger when the third-placed party is stronger. Third, following the norm leads to politicians of better quality and higher welfare. This section provides evidence on the first two, while Supplementary Appendix D addresses the third by studying the association between appointing most voted mayors and lower corruption.

6.1. *Do voters punish second-placed parties that appoint mayors?*

We test whether second most voted parties that appoint the mayor go on to lose votes compared to first most voted parties that also do so. For identification purposes, we focus on cases where the first and second placed parties have a vote share difference of 1 p.p. or less and the parties tied in seats.[53]

We estimate a triple-differences equation for vote share (v) of party p in municipality i at year t:

$$(v_{pi,t+1} - v_{pi,t}) = \alpha + \beta m_{pit} * f_{pit} + \gamma m_{pit} + \delta f_{pit} + \epsilon_{pit}, \tag{6.2}$$

where $m=1$ if party p appointed mayor, $f=1$ if party is first-place. The sample includes only first or second most voted parties and γ is the effect of appointing mayor for the second most voted and $\gamma+\beta$ is the effect for the first most voted. Our hypothesis is that $\beta>0$: first most voted parties that appoint mayors are rewarded compared to second most voted parties that do the same. Note that equation (6.2) nets out municipality-party fixed effects and time effects are absorbed into α. Standard errors are clustered at the municipality level.

The results are shown in Column (1) of Table 5, which shows that first placed parties that appoint mayors observe a subsequent growth in vote shares that is 4.8 p.p. larger than second placed parties appoint mayors. Columns (2) and (3) show similar results when a full set of region and party indicators is added (which control for region-time and party-time variation, given the first-difference specification).

6.1.1. Dynamics and parallel trends. To test whether first and second placed parties are on parallel trends, column (6) estimates equation (6.2) with lagged outcomes. It finds no significant effect, suggesting that the previous result is not driven by a pre-existing trend. To further probe the dynamics of the effect, Figure 6a provides the event-study counterpart for equation (6.2) from estimating

$$(v_{pi,t+k} - v_{pi,t}) = \alpha_k + \beta_k m_{pit} * f_{pit} + \gamma_k m_{pit} + \delta_k f_{pit} + \epsilon_{pit} \tag{6.3}$$

separately for k equal to -3, -2, -1, 1, 2, and 3. Figure 5b plots the β_k against k, as well as their 95% confidence intervals. The graph indicates no pre-existing trends (*i.e.* zero placebo effects on lagged outcomes) and suggests that the differential effect of a mayoral appointment for first-placed parties dissipates after two elections, although perhaps not fully.

53. This is the same sample that appears in column (2) of Table 3. The sample size differs since not all parties run in two consecutive elections, or local parties change names making it impossible to identify them over time.

TABLE 5
Do voters punish parties that break the norm? triple-difference estimates

	Outcome $(v_{pm,t+1}-v_{pm,t})$					Lagged outcome $(v_{pm,t}-v_{pm,t-1})$
	(1)	(2)	(3)	(4)	(5)	(6)
Mayor$_t$*Most Voted$_t$ (β)	4.834**	4.232**	3.382*	5.116*	3.504	−1.263
	(1.931)	(1.881)	(1.814)	(2.909)	(2.512)	(1.735)
Mayor$_t$ (γ)	2.868**	3.091**	2.937**	1.175	4.998**	3.822***
	(1.449)	(1.469)	(1.404)	(2.059)	(2.046)	(1.441)
Most Voted$_t$ (δ)	−1.693	−1.532	−1.399	−0.691	−2.017	−1.184
	(1.393)	(1.355)	(1.304)	(2.035)	(1.898)	(1.121)
Constant	−1.160	−0.681	6.630	0.753	−3.372***	−0.877
	(0.785)	(3.025)	(5.646)	(1.055)	(1.163)	(0.783)
Region FE		Y	Y			
Party FE			Y			
Only elections w. vote share of 3rd > median				Y		
Only elections w. vote share of 3rd < median					Y	
N	664	664	664	332	332	694

Notes: Standard errors clustered at the municipality level in parentheses. The unit of observation is a party-municipality-election. See discussion of equation (6.2) in text for specification. Outcome in columns (1)–(5) is the growth in vote share between the election immediately preceding a possible mayoral appointment (t) and the next election ($t+1$). Outcome in column (6) is growth between time t and $t-1$ (a placebo test). The sample is restricted to elections in which the two most voted parties tied in seats and their difference in vote shares was less than 1% of the total vote.

6.1.2. Interpretation. Given the triple-difference nature of the estimation, it is not clear whether the effects of Figure 6a are driven by most voted parties gaining more votes than second most voted parties that do so, or the latter losing votes. In other words, the effects are relative to the counterfactual of the other party (and can thus be interpreted as a "reward" for the most voted or a "punishment" for second most voted when they appoint a mayor). Figure 6b illustrates this by providing the double-difference event study graph for both second placed parties and first placed parties separately. In particular, we estimate equation (6.3) separately for only second most voted parties (red squares) and first most voted parties (blue circles).[54] While the second most voted party that appoints a mayor gains votes (over a second most voted party that does not), this can to be explained by the continuation of a pre-existing trend (parties that appoint mayor are in positive trajectories). First most voted parties are on a similar trajectory before appointing a mayor, but go on to gain even more after.

While Figure 6a indicates a non-zero effect four years after the mayoral appointment, it is not as conclusive regarding whether parties continue to be punished eight or twelve years later. The point estimates for are positive but not significantly different from zero, so it is not possible to discern if the punishment fully dissipates after two elections, or if it only dissipates partially and we lack the statistical power to detect the smaller effects in later elections.

6.2. *Heterogenous effects by strength of the third-placed party*

Table 6 provides evidence supporting the model's prediction that, in elections where the third most voted party has a higher vote share, the effect of being most voted should be stronger. It repeats our estimation of the main results (Table 3), but separating the sample into cases where

54. This implies we estimate $(v_{pi,t+k}-v_{pi,t})=\alpha_k+\beta_k m_{pit}+\epsilon_{pit}$ when using only the second most voted parties and $(v_{pi,t+k}-v_{pi,t})=(\alpha_k+\delta_k)+(\beta_k+\gamma_k)m_{pit}+\epsilon_{pit}$ when using only the first most voted parties.

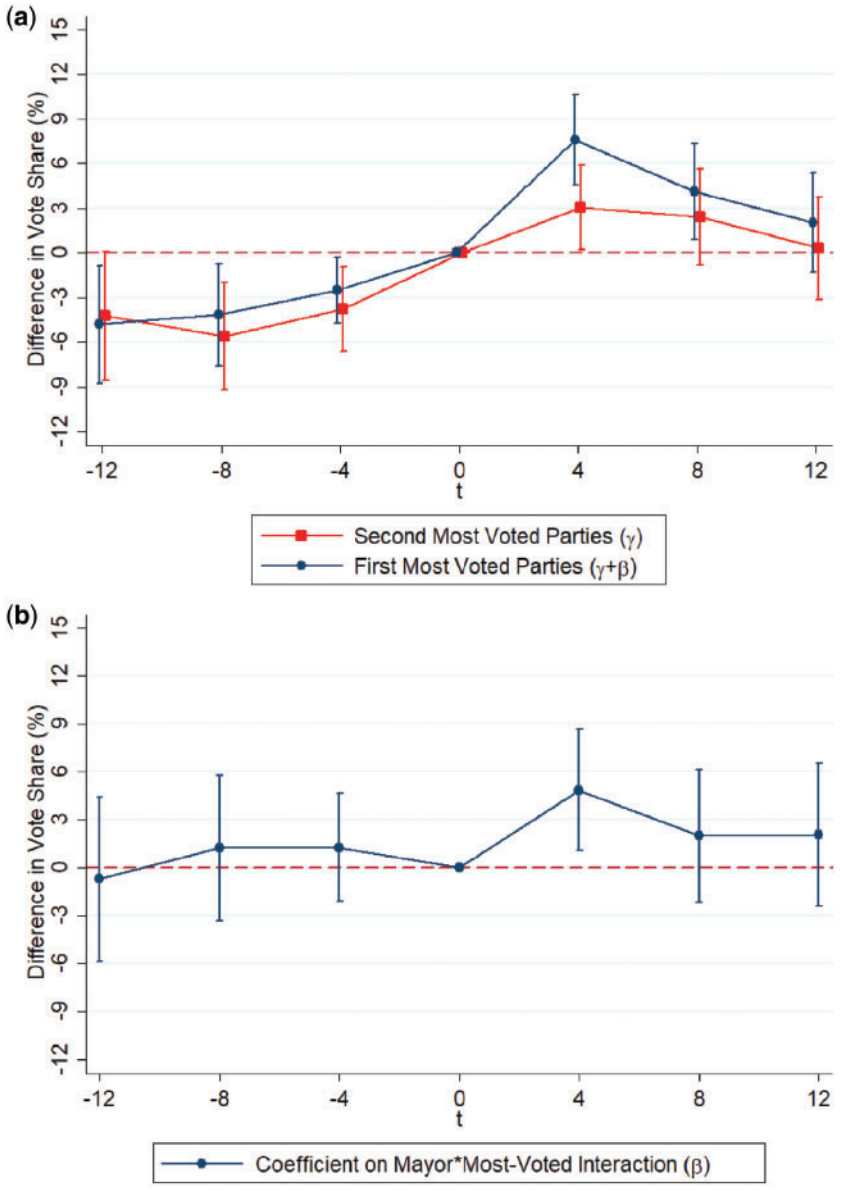

FIGURE 6

(a): Event study for effect of mayoral appointment on vote shares, triple-differences; (b) Event study for effect of mayoral appointment on vote shares, by party rank.

Notes: Whiskers represent 95% confidence intervals based on standard errors clustered at the municipality level. Sample is restricted to elections in which the two most voted parties tied in seats and their difference in vote shares was less than 1% of the total vote. Vote shares are normalized to zero at $t=0$. Red squares (blue circles) in Panel B show how the share of votes for a second-placed (first-placed) party that appoints a mayor at $t=0$ evolves relative to a second-placed (first-placed) party that does not, obtained by estimating γ and $\gamma+\beta$ from equation (6.2) with different time horizons (see text for further details). Blue circles in Panel A represent the triple-difference event study: the difference between Panel B markers, obtained by estimating β from equation (6.2) with different time horizons (see text for further details).

TABLE 6
Effect of being most voted on appointing the mayor, by strength of third-placed party

Dependent variable	2nd-pl. Mean	(1)	(2)	(3)	(4)
Panel A: third most voted party vote share above median					
Party appointed	0.290	0.306***	0.290***	0.364***	0.334***
Mayor		(0.078)	(0.056)	(0.048)	(0.062)
N		1014	468	2756	2756
Panel B: third most voted party vote share below median					
Party appointed	0.430	0.0402	0.103	0.214***	0.118*
Mayor		(0.088)	(0.066)	(0.059)	(0.070)
N		1014	408	3040	3040
p-value: test of equal effects		0.0228	0.0321	0.0498	0.0212
Specification:		Linear	Means	Quad.	Cubic
Bandwidth:		Optimal	<1%	Full	Full

Notes: 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 vote share above (below) the median of the sample used in column (1): 16.5%. 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).

the third most voted party vote share is above (Panel A) and below (Panel B) the median.[55] The effects are larger in Panel A, and it is possible to reject that the the effects in both subsamples are the same in all specifications across columns (at the 5% level). Figure A18 provides the graphical counterpart.[56]

Table A7 repeats this exercise using another criterion to separate the two subsamples. Panel A focuses on the cases where the third most voted party has been (or will be) the most voted party in at least one of the last (next) three elections. Panel B focuses on the remaining cases. Again, the effect is substantially larger in Panel A. Figure A19 provides the graphical counterpart.[57]

Finally, we test whether voters' punishment for second most voted parties that appoint mayors is more evident in cases where the third-placed party is stronger. We find that this is indeed the case, although the estimates are noisily estimated (likely given the smaller subsamples). Column (4) of Table 5 estimates equation (6.2) for the subsample where the vote share of the third-placed party is above the sample median, while column (5) repeats this for the case below the sample median. The estimated differential effect for most voted parties (β) is larger in the former case.

55. We use the median of the sample with optimal bandwidth, which is a vote share of 16.5%. In the above (below) median subsample, the average vote share of the two most voted parties is 40% (33%) each, with the third most voted obtaining 12% (22%).

56. Figure A18 is constructed similarly to Figure 1a, but for each subsample. Figure A18b illustrates why the effect for the below median subsample varies across columns in Table 6, there is a non-linearity close to the cutoff that is not captured by the specifications using the entire sample, which find a larger effect than local estimates (columns 1 and 2). Figure A13 shows that this heterogeneity holds within some council sizes.

57. This criterion is the closest empirical approximation of the conditions on Proposition 2 (that the probability the third most-voted goes is a winner in the future is sufficiently high). Compared to the criterion of Table 6, it has the drawbacks that it requires a reduced sample (it can only be defined for cases where we observe the outcomes of the last and next three elections) and that most of the observations appear on Panel B.

7. ALTERNATIVE EXPLANATIONS

While the results we document are consistent with the existence of the norm modeled in Section 5, there are other possible explanations. This section discusses two alternatives (the status quo rule and voters preferring "winners"), while Supplementary Appendix F discusses the role for politician's outcome bias, parties making informal agreements, and the central government preferring most voted parties.

7.1. *Status quo rule*

The only differential *institutional* treatment of parties by rank of their votes in Spanish municipal elections is the status quo rule described in Subsection 4.1. While, at first pass, this appears to be likely to explain our results, we believe the status-quo rule cannot be the main driver of our results. 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. Supplementary Appendix E discusses in further detail other reasons why we believe the status quo rule does not play a role in explaining our results.

7.2. *Voters prefer "winners"*

Another explanation for our results relies on assuming that voters' preferences are such that they inherently prefer the most voted party to appoint mayor. This could occur because "winning" attaches symbolic value to a party. This would in turn imply voters punish mayors from second most voted parties and tilt bargaining outcomes towards the most voted party.

First, note that this "alternative" explanation is, in some aspects, similar in spirit to our model. One of its goals is illustrating an information aggregation rationale for voters to prefer, in equilibrium, the most voted party to appoint the mayor. This alternative explanation, on the other hand, relies on directly assuming that voters prefer mayors that are the most voted. This highlights the complementary character of the distinct mechanisms. Perhaps because there is an equilibrium logic for preferring the most voted parties to appoint mayors, voters incorporate this into their preferences over time.[58]

Given that our model provides a rationale for voters to prefer most voted parties to appoint mayors, many of its empirical predictions would be the same in other explanations based on voters' inherently having such preferences. We highlight two predictions, however, that are obtained naturally from our model but require further assumptions to be explained by this alternative explanation. First, it is straightforward to extend the logic of our model to generate the second-versus-third most voted results discussed in Subsection 4.3. In the alternative explanation, one would need to assume that voters not only prefer election "winners" over "losers", but also "runner-ups" over "third-places". Second, our model predicts that the effects are stronger when the third-placed party has a higher vote share. The same prediction is not entirely clear in the alternative explanation. One could envision that the punishment for "second most voted mayors" has larger effects on bargaining outcome when the third-placed party is weaker (since this would imply that the two most voted parties are closer to winning a majority in the future, so the punishment in future elections is more likely to matter). If that is the case, this alternative explanation predicts the opposite of the results in the previous subsection. On the other hand, one could also argue

58. Note that, in our model, the concept of social norms is understood as different equilibria. In this alternative explanation, social norms would be encoded in voters' preferences.

that competition among parties is stronger when the third party has a higher vote share, yielding the opposite prediction.

8. CONCLUSION

Our main result indicates that simply being labelled the "most voted" has, in itself, a substantial effect on parties' bargaining outcomes in a legislature. This result is difficult to reconcile with existing theories of multilateral bargaining and coalition formation. The overall evidence is consistent with the existence of a norm that voters enforce by punishing parties that deviate from it. This can explain why parties support this most voted advantage even when a natural alternative, such as programmatically aligned lower-ranked parties forming a winning coalition, is available.

We believe there are two fruitful avenues for further research. One is studying the welfare consequences of this effect. Another is investigating why this effect is present in some jurisdictions and not others. In light of our model, that would require understanding the equilibrium (norm) selection mechanisms.

Acknowledgments. We thank Irma Clots-Figueras, Albert Falcó-Gimeno, Olle Folke, Nicola Gennaioli, M. Daniele Paserman, Rohini Pande, Francesco Trebbi, Leeat Yariv, three anonymous referees, and seminar participants at Bank of Spain, Berkeley Haas, Brown, Caltech, CEMFI, Chicago Harris, Columbia, Cornell, Cowles-Yale Political Economy Conference, EEA-ESEM Conference, HBS, IDB, IEB Political Economy Workshop, Kellogg, NES Development/Political Economy Workshop, Princeton, Rutgers, Spanish Economic Association Symposium, Stanford GSB, Wharton, and UBC for helpful comments. We are grateful to Albert Solé-Ollé and Pilar Sorribas-Navarro for sharing the corruption data. Vishan Nigam and Susannah Scanlan provided excellent research assistance. A previous version of this article has circulated under the title "Norms in Bargaining: Evidence from Government Formation in Spain".

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

ABC (2004), Chaves acepta el reto de Teófila y promete no gobernar si no es el más votado. http://sevilla.abc.es/hemeroteca/historico-22-01-2004/sevilla/Andalucia/chaves-acepta-el-reto-de-teofila-y-promete-no-gobernar-si-no-es-el-mas-votado/_183979.html.

AGRANOV, M., GOEREE, J. K., ROMERO, J. *ET AL.* (2018), "What Makes Voters Turn Out: The Effects of Polls and Beliefs", *Journal of the European Economic Association*, **16**, 825–856.

ANAGOL, S. and FUJIWARA, T. (2016), "The Runner-Up Effect", *Journal of Political Economy*, **124**, 927–991.

ANDERSSON, S., BERGMAN, T., and ERSSON, S. (2014), The European representative democracy data archive, release 3. Main sponsor: Riksbankens Jubileumsfond (In2007-0149:1-E). `www.erdda.se`.

AUSTEN-SMITH, D. and BANKS, J. (1990), "Stable Governments and the Allocation of Policy Portfolios", *American Political Science Review*, **84**, 891–906.

BAGUES, M. and CAMPA, P. (2017), "Can Gender Quotas Empower Women? Evidence from a Regression Discontinuity Design", Unpublished Manuscript.

BANZHAF III, J. F. (1964), "Weighted Voting Doesn'T work: A Mathematical Analysis", *Rutgers Law Review*, **19**, 317.

BARON, D. P. (1991), "A Spatial Bargaining Theory of Government Formation in Parliamentary Systems", *American Political Science Review*, **85**, 137–164.

BARON, D. P. (1993), "Government Formation and Endogenous Parties", *American Political Science Review*, **87**, 34–47.

BARON, D. P. and FEREJOHN, J. A. (1989), "Bargaining in Legislatures", *American Political Science Review*, **83**, 1181–1206.

BARRO, R. J. (1973), "The Control of Politicians: An Economic Model", *Public Choice*, **14**, 19–42.

BORDALO, P., GENNAIOLI, N., and SHLEIFER, A. (2013), "Salience and Consumer Choice", *Journal of Political Economy*, **121**, 803–843.

CALLANDER, S. (2007), "Bandwagons and Momentum in Sequential Voting", *The Review of Economic Studies*, **74**, 653–684.

CALLANDER, S. and WILSON, C. H. (2008), "Context-dependent Voting and Political Ambiguity", *Journal of Public Economics*, **92**, 565–581.

CARROLL, R. and COX, G. W. (2007), "The Logic of Gamson's Law: Pre-Election Coalitions and Portfolio Allocations", *American Journal of Political Science*, **51**, 300–313.

CASAJUS, A. and HUETTNER, F. (2018), "The Coleman-Shapley-index: Being Decisive Within the Coalition of the Interested", (ESMT Working Paper 18-03).
DALY, M. (2017), National could get many more votes than any other party and still not be in the next government. Stuff, https://www.stuff.co.nz/national/politics/97079414/national-could-get-many-more-votes-than-any-other-party-and-still-not-be-in-the-next-government.
DIERMEIER, D., ERASLAN, H., and MERLO, A. (2003), "A Structural Model of Government Formation", *Econometrica*, **71**, 27–70.
DIERMEIER, D. and MERLO, A. (2004), "An Empirical Investigation of Coalitional Bargaining Procedures", *Journal of Public Economics*, **88**, 783–797.
EL PAÍS (2004), Zapatero anuncia que solo intentará formar gobierno si logra más votos que Rajoy. https://elpais.com/diario/2004/01/19/espana/1074466803_850215.html.
EL PAÍS (2015), Clima social tras las elecciones del 24-M. http://elpais.com/elpais/2015/06/19/media/1434733713_925358.html.
EL PAÍS (2018), Rajoy critica a los ayuntamientos del cambio por "repartirse despachos para arrinconar al PP". https://politica.elpais.com/politica/2018/05/05/actualidad/1525526954_712707.html.
EUROPA PRESS (2007), Rajoy emula a Zapatero y se compromete a no gobernar si su partido no es el más votado. http://www.europapress.es/nacional/noticia-rajoy-emula-zapatero-compromete-no-gobernar-si-partido-no-mas-votado-20070904000108.html.
EUROPA PRESS (2011), El PP tendra la alcaldía de Olivenza, pueblo natal de Vara, dado que IU no apoyará a ningún otro grupo. http://www.europapress.es/nacional/noticia-22-pp-tendra-alcaldia-olivenza-pueblo-natal-vara-dado-iu-no-apoyara-ningun-otro-grupo-20110609230840.html.
FEREJOHN, J. (1986), "Incumbent Performance and Electoral Control", *Public Choice*, **50**, 5–25.
FLEGENHEIMER, M. (2016), Ted Cruz, facing hard G.O.P. path, seeks a good argument. NYTimes, https://www.nytimes.com/2016/04/22/us/politics/ted-cruz-campaign.html?_r=0.
FOLKE, O., PERSSON, T., and RICKNE, J. (2016), "The Primary Effect: Preference Votes and Political Promotions", *American Political Science Review*, **110**, 559–578.
FRECHETTE, G., KAGEL, J. H., and MORELLI, M. (2005), "Nominal Bargaining Power, Selection Protocol, and Discounting in Legislative Bargaining", *Journal of Public Economics*, **89**, 1497–1517.
GOLDER, S. N. (2010), "Bargaining Delays in the Government Formation Process", *Comparative Political Studies*, **43**, 3–32.
GONZALEZ-EIRAS, M. and SANZ, C. (2018), "Women's Representation in Politics: Voter Bias, Party Bias, and Electoral Systems", *Bank of Spain Working Paper 1834*.
IMBENS, G. and KALYANARAMAN, K. (2012), "Optimal Bandwidth Choice for the Regression Discontinuity Estimator", *The Review of Economic Studies*, **79**, 933–959.
LAVER, M. (1998), "Models of Government Formation", *Annual Review of Political Science*, **1**, 1–25.
LAVER, M. and BENOIT, K. (2015), "The Basic Arithmetic of Legislative Decisions", *American Journal of Political Science*, **59**, 275–291.
LAVER, M. and SHEPSLE, K. A. (1990), "Coalitions and Cabinet Government", *American Political Science Review*, **84**, 873–890.
LEE, D. S. and LEMIEUX, T. (2010), "Regression Discontinuity Designs in Economics", *Journal of Economic Literature*, **48**, 281–355.
LIZZERI, A. and PERSICO, N. (2001), "The Provision of Public Goods under Alternative Electoral Incentives", *American Economic Review*, **91**, 225–239.
MAGRE-FERRAN, J. and BERTRANA-HORTA, X. (2005), "Municipal Presidentialism and Democratic Consolidation in Spain", in R. Berg and N. Rao (eds), *Transforming Local Political Leadership* (New York: Palgrave Macmillan), 73–84.
MASKIN, E. and SEN, A. (2016), How majority rule might have stopped Donald Trump. https://mobile.nytimes.com/2016/05/01/opinion/sunday/how-majority-rule-might-have-stopped-donald-trump.html?referer=.
MCKINNEY, C. J. (2015), Does the largest political party form the government? https://fullfact.org/law/does-largest-political-party-form-government/.
MERILÄINEN, J. and TUKIAINEN, J. (2018), "Rank Effects in Political Promotions", *Public Choice*, **177**, 87–109.
MERLO, A. (1997), "Bargaining over Governments in a Stochastic Environment", *Journal of Political Economy*, **105**, 101–131.
MERLO, A. and WILSON, C. (1995), "A Stochastic Model of Sequential Bargaining with Complete Information", *Econometrica: Journal of the Econometric Society*, 371–399.
MORELLI, M. (1999), "Demand Competition and Policy Compromise in Legislative Bargaining", *American Political Science Review*, **93**, 809–820.
MURRAY, M. (2016), Poll: 62% say republican with most votes should be nominee. NBC News, http://www.nbcnews.com/politics/first-read/poll-62-say-republican-most-votes-should-be-nominee-n556846.
PERRAUDIN, F. (2015), Lib Dems would give biggest party time to form coalition, says Clegg. https://www.theguardian.com/politics/2015/may/05/nick-clegg-let-largest-party-after-election-form-government.
PERSSON, T., ROLAND, G., and TABELLINI, G. (1997), "Separation of Powers and Political Accountability", *The Quarterly Journal of Economics*, **112**, 1163–1202.

PONS, V. and TRICAUD, C. (2018), "Rankings Matter Even When They Shouldn't: Bandwagon Effects in Two-Round Elections" (Harvard Business School Working Paper, No. 19-026).
RAY, D. and VOHRA, R. (2014), "Coalition Formation", *Handbook of Game Theory 4*, 239–326.
SANZ, C. (2017), "The Effect of Electoral Systems on Voter Turnout: Evidence from a Natural Experiment", *Political Science Research and Methods*, **5**, 689–710.
SHAPLEY, L. S. and SHUBIK, M. (1954), "A Method for Evaluating the Distribution of Power in a Committee System", *American Political Science Review*, **48**, 787–792.
SILVER, N. (2016), Why Republican voters decided on Trump. https://fivethirtyeight.com/features/why-republican-voters-decided-on-trump/.
SNYDER Jr, J. M., TING, M. M., and ANSOLABEHERE, S. (2005), "Legislative Bargaining under Weighted Voting", *American Economic Review*, **95**, 981–1004.
STEVENSON, R. W. (2004), Confident bush outlines ambitious plan for 2nd term. NYTimes, https://www.nytimes.com/2004/11/05/politics/campaign/confident-bush-outlines-ambitious-plan-for-2nd-term.html.
SWEETING, D. (2009), "The Institutions of 'Strong' Local Political Leadership in Spain", *Environment and Planning C: Government and Policy*, **27**, 698–712.
THE GUARDIAN (2018), The Five Star Movement has earned the right to govern Italy. https://www.theguardian.com/commentisfree/2018/mar/07/five-star-movement-govern-italy-corruption-economy-berlusconi.
VALLÉS, J. M. and BRUGUÉ, Q. (2001), "El gobierno local", *Política y Gobierno en España*, 103–121.
WARWICK, P. V. and DRUCKMAN, J. N. (2001), "Portfolio Salience and the Proportionality of Payoffs in Coalition Governments", *British Journal of Political Science*, **31**, 627–649.
WELLS, P. (2015), Paul Wells: What if Harper wins a minority? Maclean's, https://www.macleans.ca/politics/ottawa/paul-wells-what-if-stephen-harper-wins-a-minority/.