

Roll-Call Vote Selection: Implications for the Study of Legislative Politics

CAITLIN AINSLEY *University of Washington*

CLIFFORD J. CARRUBBA *Emory University*

BRIAN F. CRISP *Washington University in St. Louis*

BETUL DEMIRKAYA *University of Kentucky*

MATTHEW J. GABEL *Washington University in St. Louis*

DINO HADZIC *Trinity College Dublin*


Roll-call votes provide scholars with the opportunity to measure many quantities of interest. However, the usefulness of the roll-call sample depends on the population it is intended to represent. After laying out why understanding the sample properties of the roll-call record is important, we catalogue voting procedures for 145 legislative chambers, finding that roll calls are typically discretionary. We then consider two arguments for discounting the potential problem: (a) roll calls are ubiquitous, especially where the threshold for invoking them is low or (b) the strategic incentives behind requests are sufficiently benign so as to generate representative samples. We address the first defense with novel empirical evidence regarding roll-call prevalence and the second with an original formal model of the position-taking argument for roll-call vote requests. Both our empirical and theoretical results confirm that inattention to vote method selection should broadly be considered an issue for the study of legislative behavior.


THE IMPORTANCE OF LEGISLATIVE VOTING PROCEDURES


Voting records are a key window into legislative processes around the world. Scholars use individually recorded legislative votes, or roll-call votes (RCVs), to study the influence of legislative parties on voting behavior (unity), the preferred outcomes of parties on key issues (position taking), and the character of legislative conflict (polarization).


These studies regularly rely on measures of legislator ideal points, party cohesion, and/or the dimensionality of the policy space that are derived from the RCV record.¹ It is only infrequently noted that legislators in most assemblies can vote by multiple methods (e.g., voice votes), making the RCV record a sample of legislative voting behavior.


The quality of any measurement strategy using RCV results depends crucially on how well the RCV record captures relevant aspects of *all* legislative voting behavior. The process that generates RCVs is thus a potential source of selection bias, a concern that has largely been overlooked or dismissed in past studies.² This omission may be justified. First, legislatures where roll-call voting is the standard operating procedure will likely generate large voting records through roll call. Second, where RCVs are discretionary, the rules governing the use of RCVs may make it easy to invoke the process, perhaps generating a sufficiently representative sample of votes. For example, Poole and Rosenthal (2007), in their study of the U.S. Congress, state that “[t]he selection-bias story is logically correct but empirically irrelevant” (Poole and Rosenthal 2007, 68). They discount its practical relevance because, for the


Caitlin Ainsley , Assistant Professor, Department of Political Science, University of Washington, cainsley@uw.edu

Clifford J. Carrubba , Professor, Departments of Political Science and Quantitative Theory and Methods, Emory University, ccarrub@emory.edu

Brian F. Crisp , Professor, Department of Political Science, Washington University in St. Louis, crisp@wustl.edu

Betul Demirkaya , Assistant Professor, Department of Political Science, University of Kentucky, betul.demirkaya@wustl.edu

Matthew J. Gabel , Professor, Department of Political Science, Washington University in St. Louis, mgabel@wustl.edu

Dino Hadzic , Assistant Professor, Department of Political Science, Trinity College Dublin, dhadzic@tcd.ie

This research was supported by National Science Foundation Grant # SES-1066340, awarded to Carrubba, Crisp, and Gabel and funding from the Weidenbaum Center at Washington University in St. Louis. The authors thank Bill Bernhard, Randall Calvert, and four anonymous reviewers for their helpful comments and suggestions on earlier versions of this paper. Replication files are available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/XI7ENB>.

Received: May 30, 2018; revised: July 05, 2019; accepted: April 03, 2020.

¹ For a survey of studies that estimate ideal points from roll calls, see Carroll and Poole (2014). For a survey of studies of party unity, see Sieberer (2006). For a review of alternative measures of legislators' preferences, see Section C of the Online Appendix.

² Exceptions include Carrubba et al. (2006), Clinton and Lapinski (2008), Høyland (2010), Hug (2010), Lynch and Madonna (2013), Roberts and Smith (2003), Stecker (2015), and Van Doren (1990).

U.S. Congress, they contend that the composition of the RCV record is sufficiently heterogeneous to generate estimates of legislators' ideal points and the dimensionality of the policy space.³

In this paper we examine the roll-call vote selection process and its sample properties in three steps. First, we present a novel dataset of legislative voting rules that significantly expands our understanding of the voting methods available to legislators and the institutional rules for choosing among them across contemporary legislatures. The key conclusion from our catalogue is that roll-call votes are not the standard operating procedure (SOP) in most legislatures, leaving open the possibility that sample bias is widespread.

In a second step, we present data on the prevalence of RCVs among legislative votes taken by multiple methods for a subset of these catalogued legislatures. While the raw frequency of RCVs seems to increase as thresholds for invoking the method decrease (Carey 2009), we show that the *prevalence* of RCVs is generally not higher in legislatures with the lowest thresholds for invoking the method. Thus, one cannot infer from procedures that make it easy to invoke an RCV *or* even from a large volume of RCVs that we will automatically have a useful sample of votes taken via roll call. Unfortunately, absent the observation of individual legislator-level choices on non-RCVs, we cannot directly evaluate the sample properties of the RCV record.⁴ Instead, one must turn to theory.

As a third step, we do so, examining theoretically whether position-taking incentives—a common informal argument for what gets roll-called—affect the character of the RCV sample. Using a formal model and simulations, we demonstrate that measures of ideal points, indicators of party unity, and efforts to capture the dimensionality of the policy space will be plagued by estimation problems if position-taking is the motivating story. Critically, this model is built on a set of assumptions that help maximize the chances an RCV sample is representative of all votes.⁵ We demonstrate that the estimation problems are in fact a function of the prevalence of RCVs, not the total number of them cast. For plausible levels of RCV prevalence, we find serious estimation problems. In this paper, we seek to make clear the scope of the problem and to show that some of the currently accepted assumptions about benign data-generating processes do not hold.

³ Some scholars remain concerned about the quality of the RCV sample in the United States including Clinton and Lapinski (2008), Lee (2018), Lynch and Madonna (2013), Roberts and Smith (2003), Snyder (1992), and Van Doren (1990).

⁴ For the one exception of which we are aware, see Hug (2010).

⁵ Previous studies point out similar estimation problems due to a variety of aspects of the policy-making process such as party discipline (McCarty, Poole and Rosenthal 2001), partisan agenda control, and veto players (Clinton 2007). Our results indicate that estimation problems persist even in the absence of these concerns.

PREVIOUS EFFORTS TO EXAMINE VOTE PROCEDURES ACROSS NATIONAL LEGISLATURES

We currently lack a systematic and comparable description of the types and frequency of voting methods employed in contemporary legislative chambers around the world. Apart from a few important exceptions described below, roll-call vote studies typically focus on one legislature and provide at best a cursory discussion of voting procedures and the rules governing their use. Further, we are aware of only a handful of studies that describe or evaluate the sample properties of roll-call votes (Carrubba et al. 2006; Clinton 2007; Crisp and Driscoll 2012; Hug 2010; Lynch and Madonna 2013; Roberts and Smith 2003; Saalfeld 1995; Van Doren 1990).⁶ That said, scholars have compiled notable—but still limited—cross-national comparisons of legislative voting procedures. We briefly review them below, as they provide an important reference for our analysis of an original catalogue of voting rules presented in the next section. In addition, a discussion of these studies highlights several important empirical claims that we hope to engage with our new dataset.

The Inter-Parliamentary Union (Inter-Parliamentary Union 1986) produced an impressive catalogue of legislative voting rules, based on a survey of legislative staff in 1985. More recently, four region-specific studies report the rules governing roll-call votes and their use. Based on interviews and the published standing orders in the early 2000s, Carey (2009) catalogued the SOP and the frequency of RCVs for 24 legislatures in the Americas. Crisp and Driscoll (2012) provide a slightly more recent description of most of the legislatures covered in Carey (2009) along with several additional ones from Latin America.

Two studies have focused on European legislatures. Saalfeld (1995) catalogued voting rules and reported the frequency of RCVs in 19 Western European legislatures as of the mid-1980s. Hug, Wegmann, and Wüest (2015) provide a recent and more comprehensive review of voting procedures in 54 European parliaments. Based on an expert survey, they catalogue the SOP for votes on final passage of legislation and provide information on how alternative voting methods could be invoked.

These studies all reach three fairly similar conclusions. First, vote results recorded at the level of the individual legislator are rarely the only form of legislative voting available. Second, when considering votes on most legislative business, voting by roll-call is rarely even the standard operating procedure. Thus, RCVs are almost always a sample of all votes taken, raising the potential that selection bias in RCVs is widespread. Third, the raw number of roll-call votes varies significantly across legislatures and, in general, it is inversely related to the technological and procedural obstacles for invoking an RCV.

⁶ These studies all conclude that roll-call votes provide a biased sample of legislative votes.

Hansen and Debus (2012, 715) contend that any selection bias should be minor in a legislature where the procedural rules for requesting recorded votes are low because the ease of invoking RCVs should ensure a relatively large and diverse sample of votes. Thus, exactly the chambers where RCVs are plentiful, and therefore attractive for empirical study, are likely those where the selection bias is small. However, whether this contention holds remains an open question. None of the studies described above reports the *prevalence* of RCVs among all of the forms of votes taken.⁷ Consequently, we return directly to this contention in our ensuing empirical *and* theoretical analyses—finding, in both cases, that it does not hold.

A CATALOGUE OF THE PROCEDURES FOR CHOOSING AMONG VOTE METHODS IN CONTEMPORARY LEGISLATURES

We have systematically assembled and coded information on legislative voting procedures for 145 legislative chambers across 105 countries spanning Europe, Africa, the Americas, and Asia.⁸ This procedural information was collected from the current (circa 2018) official standing orders (cameral rules of procedure) published by each legislature.⁹ We selected these legislatures because the standing orders were available in digital form from their official government website and were published in a language in which we had literacy.¹⁰ Our dataset allows for an evaluation of contemporary rules governing the choice of legislative voting methods for a significantly larger set of cases than were included in previous studies.

Further, for a subset of these legislatures, we collected both the raw number of roll-call votes and the *prevalence* of those votes as a share of votes taken by all available methods. Note that the calculation of *prevalence* requires a comprehensive count of all legislative votes. We are aware of no legislature that publicly reports this number. However, a careful reading of the plenary minutes or related documents for some legislatures provides the relevant information. We

therefore surveyed the legislatures in our dataset for electronic versions of their plenary minutes that provided this information in a consistent format and in a language in which we had expertise. Beyond four chambers that have been identified in previous studies, we found an additional 11 legislative chambers that fit these criteria.¹¹

This sample of legislatures is not enormous, but it allows a much larger view of RCV *prevalence* than has been reported in any previous study. Again, reporting *prevalence* in RCV studies is extremely rare, and even those rare instances typically involve only a single legislature. Furthermore, the available legislatures vary significantly in terms of the threshold for invoking an RCV. This variation allows us to examine the relationship between the size of the threshold and RCV *prevalence*, which is central to the argument by Hansen and Debus (2012, 715) for why the selection process for RCVs is benign. Particularly, the argument claims that if legislators and parties are focused on position-taking for their constituents, the low threshold should generate a relatively large and diverse share of votes. Our data will engage this argument more directly than has any study to date. We will also engage it in the simulations that follow, showing that prevalence is critical to the sample properties.

We rely on cameral rules for defining voting methods. These rules govern the procedures for legislative debate and decision making, including voting rules. In some instances, the rules identify a single vote procedure (e.g., rising in places), suggesting that all of a legislature's business will be conducted under that procedure. Other cameral rules define multiple voting procedures and their application to different types of decisions. For example, votes confirming presidential appointees may be set aside as decisions to be taken via a secret vote. In other cases, the cameral rules may single out a vote on the final passage of legislation for roll call, while applying other methods to votes on earlier parts of the legislative process (e.g., amendments to or votes on individual articles within a larger bill). Finally, cameral rules that permit multiple methods of voting define the mechanisms by which the SOP—whichever form of voting it is—can be set aside in favor of an alternative form of taking a vote.

For the SOP, we report voting methods defined for *most legislative business*—that is, any type of decision not explicitly singled out for the use of a specific, special voting method. For the purposes of this catalogue, we defined the voting method as an RCV if it produces a record of individual-level votes that is accessible to the public (e.g., through vote tables). Non-RCV methods are those that fail to meet this definition and commonly include voting by voice, rising in places, or secret ballot.

One might argue that only final votes on the passage of an entire piece of legislation merit attention. However, if we are interested in describing how individual legislators vote on legislative business and the character of legislative conflict, then final passage votes are only

⁷ Crisp and Driscoll (2012) report prevalence for 2 of the 24 legislatures they catalogued, which is not sufficient for analyzing this contention.

⁸ These cases include a range of regime types in terms of their level of democracy. Legislative behavior in authoritarian regimes has received a significant amount of attention (Desposato 2001; Jensen, Malesky, and Weymouth 2014; Malesky and Schuler 2010; Malesky, Schuler, and Tran 2012). We collected procedural information for both democratic and nondemocratic countries. As we discuss the entire sample, throughout the paper we also describe the patterns that emerge for exclusively democratic regimes. Our findings do not vary much across the two sets of cases.

⁹ We note a small number of exceptions where we relied on other information published by the legislature to discern relevant information.

¹⁰ We use recently adopted rules that we could not confirm are currently in operation for three legislatures: Ghana's Parliament, Micronesia's Congress, and Papua New Guinea's Parliament. The inclusion or exclusion of these rules from our analysis has no effect on the conclusions we draw about voting rules.

¹¹ We employed text analytic tools to scrape the relevant vote information from the plenary minutes.

one of many types of relevant votes. This comprehensive view of votes is consistent with much of the empirical literature that employs roll-call vote data to learn about legislative behavior (e.g., Bräuninger, Müller, and Stecker 2016; Carey 2007; Carson et al. 2010; Crespin and Rohde 2010; Jenkins 2008; Poole and Rosenthal 2007). After all, votes setting the day's agenda or determining the fate of a single (poison) amendment can significantly affect the substance of legislative business and elicit legislator's ideological positions, the influence of party leadership, and other relevant aspects of legislative behavior. Singling out votes of final passage for attention would ignore all the critical business that occurs before what may end up being a pro forma vote on a carefully selected and highly vetted item. Consequently, for each chamber in our sample we collected the following information:

- whether the SOP is RCV or non-RCV as well as the precise method of casting a vote (electronic, voices, show of hands, rising in places, etc.),
- whether legislative actors can invoke a voting method that departs from the SOP (from RCV to non-RCV or non-RCV to RCV), and
- the legislative actors who can invoke a voting method that departs from the SOP.

Section A of the Online Appendix presents this information in Table OA.A1 for all 145 chambers.

Before discussing the results, we first evaluate the validity of our catalogue of voting rules compared with another recent source based on an expert survey: the Hug, Wegmann, and Wüest (2015) project. Of the 48 chambers where our studies overlap, our classifications agree for the vast majority (43).¹² For the five instances of disagreement, we found that our classification comported best with the actual practices reported in the plenary minutes as of 2018. We know that in one case of apparent disagreement (Sweden), voting procedures actually changed between the time of the expert survey and our cataloguing of cameral rules. In general, this exercise gives us confidence in the quality of our dataset.¹³

We now turn to the survey of 145 legislatures. Two empirical patterns are particularly noteworthy. First, for 105 (72%) of the legislatures, the SOP for most legislative business is a non-RCV method.¹⁴ Second, RCVs are the exclusive vote method in only 18 (12%)

of the legislatures.¹⁵ Thus, in 88% of the chambers, *non-RCVs* are either the SOP or they can be invoked by various legislative actors—making RCVs a sample of all votes taken.

These findings, by themselves, do not necessarily raise a serious issue for research on legislative politics. It is possible that the legislatures of most interest to scholars are largely those where RCVs are mandatory, or at least the SOP. To see if this is so, we counted the number of articles in the Social Sciences Citation Index that reference each legislative chamber. In Table OA.B1 (presented in Section B of the Online Appendix) we present a list that includes each chamber, the number of articles focused on it, and the share of total articles these represent.¹⁶ The results indicate that the RCV-selection concern is a potential issue for the overwhelming majority of research on legislative politics. The vast majority (81%) of all articles focus on chambers using non-RCV methods as the standard procedure.¹⁷ And, interestingly, the chambers where the selection issue is absent—those with mandatory RCVs—have not uniformly attracted scholarly attention. Indeed, a third of these were the focus of two or fewer publications during the 19-year period.

It is worth noting that the bulk of legislative research involves only a fraction of the 145 chambers catalogued here. The 30 most studied legislative chambers account for about 89% of all articles about legislative politics. Only four of these feature a mandatory RCV procedure, and they account for a small share (7%) of all articles that focus on this set of chambers. Similarly, chambers that feature an RCV method for the SOP, but allow non-RCVs, represent only six of the top 30 chambers and are seldom the focus of published research.¹⁸ In contrast, legislatures with non-RCV procedures as the SOP attract the lion's share of scholars' attention—around 82% of the articles focusing on the 30 most-studied chambers. In sum, the potential selection problems for RCV analysis apply to a large share of the legislatures that are the focus of scholarly research.

We therefore need to understand the rules for invoking RCVs—and the norms for their use—to determine the severity of any selection-bias issues. As we report in Table OA.A1 in the Online Appendix the rules vary considerably. In some instances, the procedural threshold to invoke an RCV is *quite* high. For eight chambers, a majority of those present or even a majority of total members need to agree to a roll call in order for one to be held. In an additional 10 chambers, at least 20% (but

¹² The two studies vary in that Hug, Wegmann, and Wüest (2015) focus on rules for votes on final passage, and we focus on the voting procedures for most legislative business. Also, Hug, Wegmann, and Wüest (2015) count votes reported by party bloc as “open,” while we do not consider these as roll calls.

¹³ We conducted a basic analysis that is modeled after the one presented in Hug, Wegmann, and Wüest (2015) and report the results in Section D of the Online Appendix. More specifically, we examined the relationship between a legislature's standard voting method and a number of potential determinants. In short, neither study gained much insight into what explains differences in procedure. For a more detailed discussion of the analysis, please see Section D.

¹⁴ This result is similar whether for legislatures in consolidated democracies or in other regimes. In consolidated democracies, defined as countries scoring 6 or higher on the Polity IV index for

the decade 2008–2017, the SOP is a non-RCV method for around 69% of the chambers.

¹⁵ For consolidated democracies, it is somewhat more common for RCVs to be the exclusive vote method. RCVs are mandatory in 17% of these legislatures.

¹⁶ For the years 1990–2018, we counted the number of articles with a topic related to each chamber. Details for the search algorithm and its results are found in Section B of the Online Appendix.

¹⁷ This result is identical if we exclusively focus on consolidated democracies.

¹⁸ They were featured in about 11% of the articles that cover the top 30.

TABLE 1. Voting Procedures and Practices for Most Legislative Business

Country	Chamber	RCV Threshold	Period	Total Votes	# of RCVs	% RCVs
Argentina	lower	10% of those present	2007	535	194	36.3
Australia	lower	2 members	2016–2019	1,530	469	30.7
Barbados	lower	1 member	2014–2019	2,080	3	0.1
Barbados	upper	1 member	2014–2019	1,118	10	0.9
Belgium	lower	8 members, speaker	2014–2019	25,417	2,670	10.5
European Parliament	lower	32 members	1999–2000	2,405	173	7.2
France	lower	speaker, 1 party, referral committee, government, “conference of presidents”	1998–2012	62,185	2,215	3.6
France	upper	speaker, 1 party, referral committee, government, 30 members	2006–2013	32,636	1,482	4.5
Guatemala	unicameral	6 members	2004–2011	9,584	107	1.1
Mexico	lower	6 members	2006–2007	702	123	17.5
Netherlands	lower	1 member	2008–2015	18,594	96	0.5
New Zealand	unicameral	1 member with speaker approval	2003–2015	8,291	162	2.0
Switzerland	lower	30 members	1995–2003	6,052	3,999	66.1
Uruguay	lower	1/3 of those present	2001–2005	6,046	25	0.4
Uruguay	upper	1 member with speaker approval ^a	2000–2004	6,070	17	0.3

Note: The entry for the Swiss lower chamber and Guatemala’s unicameral legislature reflect older standing orders. The rules have since been revised, and the current SOP in both chambers is an RCV (open) method (see Table OA.A1). For this exercise, we use the older rules because we have data on RCV prevalence for periods during which a non-RCV method was the SOP in these chambers. This enables us to examine the relationship between thresholds for invoking an RCV and the prevalence of such votes.

^a We determined this from the observed practice described in the plenary minutes. The chamber’s rules of procedure do not indicate how to invoke a roll-call vote.

less than 50%) of the chamber needs to approve a request for the use of roll-call procedures. High procedural thresholds are not the norm though. Indeed, 32 legislatures require only a request by a single legislator to invoke an RCV. The widespread use of low or modest thresholds may mitigate or even eliminate the selection problem, provided the low procedural obstacles to invoking RCVs generate a large and diverse sample of votes. We turn to that question next.

In Table 1, we present RCV prevalence information for 15 legislative chambers where the SOP is a non-RCV method and the thresholds for invoking an RCV are relatively low. Five of these chambers require the bare minimum—a single member of parliament—to

invoke roll-call voting. Several others require the support of a very small share of the legislature. For example, the 6-member threshold in Mexico’s lower chamber corresponds to slightly over 1% of the total membership. The highest threshold in our sample of cases is one-third of members in attendance, required in Uruguay’s lower house. Finally, some legislatures, such as the French Senate, do not define a preset number of members as a threshold, allowing a range of institutional actors (e.g., parties, committees, or the government) as well as regular members of parliament to request an RCV.

The table also presents information on the share of votes by roll call (i.e., the prevalence of RCVs) in these

legislatures. As described earlier, we successfully identified information on the universe of legislative votes for significant periods in 11 chambers where the publicly available documents (e.g., the plenary minutes) present sufficient detail of legislative proceedings for us to identify systematically votes of all sorts. To these, we add the relevant information for the four legislatures (the European Parliament, the lower chambers of Argentina and Mexico, and the Swiss lower chamber) presented in past studies (Carrubba et al. 2006; Crisp and Driscoll 2012; Hug 2010).¹⁹ This sample of legislatures—albeit a convenience sample—provides the largest comparison of RCV prevalence to date, allowing us to examine how low thresholds for invoking a roll call are related to their prevalence among all votes taken.

Despite the lack of major procedural obstacles across most of these legislatures, the average prevalence of RCVs is only 12%. What is more, even in legislatures where all final passage votes must be taken by RCV (Argentina, Belgium, Mexico, and Switzerland), it is still the case that less than half of votes are taken by roll call. Perhaps the most striking observation is that the five legislatures with the absolute minimum threshold (1 member) rarely vote by roll call: the average prevalence is *less than 1%*. The fact that RCVs tend to constitute such a small share of total votes, despite the ease of invoking the procedure, only increases our concern that many roll-call analyses may involve unrepresentative samples of legislative votes. Moreover, the threshold is *not* inversely related to prevalence among these chambers. The correlation between the share of votes taken by RCVs and the percentage of legislators required to request an RCV is 0.18 ($p = 0.51$).²⁰ Note that these results do not depend on any one legislature; with one exception, the correlation is statistically insignificant if we drop individual legislatures from the estimation. If we drop Uruguay's lower chamber, the correlation is 0.68 ($p < .01$)—a relationship that contradicts the current thinking.

Finally, it is important to note that the relatively high prevalence of RCVs in the Swiss lower house proved to be no insurance against selection problems. Recall that Hug (2010) is unique in that it compared the individual-level voting behavior of legislators on votes they knew to be recorded and on votes they expected to go unreported at the individual level. That comparison revealed a distinct selection bias—partisan blocs were discernibly less cohesive when using non-RCV voting methods.

Let us quickly repeat the conclusions we can draw from our overview. It seems very likely that in most legislatures the votes taken via roll call might be a biased sample of all votes. For most legislative business, roll call is the SOP for only around a quarter of

chambers throughout the world. In some of these cases the procedural obstacles for invoking a roll call can be quite extreme. Even where the threshold for invoking a roll call is quite low, this does not predict the prevalence of votes taken by that method.

These findings are concerning for the extensive, unquestioned use of roll-call voting records to calculate measures of individual ideological location, partisan homogeneity/discipline, or dimensionality of the policy space. Unless the selection process is benign, traditional RCV analyses likely produce biased results for a broad range of legislatures.

MODELING THE RCV SELECTION PROCESS

Roll-call votes are used extensively to estimate three legislative phenomena of interest: the location of legislator ideal points, party unity (or cohesion), and the dimensionality of policy contestation. If the RCV selection process is not entirely benign, RCV estimates of these three quantities could be misleading including being biased, imprecise, or otherwise inaccurate. The presence of these problems is very difficult to evaluate in a purely empirical manner. With one exception (Hug 2010), the legislative records of parliaments lack necessary information about individual legislative behavior on non-RCVs. Less direct empirical comparisons such as those comparing attributes (e.g., policy area) of RCVs and non-RCVs may indicate potential sources of problems (Carrubba, Gabel, and Hug 2008; Crisp and Driscoll 2012), but any conclusions about the nature of the problem (e.g., direction and magnitude in estimation bias) require theoretical interpretations of the data-generating process.

Theoretically, RCVs can have one of two relationships with the underlying universe of votes. One option is that making a vote public record causes legislators to vote differently than they would have if the RCV had not been requested. The other option is that RCVs simply make public specific voting activity, without changing how individuals vote. Previous work demonstrates the dramatic effect on RCV samples and their properties if calling a vote is associated with changing votes because of party disciplining (Carrubba, Gabel, and Hug 2008). What we still do not know is whether RCVs provide a problematic understanding of legislative behavior if they do not alter voting behavior. This question is particularly important given that scholars have questioned whether RCVs are actually used to influence voting behavior (Carey 2009; Thierse 2016). Further, this is the most favorable case for RCVs to be an innocuous sample of votes.

What follows is a simple theoretical model designed to characterize a position-taking story in which RCVs are requested to highlight certain votes (but not to change an individual legislator's vote decision). We then use that model to create a simulation exercise in which we demonstrate how ideal point estimates, unity scores, and estimates of dimensionality are influenced by this selection process.

¹⁹ Sieberer et al. (Forthcoming) report the RCV prevalence for final passage votes in the German Bundestag as about 5%. Because they only reported prevalence for final passage votes, we do not include them in our comparison.

²⁰ For the French upper chamber, we estimated a 5% threshold, as the smallest political party consisted of between 4% and 6% of the legislature during the period under study.

A POSITION-TAKING MODEL OF ROLL-CALL REQUESTS

If RCVs are intended, in some fashion, to record party positions on specific issues, we will refer to this motivation as “position-taking.” A variety of position-taking motivations have been posited, including signaling to the public your support for a popular policy, the opposing party’s support for an unpopular policy, your party’s unity around some issue, and/or the opposing party’s disunity over the issue (Carey 2007, 2009; Crisp and Driscoll 2012; Kreppel 2002; Saalfeld 1995, 548). For this simulation, we model parties as caring about relative demonstrated unity. That is, if a vote is made public, each party wants to show that it is more unified about some issue than the other party is. We focus on preferences over unity because it captures a selection mechanism posited in the literature in which those requesting roll calls have preferences over exactly the output of interest we are trying to study. We model those preferences as being over relative unity since parties simultaneously have an incentive to try to show well while embarrassing the other party. Note that we use the term unity for voting division on a single motion and cohesion as the average unity across a set of votes. Assume there are two parties $i \in \{\ell, r\}$ competing over a unidimensional policy space, \mathbb{R}^1 .²¹ Each party consists of a set of party members (backbenchers) and a party leader, I_i , whose ideal policy is located at the median of his or her party. Without loss of generality, we fix the party leaders’ ideal policies at $I_r = 1$ and $I_\ell = 0$.

For an N -person legislature, we define the set of backbench legislators’ ideal points as being drawn from a uniform distribution around their party leader’s ideal point, such that $x_{ij} \sim U[I_i - d, I_i + d]$ for $j \in \{1, \dots, N\}$ and $i \in \{\ell, r\}$. The parameter d measures how potentially varied that the preferences of the parties’ members are, and it will be referred to as *potential party heterogeneity* throughout.²² For $d \leq 0.5$ party member policy ideal points will not overlap. As d increases, the parties become less differentiated. Finally, let $\phi \in (0, 1)$ denote the proportion of the legislative chamber controlled by the right party, such that there are a total of $n_r = (\phi * N) - 1$ backbenchers in the right party and $n_\ell = [(1 - \phi) * N] - 1$ backbenchers in the left party. For $\phi \leq 0.5$ the left party is the majority party. As ϕ increases, the right party becomes increasingly dominant.

The structure of the model proceeds as follows. Parties are competing over potential motions made in the one-dimensional policy space. A *motion* will consist of a pair $m = \{b, sq\}$, where b is the location of the proposal and sq is the location of the status quo. First, party leaders must choose whether to initiate a motion by proposing (b) or not, where not making a proposal preserves the status quo (sq). Motions with $b < sq$ ($b > sq$) are treated as possible proposals by the left (right) party. The agenda that determines possible

motions is described in detail below. Critically, it is an agenda-setting process with a minimum of control.

If a proposal is made, we allow any member of the chamber to initiate a roll-call vote, where $R_i = \{RCV, -RCV\}$. We make this assumption to ensure maximal ease of making an RCV request—something our empirical work indicates is reasonably common. Requesting a roll-call vote entails a cost, $k \geq 0$. These costs reflect the potential time and imposition of recording the votes. The smaller the value of k , the less the inconvenience. Once a bill proposal has been made and a voting procedure determined, individual votes are cast for either the proposal or status quo: $v_j = \{b, sq\}$. The outcome of the vote $z \in \{b, sq\}$ is the majority winner.

Preferences over relative revealed unity for a given vote are based upon simple preferences over each party’s unity score. The unity score of the proposing (opposition) party is defined as the proportion of that party voting in favor of (against) the proposal. A legislator’s (leader or backbencher) relative unity score is the difference between the two parties’ unity scores.

Formally, a legislator’s relative unity score for a bill their party supports is $V^*c_i = \frac{B_i(b) - B_{-i}(sq)}{n_i - n_{-i}}$, where $V = \{0, 1\}$ is an indicator function equaling 1 only if an RCV is requested and $B_i(v)$ is the number of legislators in party i voting for policy v . If a legislator’s party does not support the proposal, his or her relative unity score is $V^*c_i = \frac{B_i(sq) - B_{-i}(b)}{n_i - n_{-i}}$. Legislators voting against their party’s majority position receive a zero payoff from the parties’ relative unity.²³ Legislators will have an incentive to request RCVs when the relative unity advantage, c_i is sufficient to outweigh the cost of making the request, k .

Altogether, payoffs are as follows. As described above, backbenchers and party leaders care about both the policy outcome and the relative revealed unity (if an RCV is requested) for a given vote. Because we want to model RCV requests as highlighting, not changing, voting behavior, we assume legislators’ incentives to vote sincerely dominate strategic incentives to change their vote to maximize unity benefits.²⁴ Thus, once a proposal is up for a vote, legislators always vote sincerely based on minimizing the Euclidean distance between their own policy ideal point and the two policy alternatives. Prior to that, legislators decide whether to request an RCV conditional on common knowledge over

²³ Because members supporting a party’s majority position have identical utility from relative revealed unity, the results of this simulation are robust to assuming only party leaders can request RCVs, or a group of legislators equal to 50% + 1 legislators of the minority party $\lceil \frac{N}{2}(1 - \phi) + 1 \rceil$ is required in order to request an RCV. Unity preferences for party members voting against their party’s majority position are set to zero for convenience. They could equivalently have matching preferences to the rest of their party or to the opposition party’s preferences and nothing changes.

²⁴ We already know strategic voting will exacerbate estimation problems: “In situations where strategic voting is widespread—where voting is motivated by factors unrelated to one’s preferences regarding the content of policies under consideration—one cannot assume that the variation in the data can be reasonably interpreted as reflecting the preferences of individual legislators” (Carroll and Poole 2014, 116).

²¹ Later, we introduce a second policy dimension in order to examine dimensionality reduction.

²² The results hold when parties have different levels of potential heterogeneity (see Online Appendix E.2.3).

the distribution of ideal points and anticipated sincere voting behavior. Finally, in deciding when to make a policy proposal, leaders maximize their utility over policy outcomes and relative revealed unity payoffs.

We parameterize how much they value the policy outcome compared with demonstrating high relative cohesion by the following equation:

$$u(I_i, z) = -\alpha|I_i - z| + V^*c_i, \quad (1)$$

where $\alpha > 0$. The larger the α , the more they value getting policy victories over demonstrating relative party unity.

The RCV request incentives are simple and intuitive. To demonstrate this fact, assume party leaders do not care about policy outcomes, $\alpha = 0$. If RCVs are costless, $k = 0$, all votes will be roll-called; because $u(I_i, b) = V^*c_i(b)$ and $c_i(b) = -c_{-i}(b)$, either $u(I_i, b) \geq 0 \geq u(I_{-i}, b)$ or $u(I_{-i}, b) \geq 0 \geq u(I_i, b)$. Intuitively, the unity advantage on any given vote is zero-sum. Members of either the right party or the left party benefit. Sensibly, one of them will always want to call an RCV if there is no downside. Now suppose requests are costly, $k > 0$. Once this is the case, not all votes will receive a roll call. The unity advantage must be sufficiently large to merit the request. Further, in the resulting sample, the proposing party will always have higher revealed unity than the non-proposing party does because $u(I_i, b) = c_i(b) - k \geq 0$ must hold for the proposing party for a proposal to be made. Again, this makes perfect sense. A party that only cares about relative unity is not going to make a proposal that works to its disadvantage.

Based on this model, we can generate results for how any possible motion will be treated. We can characterize whether a proposal will be made, if that proposal will receive a roll call, how legislators will vote, and whether the proposal will pass. We turn to simulations of the analytic model because we are not interested in individual votes so much as aggregate properties compiled across an entire legislative agenda for a wide sample of possible legislatures. The simulation will consist of a set of natural legislatures with fixed membership (between elections). For each legislature, we randomly draw a value for the parameters $d \sim U[0.5, 2]$ and $\phi \sim U[0.2, 0.8]$. As mentioned above, these values ensure a wide range of possible legislative chambers can be generated, everything from a chamber with no party overlap to one in which the parties are relatively indistinguishable and from one with a very dominant right party to one with a very dominant left party. Based on these two values, we draw a set of legislators to populate a 100-member chamber. Unless otherwise stated, we set $\alpha = 0.5$ and $k = 0.1$. An $\alpha = 0.5$ reduces the weight on policy relative to unity. A $k = 0.1$ is substantively equivalent to requiring a relative unity edge of at least 10% to merit making an RCV request. On average, these parameter values correspond to natural legislatures in which 28–36% of all votes are taken by roll call.²⁵ Given the empirical findings

presented in the previous section, this means the simulated samples of roll-call votes are far more complete than those we typically observe in the real world. In the following analyses where $k \leq 0.2$, the issues we highlight are the lower bounds of the problems that can arise in applied empirical analyses of roll-call votes.

Next, we create a possible legislative agenda for this legislature by drawing 1,000 motions. A motion consists of a status quo location $sq \sim U[0, 1]$ and possible proposal $b \sim U[0, 1]$. As previously mentioned, motions where $b < sq$ are proposals that the left party leader ($I_\ell = 0$) can choose to make. The rest are proposals the right party leader ($I_r = 1$) can choose to make. This way of characterizing the legislative process is tantamount to weak agenda control by the proposing party. Each party can avoid making self-defeating proposals, but they cannot control what comes up for consideration. It has the advantage of allowing the stochastic process to capture out-of-model constraints on agenda control by party leadership. It also offers a way of populating potential proposals by the minority party. We take this approach to ensure that the RCV selection process, not agenda control, is driving the findings of our model.²⁶

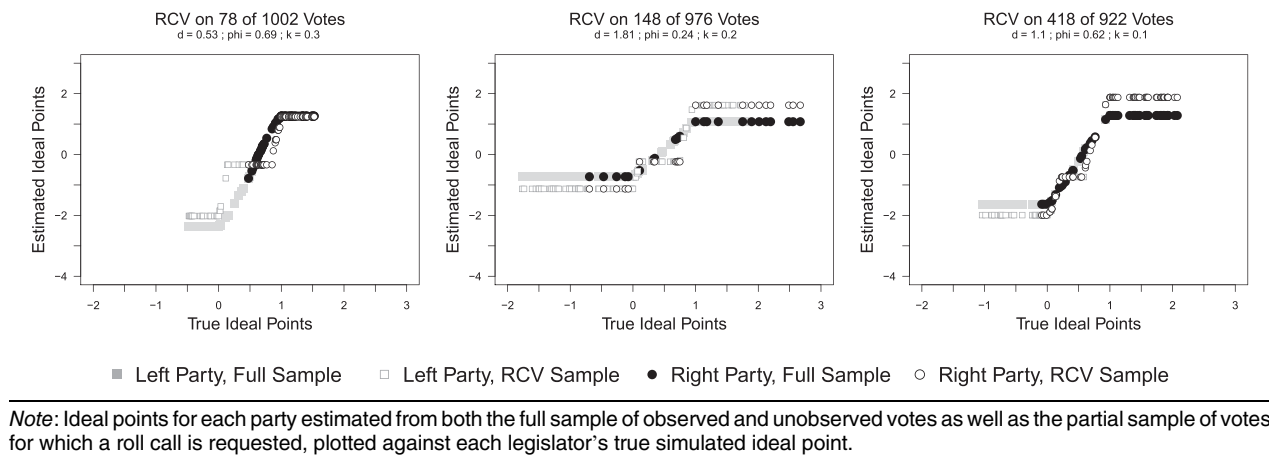
By simulating the roll-call vote selection process, we have the advantage of knowing how individual legislators vote when they are *not* voting via roll call—something only Hug (2010) has seen in the “real world.” Using these simulations, we conduct standard empirical analyses on the full sample of observed and unobserved votes as well as on the sample of roll call votes. In doing so, we are able to characterize and compare results on our three key quantities of interest: legislator ideal points, cohesion scores, and policy dimensionality. Figure 1 provides three illustrations of the relationship among ideal points estimated using (1) all votes and (2) only those taken by roll call, and compares those with the true policy ideal point locations of the simulated legislators.²⁷ The left panel corresponds to a chamber with very infrequent use of roll call votes (8% of total votes), the middle panel to one with moderate use (15%), and the rightmost panel one with extensive use (45%). In each panel, the x -axis measures the location of the true ideal points and the y -axis the location of the estimated ideal points from ideal point estimation models that use either the sample of all votes or only those votes that receive a roll-call request. The filled grey squares and black circles correspond to the ideal point estimates from the full sample compared with the legislators’ true (simulated) ideal points. The open grey squares and black circles correspond to the ideal point estimates from the RCV sample compared with the legislators’ true (simulated) ideal points.

If the estimated ideal points map to the actual ideal points, the rank orders will be preserved: the larger the true ideal point, the larger the estimated ideal point,

²⁵ See Online Appendix E.1.2 for heat maps demonstrating how RCV selection varies in α and k .

²⁶ Online Appendices E.2.6, E.3.4, and E.4.4 show the robustness of the results to a completely exogenous agenda.

²⁷ All ideal points are estimated with the expectation-maximization techniques from Imai, Lo, and Olmstead (2016) and implemented in the corresponding package `emirt`.

FIGURE 1. Estimates of Ideal Points from Full and RCV Samples

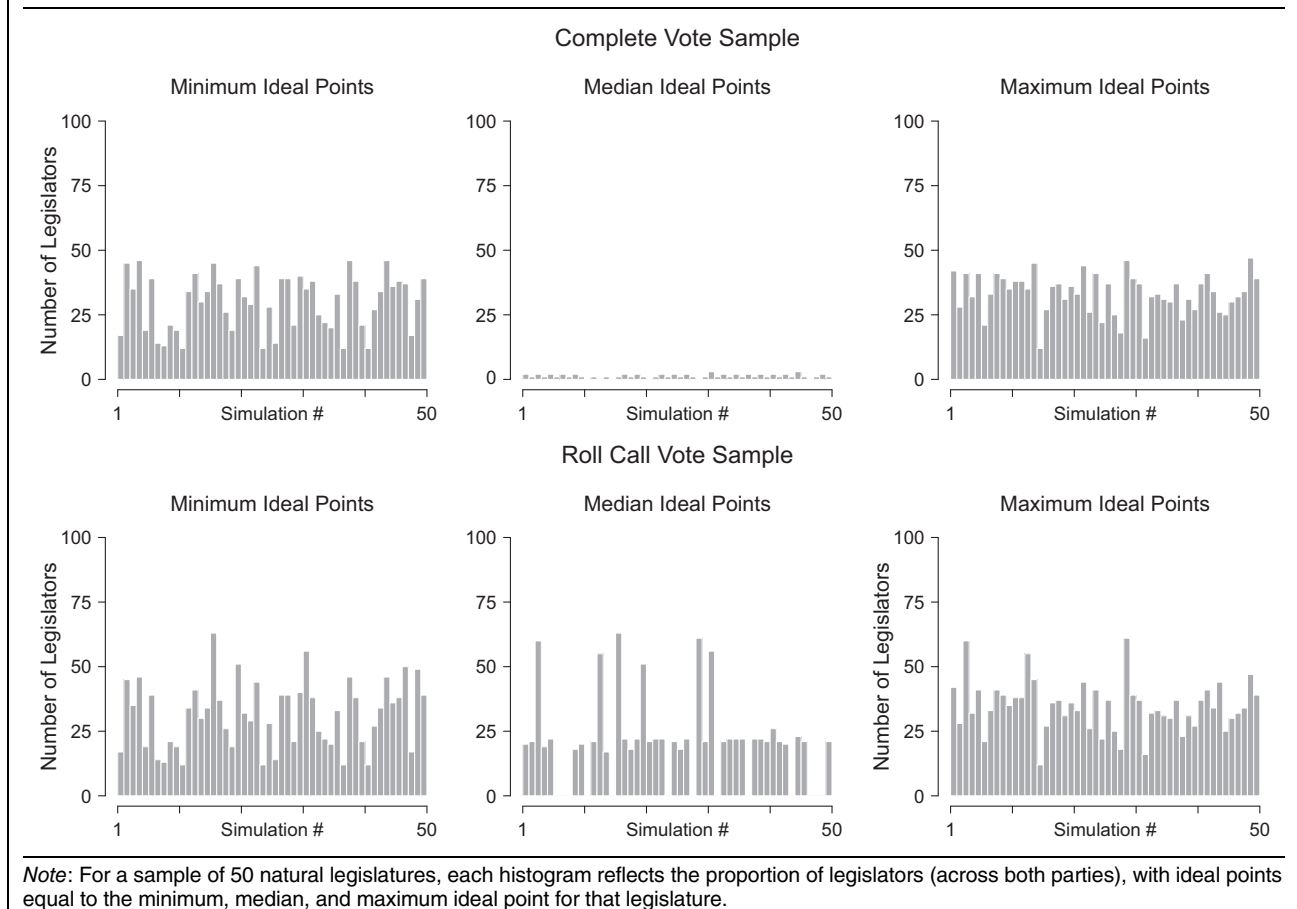
and therefore the points will be monotonically increasing on the diagonal. As can be seen, with the exception of the ideologically extreme candidates, the ideal point estimation on the full sample of votes is able to recover the rank ordering of legislator ideology quite well. As is standard for these sorts of models, extreme members are difficult to distinguish from one another due to a small number of discriminating votes in the tails of the distribution. The ideal point estimates based just on RCVs fair more poorly. They struggle not only to differentiate extreme legislators but also to discriminate among the more moderate, or interior, legislators: that they lie on a horizontal line indicates that while the true ideal points (x -axis) are different, the estimator (y -axis) does not pick up any variation. While the extent of the problem increases as the RCV sample shrinks, it is present in all three contexts. The key is that legislators choose not to request roll calls on votes that divide moderates because those votes do not yield a sufficient unity advantage on that vote to one party or the other.

To demonstrate the consistency of this pattern across possible natural legislatures, Figure 2 depicts the number of legislators in each natural legislature for which our ideal point estimates would yield the identical minimum ideal point (left panels), the identical median ideal point (middle panels), and the identical maximum ideal point (right panels). Estimates based on the full sample are presented in the top row, and estimates based on just roll calls are presented in the bottom row. In the full sample, identical estimates occur on the ideological extremes, but they are extremely rare in the middle. For this randomly drawn sample of 50 natural legislatures, the largest number of non-extreme legislators receiving identical ideal point estimates is three. Conversely, in the RCV sample a substantial proportion of moderate legislators—in a significant number of cases even the majority—are consistently identified as having identical ideal points.²⁸

²⁸ The simulations assume perfect spatial voting (i.e., no vote errors by legislators). Spatial voting that is perfect or near perfect (low levels of voting errors) is a reasonable assumption for legislatures

Turning to party unity, in Figures 3a–d we consider how cohesion scores (averaged unity scores across a full legislative agenda) generated from RCVs compare with cohesion scores generated from unobserved votes. The sample is based on 1,000 votes per natural legislature for 50 different natural legislatures. The more consistently that the cohesion scores that are generated from observed roll-call votes reflect the cohesion scores that are based on unobserved votes taken by some other means (e.g., show of hands), the safer it is to assume that a measure of cohesion based on an RCV sample reflects cohesion on all votes. The figures show cohesion scores based on RCVs (dashed lines) and unobserved votes (solid lines) for the left and right parties as potential party heterogeneity, d , and size of the right party, ϕ , vary. The black solid and dashed lines are for proposals made by the left party and the grey lines are for proposals made by the right party. As can be seen, the proposing party's cohesion scores based on RCV votes are systematically higher than those that are based on unobserved votes, whereas the opposite is true for the non-proposing party. This result holds across possible values of potential party heterogeneity and majority party size. The logic behind the result is based on the intuition discussed previously. The proposing party will favor moving forward on votes that benefit them, where the benefit is based in part on their relative unity on the vote if they anticipate the vote being roll-called. Thus, we should expect the proposing party to look more unified on roll calls than they do on

characterized by a multiparty system of disciplined parties (Rosenthal and Voeten 2004). Such legislatures are common in our sample. As shown with other forms of RCV selection problems, greater levels of voting error mitigate the problem of distinguishing moderate legislators (Clinton 2012; Hirsch 2011; Snyder and Groseclose 2000). However, the problem with moderate legislators holds in our simulations when we introduce small or moderate levels of vote error, using the range of vote error in Clinton (2012) and Hirsch (2011). While the point estimates of moderate legislators are no longer identical when there are high levels of voting error, we still cannot distinguish them from one another due to a lack of precision resulting from increased uncertainty.

FIGURE 2. Collocated Ideal Point Estimates in Full and RCV Samples

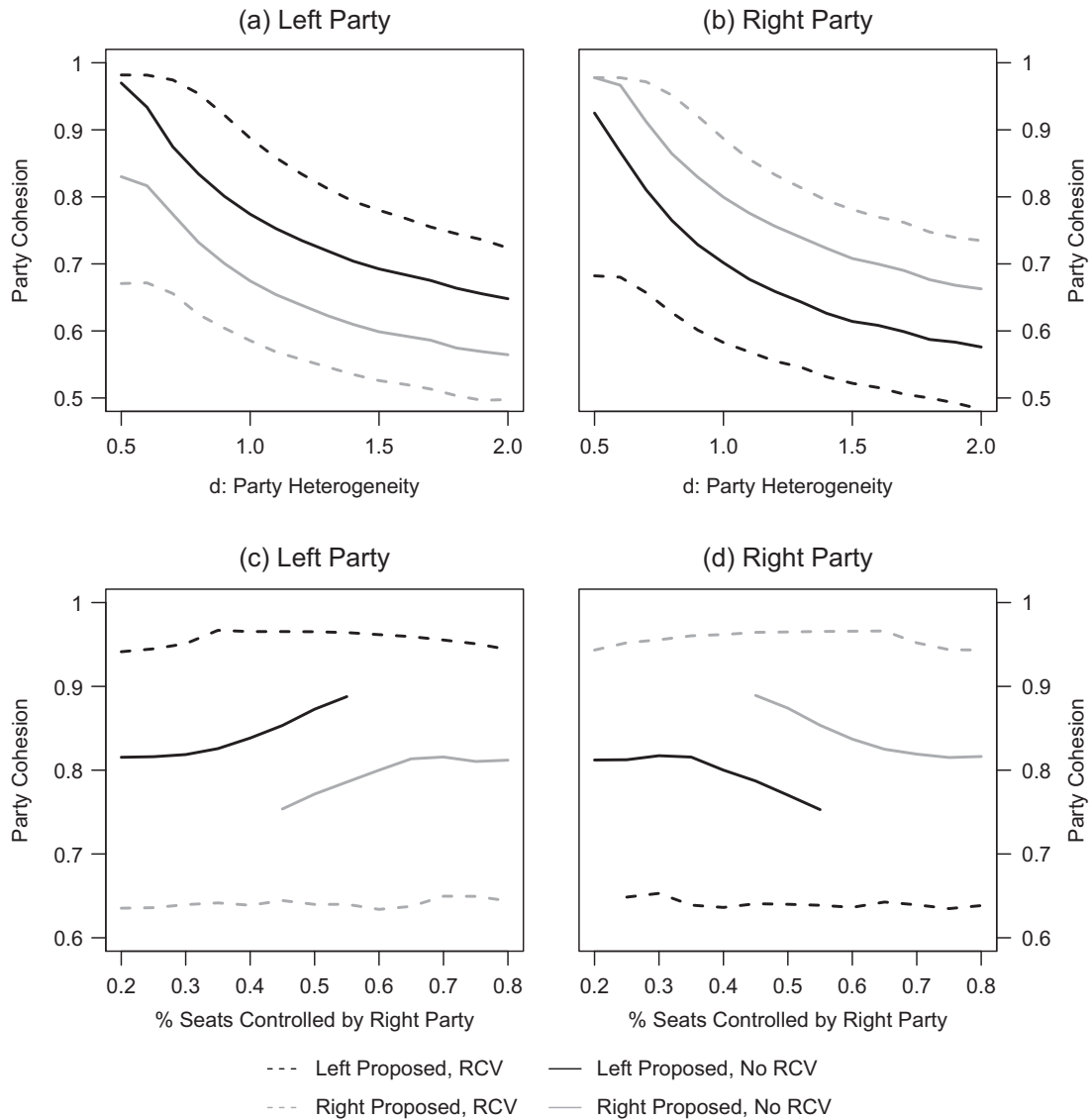
unobserved votes and the non-proposing party to look worse on RCVs than on unobserved votes. The result is that we cannot assume that cohesion scores generated from RCVs generalize to cohesion scores based on what in the “real world” would be unobserved votes.

Of course, one might suppose that these systematic differences wash out in aggregate (Carey 2009, 48; Poole and Rosenthal 2007). Because of varied agendas and types of votes, on average the RCV cohesion scores derived from a legislative session should accurately map to the non-RCV cohesion scores for that same session. Figure 4 addresses this claim. We simulate 50 randomly drawn natural legislatures and take the difference between the cohesion score based on the unobserved sample and the RCV vote sample. We assume the cost of requesting a roll call vote is relatively low ($k = 0.2$), creating legislative dockets with an average of 479 total votes of which 32% are taken via roll call. Figure 4 shows the distribution of these differences across natural legislatures where positive (negative) values here correspond to parties being on average more (less) cohesive in the unobserved sample. As the density plot demonstrates, the distribution of differences by natural legislature is frequently quite substantial. The average absolute difference for both parties is 8.2%, meaning by looking at roll-call votes alone we would expect on average 8 more/fewer per

100 to vote with their party than actually do in the full sample of votes. In 39% of the simulated natural legislatures, the difference between RCV and non-RCV cohesion scores is at least 10% for one of the two parties. Critically, cohesion is not being systematically over- or underestimated. For approximately half of the time, the RCVs are overestimating cohesion, and for half of the time they are underestimating cohesion. For this sample of simulated legislatures, the cohesion we observe is greater than that which we do not observe for 60% of the legislatures. Thus, the only thing we are confident of is that they are frequently unrepresentative.²⁹ Critically, this problem emerges here for a set of natural legislatures where a relatively high percentage of votes (32%) are taken via roll call. In Online Appendix E.3.3, we replicate this simulation for a set of natural legislatures in which only 14.5% of total votes are taken via roll call and show that the difference in cohesion for observed and unobserved votes exceeds 10% for over half of the sample.

²⁹ Note that these differences are not simply a function of unusually distributed legislative chambers. For the full sample of chambers, the mean difference is 1.8% and the standard deviation is 0.09. For a sample of chambers in which potential party heterogeneity is more moderate, $d \in [0.75, 1.5]$, and the parties are more balanced, $\phi \in [0.35, 0.65]$, the mean difference is 1.1%, the absolute difference is 8.9%, and the standard deviations are 0.10 and 0.05, respectively.

FIGURE 3. Party Cohesion in Full and RCV Samples



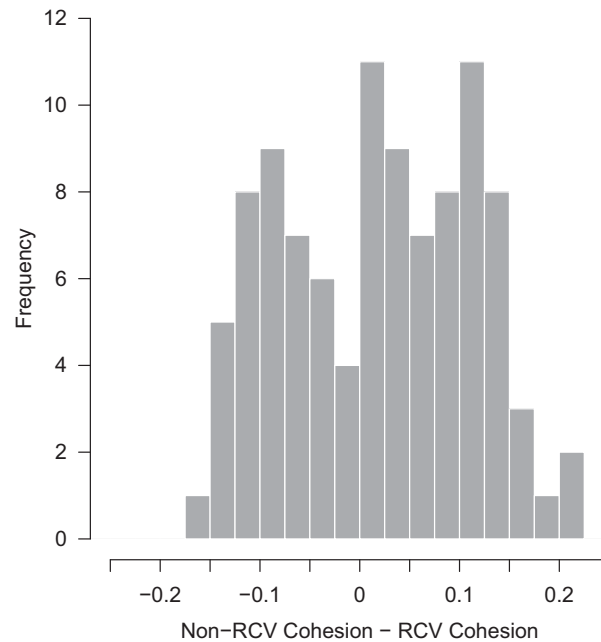
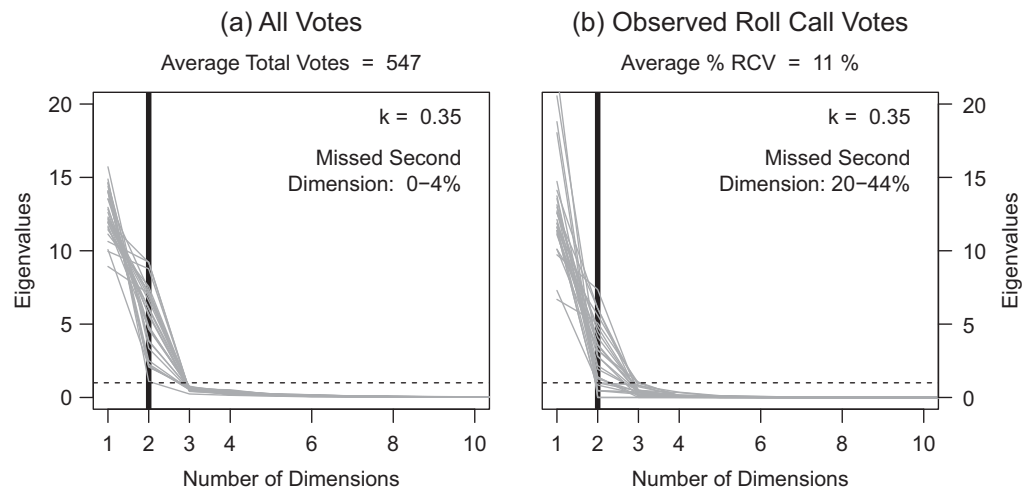
Note: Each figure shows for the left and right party, respectively, across values of party heterogeneity (d) and seats controlled by the right party (ϕ), the differences in cohesion across both unobserved (non-RCV) votes and observed (RCV) votes in, divided by which party made the proposal.

Simply put, one cannot assume that in any given natural legislature RCV cohesion scores will be representative of non-RCV cohesion scores.

Finally, to examine the effect of the roll-call vote sample on the ability to discern the dimensionality of legislative preferences, we can compare the results of ideal point estimation models on the full simulated sample of votes with those on the sample of only votes where an RCV was invoked. For this analysis, we generate natural legislatures with two distinct policy dimensions, meaning each legislator has an ideal policy position in each dimension. Further, we allow the degree of party heterogeneity (d) to vary across policy dimensions. While party heterogeneity within the first and second policy dimensions (d_1 and d_2) is constant across parties, we assume the degrees of heterogeneity are random draws from the uniform distribution:

$d_k \sim U[0.5, 2]$ for $k \in \{1, 2\}$. Thus, legislators' ideal points on each dimension remain centered around their party leader's ideal point I_i , but they are randomly drawn from the distributions $x_{ijk} \sim [I_i - d_k, I_i + d_k]$, where $i \in \{\ell, r\}$ indicates legislator j 's party affiliation and $k \in \{1, 2\}$ indicates the policy dimension. For each natural legislature, we generate 500 motions per dimension. We set $k = 0.35$ and $\alpha = 0.5$, which for our sample will induce an average of 547 total votes and 62 roll-call votes (i.e., approximately 11% of the sample will be roll-called on average) for each natural legislature.³⁰

³⁰ We set $k = 0.35$ to generate an RCV prevalence (i.e., about 10%) similar to what we observe in Table 1. In Online Appendix E.4.1, we demonstrate the robustness of the results.

FIGURE 4. Difference in Party Cohesion on Unobserved (Non-RCV) and Observed (RCV) Votes**FIGURE 5. Recovering Multiple Preference Dimensions with Full and RCV Samples**

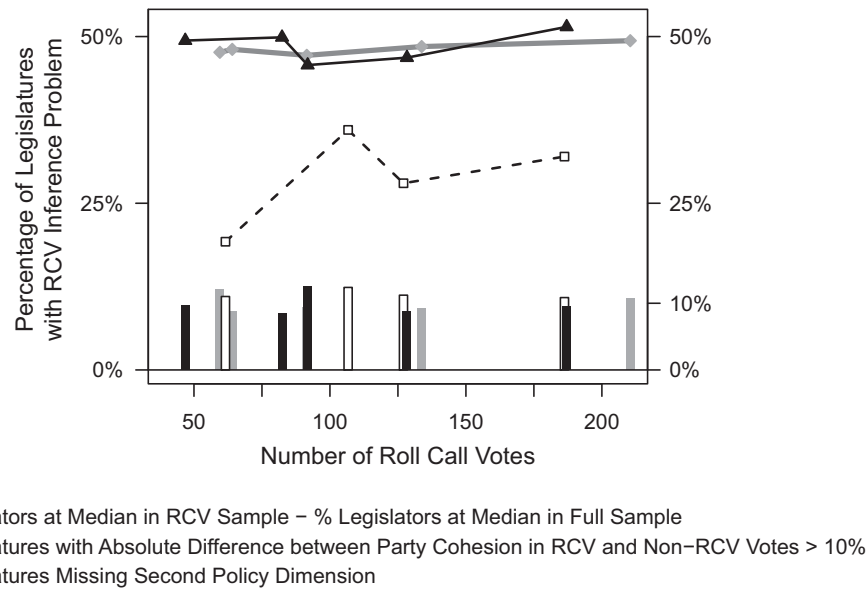
Note: For the samples of all votes (a) and subsample of only observed roll-call votes (b), each scree plot shows the eigenvalues from the estimation of Poole and Rosenthal W-Nominate scores.

The central result from this exercise is illustrated in the pair of scree plots in Figure 5. For each of the 25 simulated natural legislatures, Poole and Rosenthal W-Nominate scores are estimated for the $n = 100$ legislators, once using the full sample of votes and once using just the subsample of votes strategically selected to be done by RCV.³¹ Thus, we estimate two models for each of the 25 natural

legislatures. To examine whether the sample of RCVs is able to recognize the two dimensions of policy preferences as well as the full sample of votes does, we can compare the eigenvalues across the two samples. Two metrics have been traditionally employed for evaluating the number of dimensions based on these eigenvalues.³²

³¹ For simulations examining dimensionality, all analyses are conducted using the R package *wnominate* (Poole et al. 2011).

³² These approaches are subject to some criticism for their subjectivity (Aldrich, Montgomery, and Sparks 2014). However, since they have been commonly used, we use them for our exercise here.

FIGURE 6. Effects of Increasing the Number of Roll-call Votes

Note: Each line depicts how increasing the number of roll-call votes while holding prevalence fairly constant affects the average percentage of (a) additional legislators with collocated ideal points at the median under the RCV sample compared with the full sample of votes (solid black line), (b) natural legislatures with observed (RCV) cohesion $\pm 10\%$ different from unobserved cohesion (grey line), and (c) natural legislatures for which our estimates would not pick up a second policy dimension (dashed black line). The vertical bars across the bottom indicate the prevalence of roll-call votes for each simulation, ranging from 9% to 12% on average.

The first approach is to include only the number of dimensions prior to (i.e., to the left of) where we observe an “elbow” (Cattell 1966; Poole 2014). As can be seen in Figure 5a for the sample of all votes, there is a clear elbow and consistent leveling off of eigenvalues at the third dimension, indicating that the model is accurately picking up on the two-dimensional policy space of the simulated data in all but one of the 25 natural legislatures. The alternative approach proposed by Kaiser (1960) uses an “eigenvalue-greater-than-one” rule. Following this heuristic would lead us to correctly infer the existence of a two-dimensional policy space in each of the 25 natural legislatures using the full sample of votes.

Figure 5b presents the same information for the RCV subsample. As can be seen, this subsample less consistently identifies a second dimension. Using the “elbow” criteria, the second dimension would be missed in 10 of the 25 (or 40% of the) natural legislatures, as compared with only one of the 25 (4%) with the full sample of votes. Using the eigenvalue-greater-than-one criteria, 20% of the natural legislatures—or five of the 25—would miss the second dimension.³³ This second approach appears to be a lower bound, given the concentration of values just above one and the substantial leveling out of the eigenvalues that occurs after for an additional five of the simulated natural legislatures.

Note that losing the second dimension does not mean that we are not observing votes on motions on the second dimension. In this simulation, we observe a nearly even split of motions and votes on each dimension in the full sample. Of the, on average, 547 total motions in each natural legislature, an average of 243 votes (50%) are from the first dimension and 243 votes (50%) are from the second dimension. If we assign as the first dimension the one on which parties are more heterogeneous (i.e., $d_1 > d_2$), there is nothing systematic to suggest that party leaders propose more frequently on one dimension versus the other. However, this same roughly equal distribution of motions across dimensions is not preserved for the subsample of votes receiving a roll-call request. In the roll-call vote sample, only 42% of all roll-called votes come from motions on the first (more heterogeneous) policy dimension.

This suggests one possible reason RCVs can cover a variety of substantive topics and yet common methods only estimate one to two dimensions. Legislators are not requesting roll-call votes on the votes that would allow our estimation strategies to tease out the existence of the second dimension. That is, the votes that remain unobserved are the ones that distinguish the ordering of the ideal points in the second dimension from that of the first.

CONNECTING THE MODEL TO THE EMPIRICS

All of this analysis suggests strong reason for caution over using RCV samples to draw general conclusions

³³ In Online Appendix E.4.1, we replicate this analysis for a variety of values of k , demonstrating that for roll-call vote samples that contain 8% of the total votes, the likelihood of the analyses missing a second dimension can increase to well over 50%.

about legislative behavior and conflict. However, one could reasonably question to what degree it speaks to actual legislatures we observe in the world. In fact, we have designed the simulations to not only map closely to our empirical referents but also to be somewhat generous compared to the true data.

For our ideal point estimation results we choose values for the cost of requesting votes, k , and the relative importance of policy outcomes, α , to ensure that our simulation spans a wide range of actual prevalence. While increasing prevalence does improve the estimation properties with respect to ideal points, at no point is prevalence high enough to eliminate the concern. Increasing prevalence of roll-call votes to 64%—just 2% shy of the maximum prevalence of roll-call votes we observe empirically for Switzerland—results in 10.3% of ideal points being collocated at the median (i.e., we cannot discriminate among them). This number is clearly an improvement, but it is still substantially higher than the, on average, 2.8% that are collocated at the median when estimated using all votes. Similarly, when we increase the prevalence of roll-call votes to 65%, our estimates of cohesion from the roll-call vote sample would differ from the unobserved sample by $\pm 5\%$ for 41% of natural legislatures and $\pm 10\%$ for 4% of natural legislatures.³⁴ While this is a nontrivial improvement on what we observed in Figure 4, the issue we highlight about inferring cohesion on unobserved behavior from only observed voting behavior remains. The dimensionality estimates improve most quickly with increased prevalence. Once the roll-call vote sample reaches 25% of all votes, estimations based on full and RCV samples start to converge. Models based on the RCV sample correctly identify the second policy dimension for between 80% and 96% of all natural legislatures, whereas the models based on the full sample do so for 96% of natural legislatures.³⁵ This constitutes a dramatic improvement for those few legislatures in our sample that reach this threshold of prevalence.

On a final, related note, the results of the simulations also speak back to a central point made in the empirical analysis emphasizing the importance of RCV prevalence rather than absolute number. Our simulations reinforce this point. What matters for the estimates at hand is the percentage of total votes taken by roll call, not the size of the RCV sample itself. Holding the prevalence of roll call votes roughly constant at an average of around 10% of the total votes, we show in Figure 6 that increasing the quantity of roll-call votes (x -axis) does not improve our ability to (1) discern among the ideology of moderate legislators (the solid black line indicates the number of additional legislators collocated at the median in the RCV sample compared with the full sample of votes); (2) draw inferences about

cohesion in observed votes (the grey line corresponds to the percentage difference in cohesion between the observed and unobserved sample of votes); and (3) correctly identify policy dimensions (the dashed line reflects the percentage of natural legislatures in which RCV samples would lead us to miss the second policy dimension). For each of these quantities, if the lines converged on zero as the sample size increased, it would indicate that, all else being equal, increasing the sample of RCVs is a viable solution to these issues. However, this is clearly not the pattern we observe.

CONCLUSION

Though roll-call vote data is central to much of the current empirical research being done in legislative studies, relatively little attention has been given to empirically or theoretically understanding the data-generating process. To help fill this void, first, we provided a systematic cross-national cataloging of voting rules from national legislatures. This demonstrated the breadth of the sample-bias issues that might possibly confront scholars of legislative studies around the world. We showed that SOPs for voting—and the rules for setting aside those procedures for an alternative one—vary widely. Moreover, we found that there is no simple relationship between the endogenous rules for selecting a vote method and the relative frequency with which roll-call procedures are employed (RCV prevalence). Further, both our empirical and theoretical analyses indicate that the resulting vote samples are likely unrepresentative of general legislative voting. The potential bias is particularly problematic in the legislatures that receive the most scholarly attention. Nearly 80% of articles that are written about legislative politics focus on cases where RCVs are not the SOP.

We also provide a rigorous theoretical treatment of the position-taking explanation for roll-call selection. In this modeling exercise we examined what happens to three key quantities of interest—individual legislators' ideal points, party cohesion, and policy dimensionality—if RCV requests are made based on an incentive to highlight voting behavior on particular motions. Critically, we make the selection process as benign as possible. We minimize party agenda-setting powers, we require only a single legislative member to invoke an RCV, and we do not allow for strategic voting based on the decision to request a roll call. Using a set of extensive simulations, we demonstrate the limits of using RCV samples for these estimates. We show that this sample will struggle to differentiate among moderate members of the legislature, that cohesion scores based on RCV samples will not replicate ones based on unobserved votes (and the differences will both be large and in an unpredictable direction), and RCV samples will be biased towards missing higher levels of dimensionality. All of these results arise from a simple strategic incentive. Party members (and leaders) have an incentive to request RCVs when their party is going to have a noticeably higher level of unity than the opposition. These findings hold for RCV samples that

³⁴ In Online Appendix E.3.3, we repeat the analyses of cohesion, showing how differences between observed and unobserved cohesion are decreasing in the prevalence of roll-call votes but persist even when roll-call votes are highly prevalent.

³⁵ In Online Appendix E.4.1, we demonstrate how increased prevalence of roll-call votes affects the proportion of natural legislatures for which the two policy dimensions are correctly identified.

map closely to the prevalence of RCVs found empirically in the sample identified in Table 1. Thus, the results are not simply speculative. They map to real world conditions.

This paper highlights the importance of understanding the rules governing how a roll-call vote can be invoked and the effects that those rules have on the voting record. One of the most prominent areas of research in legislative institutions has focused on understanding how formal institutions affect legislative outcomes. Scholars have examined how organizational structure (e.g., committee systems) and specific procedural rules—even some quite obscure ones—offer strategic opportunities to shape the character and final disposition of legislative proposals. Our study identifies a largely overlooked procedural rule that could have a significant influence on how we understand legislative politics. Because legislators may change their votes depending on whether the results are public, the selection of votes for roll call can have powerful effects on outcomes.³⁶ The study of the design and use of legislative voting procedures offers fruitful avenues for future research on legislative institutions.

SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S0003055420000192>.

Replication materials can be found on Dataverse at: <https://doi.org/10.7910/DVN/XI7ENB>.

REFERENCES

- Aldrich, John H., Jacob M. Montgomery, and David B. Sparks. 2014. "Polarization and Ideology: Partisan Sources of Low Dimensionality in Scaled Roll Call Analyses." *Political Analysis* 22 (4): 435–56.
- Bräuninger, Thomas, Jochen Müller, and Christian Stecker. 2016. "Modeling Preferences Using Roll Call Votes in Parliamentary Systems." *Political Analysis* 24 (2): 189–210.
- Carey, John M. 2007. "Competing Principals, Political Institutions, and Party Unity in Legislative Voting." *American Journal of Political Science* 51 (1): 92–107.
- Carey, John M. 2009. *Legislative Voting and Accountability*. Cambridge: Cambridge University Press.
- Carroll, Royce, and Keith T. Poole. 2014. Roll-Call Analysis and the Study of Legislatures. In *The Oxford Handbook of Legislative Studies*, ed. Shane Martin, Thomas Saalfeld, and Kaare W. Strøm. New York: Oxford University Press, 103–24.
- Carrubba, Clifford J., Matthew Gabel, Lacey Murrah, Ryan Clough, Elizabeth Montgomery, and Rebecca Schambach. 2006. "Off the Record: Unrecorded Legislative Votes, Selection Bias, and Roll-Call Vote Analysis." *British Journal of Political Science* 36 (4): 691–704.
- Carrubba, Clifford J., Matthew Gabel, and Simon Hug. 2008. "Legislative Voting Behavior, Seen and Unseen: Adjusting for Selection Effects in Roll Call Vote Analysis." *Legislative Studies Quarterly* 33 (4): 543–72.

- Carson, Jamie L., Gregory Koger, Matthew J. Lebo, and Everett Young. 2010. "The Electoral Costs of Party Loyalty in Congress." *American Journal of Political Science* 54 (3): 598–616.
- Cattell, Raymond B. 1966. "The Scree Test for the Number of Factors." *Multivariate Behavioral Research* 1 (2): 245–76.
- Clinton, Joshua D. 2007. "Lawmaking and Roll Calls." *Journal of Politics* 69 (2): 455–67.
- Clinton, Joshua D. 2012. "Using Roll Call Estimates to Test Models of Politics." *Annual Review of Political Science* 15 (1): 79–99.
- Clinton, Joshua, and John Lapinski. 2008. "Laws and Roll Calls in the U.S. Congress, 1891–1994." *Legislative Studies Quarterly* 33 (4): 511–41.
- Crespin, Michael H., and David W. Rohde. 2010. "Dimensions, Issues, and Bills: Appropriations Voting on the House Floor." *The Journal of Politics* 72 (4): 976–89.
- Crisp, Brian F., and Amanda Driscoll. 2012. "The Strategic Use of Legislative Voting Procedures." *Legislative Studies Quarterly* 37 (1): 67–97.
- Desposato, Scott W. 2001. "Legislative Politics in Authoritarian Brazil." *Legislative Studies Quarterly* 26 (2): 287–317.
- Hansen, Martin Ejnar, and Marc Debus. 2012. "The Behaviour of Political Parties and MPs in the Parliaments of the Weimar Republic." *Party Politics* 18 (5): 709–26.
- Hirsch, Alexander V. 2011. "Theory Driven Bias in Ideal Point Estimates: A Monte Carlo Study." *Political Analysis* 19 (1): 87–102.
- Høyland, Bjørn. 2010. "Procedural and Party Effects in European Parliament Roll-Call Votes." *European Union Politics* 11 (4): 597–613.
- Hug, Simon. 2010. "Selection Effects in Roll Call Votes." *British Journal of Political Science* 40 (1): 225–35.
- Hug, Simon, Simone Wegmann, and Reto Wüest. 2015. "Parliamentary Voting Procedures in Comparison." *West European Politics* 38 (5): 940–68.
- Imai, Kosuke, James Lo, and James Olmstead. 2016. "Fast Estimation of Ideal Points with Massive Data." *American Political Science Review* 110 (4): 631–55.
- Inter-Parliamentary Union. 1986. *Parliaments of the World: A Comparative Reference Compendium*. London: Macmillan.
- Jenkins, Shannon. 2008. "Party Influence on Roll Call Voting: A View from the U.S. States." *State Politics & Policy Quarterly* 8 (3): 239–62.
- Jensen, Nathan M., Edmund Malesky, and Stephen Weymouth. 2014. "Unbundling the Relationship between Authoritarian Legislatures and Political Risk." *British Journal of Political Science* 44 (3): 655–84.
- Kaiser, Henry F. 1960. "The Application of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement* 20 (1): 141–51.
- Kreppel, Amie. 2002. *The European Parliament and Supranational Party System: A Study in Institutional Development*. New York: Cambridge University Press.
- Lee, Frances E. 2018. "The 115th Congress and Questions of Party Unity in a Polarized Era." *Journal of Politics* 80 (4): 1464–73.
- Lynch, Michael S., and Anthony J. Madonna. 2013. "Viva Voce: Implications from the Disappearing Voice Vote, 1865–1996." *Social Science Quarterly* 94 (2): 530–50.
- Malesky, Edmund, and Paul Schuler. 2010. "Nodding or Needling: Analyzing Delegate Responsiveness in an Authoritarian Parliament." *American Political Science Review* 104 (3): 482–502.
- Malesky, Edmund, Paul Schuler, and Anh Tran. 2012. "The Adverse Effects of Sunshine: A Field Experiment on Legislative Transparency in an Authoritarian Assembly." *American Political Science Review* 106 (4): 762–86.
- Mazzei, Patricia. 2018. "Even Popular Gun Law Proposals Have a Tough Road in Florida." *New York Times*, March 5, 2018.
- McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2001. "The Hunt for Party Discipline in Congress." *American Political Science Review* 95 (3): 673–87.
- Mueller, Wolfgang, and Ulrich Sieberer. 2014. "Procedure and Rules in Legislatures." In *The Oxford Handbook of Legislative Studies*, ed. Shane Martin, Thomas Saalfeld, and Kaare Strom. Oxford: Oxford University Press, 311–32.
- Poole, Keith T. 2014. *Spatial Models of Parliamentary Voting*. New York: Cambridge University Press.

³⁶ These types of changes in votes toppled three Italian governments (Mueller and Sieberer 2014). For a more recent example, consider the March 2018 votes in the State of Florida Senate on banning AR-15-style rifles. The measure passed on a voice vote, but it was subsequently reversed on a roll-call vote (Mazzei 2018). Our survey of legislatures shows that this sort of institutional manipulation is possible in a wide range of legislatures.

- Poole, Keith T., and Howard Rosenthal. 2007. *Ideology and Congress*. New Brunswick, NJ: Transaction Publishers.
- Poole, Keith T., Jeffrey Lewis, James Lo, and Royce Carroll. 2011. "Scaling Roll Call Votes with `wnominate` in R." *Journal of Statistical Software* 42 (14): 1–21.
- Roberts, Jason M., and Steven S. Smith. 2003. "Procedural Contexts, Party Strategy, and Conditional Party Voting in the U.S. House of Representatives, 1971–2000." *American Journal of Political Science* 47 (2): 305–17.
- Rosenthal, Howard, and Erik Voeten. 2004. "Analyzing Roll Calls with Perfect Spatial Voting: France 1946–1958." *American Journal of Political Science* 48 (3): 620–32.
- Saalfeld, Thomas. 1995. "On Dogs and Whips: Recorded Votes." In *Parliaments and Majority Rule in Western Europe*, ed. Herbert Doring. New York: St. Martin's Press, 528–65.
- Sieberer, Ulrich. 2006. "Party Unity in Parliamentary Democracies: A Comparative Analysis." *Journal of Legislative Studies* 12 (2): 150–78.
- Sieberer, Ulrich, Thomas Saalfeld, Tamaki Ohmura, Henning Bergmann, and Stefanie Bailer. Forthcoming. "Roll-Call Votes in the German Bundestag: A New Dataset, 1949–2013." *British Journal of Political Science*.
- Snyder, James M., Jr. 1992. "Committee Power, Structure-Induced Equilibria, and Roll Call Votes." *American Journal of Political Science* 36 (1): 1–30.
- Snyder, James, and Timothy Groseclose. 2000. "Estimating Party Influence in Congressional Roll-Call Voting." *American Journal of Political Science* 44 (2): 193–211.
- Stecker, Christian. 2015. "How Effects on Party Unity Vary across Votes." *Electoral Studies* 21 (5): 791–802.
- Thierse, Stephan. 2016. "Going on Record: Revisiting the Logic of Roll-Call Vote Requests in the European Parliament." *European Union Politics* 17 (2): 219–41.
- Van Doren, Peter M. 1990. "Can We Learn the Causes of Congressional Decisions from Roll Call Data?" *Legislative Studies Quarterly* 15 (3): 311–40.