

Using Roll Call Estimates to Test Models of Politics

Joshua D. Clinton

Department of Political Science and Center for the Study of Democratic Institutions, Vanderbilt University, Nashville, Tennessee 37235-1817; email: josh.clinton@vanderbilt.edu

Annu. Rev. Polit. Sci. 2012. 15:79–99

First published online as a Review in Advance on March 8, 2012

The *Annual Review of Political Science* is online at polisci.annualreviews.org

This article's doi:
10.1146/annurev-polisci-043010-095836

Copyright © 2012 by Annual Reviews.
All rights reserved

1094-2939/12/0615-0079\$20.00

Keywords

ideal point estimation, formal theory, empirical implications of theoretical models

Abstract

Measuring the preferences of political elites is critically important for analyzing the determinants and consequences of elite behavior. The decisions that elites make when casting roll call votes seem to provide an ideal opportunity for measuring elite preferences and testing theories of the political process. The fact that the resulting ideal points are a consequence of applying a statistical model to a model of individual choice, however, may affect their usefulness for measuring elite preferences and testing predictions regarding individual and collective decision making. When analyzing roll call votes, scholars should be mindful of how their decisions may affect the estimates from their analyses. I use simulations to illustrate how the nonrandom selection of roll calls may affect the ability to estimate ideal points that accurately reflect the preferences responsible for generating the observed votes, and I discuss work integrating the many models involved in the production and consumption of roll call estimates.

INTRODUCTION

Why should political scientists care about the analysis of roll call votes? Such analyses may seem to be motivated primarily by a desire to manipulate Greek characters and write computer code to analyze matrices of ones and zeros, but this could not be further from the truth. The underlying motivation for analyzing votes recorded in a legislature or court is to measure characteristics that are essential for understanding the causes and consequences of politics—the policy preferences of political elites.

Measuring elite preferences is extremely important because much of what we are interested in as political scientists involves either the determinants or consequences of policy preferences. Empirical studies of elite behavior face the difficulty that elites' preferences are unobservable, and we often lack a reliable analogue to public opinion surveys that are typically used to characterize citizens' preferences. Moreover, even if elite surveys were more plentiful, elites' incentives would suggest justifiable concerns about the representativeness of responding elites and the meaning of the received responses. The difficulty of measuring elite preferences does not alleviate the necessity of the task, however, and many have turned to roll calls in the hope of using the observed “yeas” and “nays” to estimate preferences.

A tremendous amount of time and energy in nearly every subfield of political science has been devoted to analyzing roll calls or using the resulting estimates in secondary analyses. Starting with Lowell (1902), political scientists have generally used roll call votes to investigate two types of questions: why do legislators behave as they do (e.g., Bartels 1991, Ansolabehere et al. 2001, Bafumi & Herron 2010), and how do institutions affect the choices that are made (e.g., Krehbiel 1998, Binder 1999, Chiou & Rothenberg 2003, Cox & McCubbins 2005, Wiseman & Herron 2008, Wiseman & Wright 2008)?

Roll calls appear to offer an ideal opportunity for measuring the preferences of political elites because we observe elites registering their individual support on many proposed policies. Although the term “roll call” technically refers to the recording of a vote in a deliberative body such as a legislature, the methods that were originally developed to analyze votes in the US Congress have been applied to binary decision-making behavior by nonlegislators in executive and judicial branches throughout the world (e.g., decisions made by a Supreme Court justice, public positions taken by a president or nominee, votes by a country in the United Nations' General Assembly), as well as in groups not formally affiliated with government (e.g., interest groups). Though cognizant of the breadth of applications, for the sake of clarity, I hereafter refer to the binary decisions being analyzed as “votes” and the political actors who cast the observed votes as “legislators” even though neither characterization may be accurate.

Scholarship largely, but not always, focuses either on the statistical and computational issues involved in analyzing roll calls (e.g., Poole & Rosenthal 1985, 2007; Heckman & Snyder 1997; Martin & Quinn 2002; Clinton et al. 2004a) or else on using the resulting estimates in second-stage regression analyses to evaluate theories of the political process. The exceptionally difficult computational problems facing pioneers like Poole & Rosenthal (1985) have been largely solved, but interpreting the measures that result from the analysis of roll calls remains an important and ongoing research agenda. At issue is whether the implicit division of labor between the production and consumption of roll call estimates obscures consequential connections between how the matrix of roll calls is analyzed to produce estimates of elite preferences and how those estimates are used in subsequent analyses to evaluate models of the political process. Because every parameter estimated from roll calls depends on the assumed model of individual choice (and perhaps also on how the model is statistically implemented), interpreting the roll call estimates requires understanding the consequences of the many modeling decisions that are made when analyzing roll call votes.

Rather than attempt the impossible task of summarizing the many important contributions that scholars have been able to make because of roll call analyses, I focus my review on some issues that scholars working with roll calls should consider when analyzing and interpreting the resulting estimates. I focus on the analysis of roll calls, but many of the same considerations likely apply to using any observed behavior to measure participants' preferences. Although I do not draw explicit connections, similar issues may arise when scholars try to measure elite preferences by using political speeches or text (e.g., Laver et al. 2003), the sponsorship and cosponsorship behavior of legislators (e.g., Desposato et al. 2011), the decision of presidents or nominees to take a public policy position on an issue or not (e.g., Bailey 2007, Bertelli & Grose 2011), and the decision to contribute to a campaign or not (e.g., McCarty et al. 2006, Bonica 2011, Chen & Johnson 2011).

TESTING A MODEL USING A MODEL OF A MODEL

What does it mean to test a theory? Even if we can at best only disprove a theory (Popper 1961), the claim that formal models of politics can be tested and disproved is a contested proposition (see, e.g., the positions of Riker 1980, Morton 1999, Diermeier & Krehbiel 2003, Clarke & Primo 2007, Granato et al. 2010). A model is a simplified representation of reality, and it necessarily provides an incomplete account of the interactions. Even so, scholars are typically—but not always—interested in whether some models are better able to predict observable variation. Evaluating the predictions of a formal model using roll calls faces at least two difficulties.

First, the strategic interactions being modeled may be responsible for only a small amount of the variation in the data. A model may perfectly describe the effect of the political interactions being modeled, and participants may react to the constraints and incentives in a manner consistent with the formal model, but the variation in the data due to the modeled effect may still be small relative to other unmodeled pressures. Put differently, a model may provide an accurate depiction of the modeled interactions but there may be minimal observable effects because of unmodeled influences.

Second, the modeling decisions made when analyzing roll calls may make it difficult to correctly characterize the variation of theoretical interest. Unlike the preferences of formal models that are exogenous primitives (McCarty & Meirowitz 2007), the ideal point estimates produced from roll call analyses reflect the preferences members reveal in their behavior as interpreted through the application of a statistical model to an underlying model of individual choice. Even if the variation predicted by the model is empirically discernible, the predictions of a formal model might still not be supported by roll call estimates if (a) the formal model is wrong, (b) the formal model is correct but a statistical model used to generate or evaluate the estimates is wrong, or (c) both the formal and statistical models are wrong. It is obviously impossible to ensure that our models are correct, but we can control whether our models share common behavioral assumptions. If the models of individual choice assumed by Models A and B differ, for example, it is unclear what the relationship between the predictions of Model A and the estimates of Model B implies about Model A's predictive ability.

Because both the formal and statistical models are representations of the political process, a goal of theory testing is to minimize the differences between the models so that if we choose to reject a theory based on its inability to predict observable variation, we reject it under the most favorable conditions. This requires determining whether the model responsible for generating the roll call estimates is consistent with the models being tested and, if not, to understand the possible consequences.

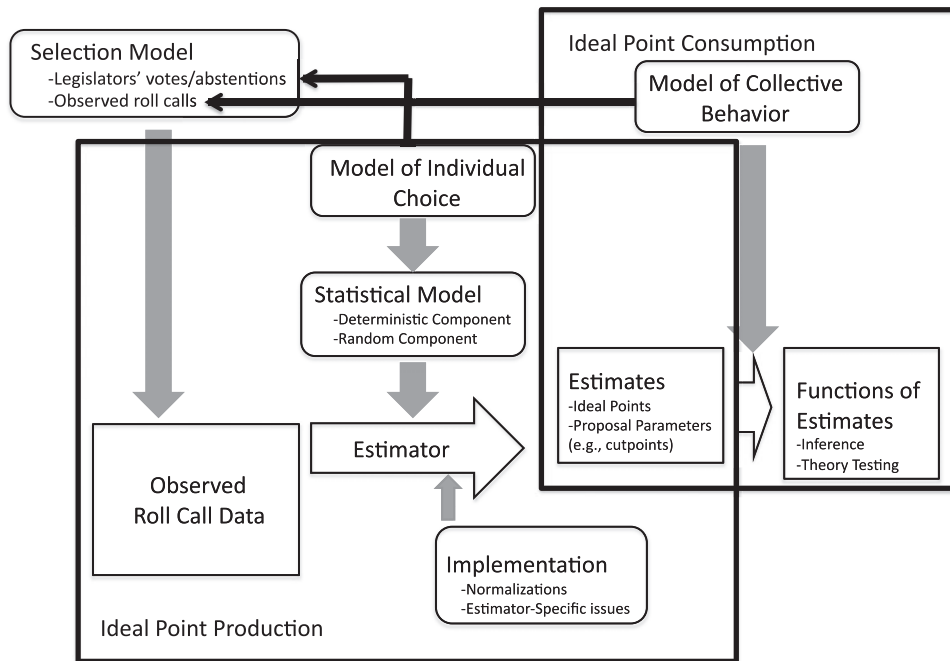


Figure 1
Decisions made in the production and consumption of roll call estimates.

THE RELATIONSHIP BETWEEN MODELS

The analysis of roll calls occupies a prominent place in political science because it strives to accomplish the elusive goal of measuring the preferences of political elites by using a statistical model based on the commonly used spatial voting model of individual choice (Enelow & Hinich 1984). This section builds on the considerations raised above by reviewing some aspects of the commonly assumed models that may be relevant when using the resulting estimates to evaluate models of the political process.

Roll call analysis typically involves analyzing a matrix of binary choices (see Voeten 2000 for an application to voting in the United Nations with three outcomes) where the analyst knows only the identity of the legislator casting the vote on each issue. Interpreting what the resulting matrix of ones and zeros implies about the politics and preferences responsible for producing public policy requires a lot of additional assumed structure (see Kalandrakis 2010 for an overview). **Figure 1** outlines the various decisions involved in the production and consumption of ideal points and some of the connections between them.

As **Figure 1** makes clear, producing ideal points requires applying a statistical model to a model of individual choice. The meaning of the recovered estimates therefore depends critically on the modeling decisions that are made. Moreover, interpreting what the roll call estimates imply about the nature of politics often, but certainly not always, involves using the estimates to assess predictions from yet another model—a model of the political process (the “ideal point consumption” problem). For example, a study using ideal points to study the effects of congressional institutions necessarily involves using three models: a model of individual choice that explains why members vote as they do, a statistical model that implements the assumed individual choice model, and

a model of the congressional institution. Selection models may also be potentially implicated as legislators miss votes and only some issues are considered on the agenda.

As the previous section argues, the potential for mistaken inferences partially depends on the consistency of the multitude of models used in the analyses of roll calls. If the models assume different data-generating processes or contradict the model of the political process of interest, it is difficult to imagine that much can be learned from the resulting roll call estimates.

There are many decisions in **Figure 1** that may be of consequence, but I focus on three aspects of the individual choice model that are common to prevalently used roll call estimators [i.e., the various flavors of NOMINATE (Poole & Rosenthal 2007) and the Bayesian item response models of Clinton et al. 2004a, Martin & Quinn 2002, and Bafumi et al. 2005]: the meaning of the estimated ideal point, the nature of legislators' voting behavior, and the assumed voting error.¹ To be clear, I raise questions about the model of individual choice not because existing analyses are incorrect, but to highlight aspects that may be worthy of more careful consideration when evaluating theoretical predictions using "off-the-shelf" estimates or estimators (perhaps especially when applied to contexts other than the US Congress).

What Is an "Ideal" Point?

All roll call models assume that legislators have a most-preferred point in the policy space that they would implement if they could. Legislators' policy preferences are satiable, and the so-called ideal point may reflect personal preferences, constituency preferences, or any combination of factors that induce the legislator to prefer a particular policy. In fact, an entire literature explores why legislators take the positions they do. Whether preferences are defined over policy outcomes or public positions is unclear, and the meaning likely varies by application. Work focused on the representation of interests and constituents (e.g., Bartels 1991, Ansolabehere et al. 2001, Bafumi & Herron 2010) typically requires only that ideal points reflect legislators' public positions, but some assume that ideal points characterize true preferences over policy outcomes (e.g., Chiou & Rothenberg 2003, Bianco et al. 2004, Bianco & Sened 2005).

Given a most-preferred outcome or position in the policy space, the model of individual choice in roll call analyses assumes that legislators vote for the closest expected alternative in every pairwise comparison. Legislator i votes yes on vote t (i.e., $y_{it} = 1$) if the policy position associated with voting yes ($\theta_{y(t)}$) is closer to her most-preferred ideal point x_i than the position associated with voting no ($\theta_{n(t)}$). The specific utility function depends on the estimator being used [the NOMINATE family of estimators assume Gaussian utility (Poole & Rosenthal 2007) and Bayesian item response models typically assume quadratic utility], but the basic choice model is: $y_{it} = 1$ if $|x_i - \theta_{y(t)}| < |x_i - \theta_{n(t)}|$ and 0 otherwise. In words, the model of individual choice assumes that only the relative distance between a legislator's most-preferred policy and the two alternatives being considered is relevant for determining her vote (the effects of voting error are discussed below).

It seems reasonable to wonder whether a legislator might cast votes for reasons other than the relative proximity of the alternative to the legislator's most-preferred policy. For example, it may be that legislators are sometimes pressured to cast votes for more distant alternatives or that they are willing to "sell" their votes to interested buyers such as interest groups. There are many models of vote buying (e.g., Snyder 1991, Snyder & Groseclose 2000), and there is a robust debate concerning whether political parties influence the voting behavior of individual members

¹Work by Carroll et al. (2009) and Clinton & Jackman (2009) explores how other differences between the estimators may matter.

(e.g., Cox & McCubbins 1993, 2005; Glazer et al. 1995; Krehbiel 2000; Snyder & Groseclose 2000; McCarty et al. 2001; Cox & Poole 2002; Smith 2007; Roberts & Bell 2008). At issue is whether the model of individual choice used to analyze roll calls conflicts with the assumptions of the theory being tested.

Both the roll call model and the theoretical model assume the existence of a true ideal point x_i , but analyzing roll calls can only recover the preferences that rationalize the observed votes given the assumed model of individual choice as implemented via a statistical model. It is therefore possible that the preferences revealed in the pattern of votes differ from true preferences. If legislators vote for alternatives because of nonspatial considerations, the model of individual choice that is assumed by common estimators is incorrect. For example, if some legislators are pressured, the ideal point responsible for generating the observed votes will be a function of both the legislator's true ideal point (x_i) and the effect of being pressured (p_i). In that case, the estimated ideal point will be $x_i^* = x_i + p_i$. Moreover, the pressure p_i will likely vary depending on x_i and vary across votes. Absent additional information (or assumptions), it is impossible to tell from the votes themselves whether such pressure exists [but see work by Evans (2008), who uses whip counts to help locate possible party pressure].

If such pressure exists, there are three possible sets of ideal point estimates: the ideal points that reflect the effects of external pressure (x^*), the ideal points in the absence of any pressure (x), and the standard "off-the-shelf" estimates that ignore the distinction and therefore contain both pressured and unpressured ideal points. The magnitude of the resulting error depends on the amount of pressure and vote buying that occurs, but if such pressure exists, the standard estimates may be inappropriate. Because of this, some scholars have changed the behavioral model of the canonical roll call model to allow for the possible effects of pressure on voting behavior. McCarty et al. (2001), for example, model party pressure using separate cutpoints, and Snyder & Groseclose (2000) use different voting models depending on whether the vote is lopsided or not to estimate pressured and nonpressured ideal points. Determining the correct estimates for the application of interest is important because if the model of individual choice assumed by the roll call estimator is inconsistent with the model of choice assumed by the theoretical model of interest, it is difficult to interpret what the former reveals about the latter.

Do Legislators Behave Uniformly?

A second potential consideration involves the meaning of the proposals being voted on. Are votes cast sincerely based only on the proposals involved in each pairwise comparison, or do legislators vote strategically based on their expectations? Moreover, is there variation in this behavior across votes, legislators, or both?

If legislators behave identically, whether $\theta_{y(t)}$ represents the actual proposal being considered in vote t or the policy outcome that is expected to occur should the proposal pass (i.e., the sophisticated equivalent) is irrelevant for estimating legislators' ideal points (Poole & Rosenthal 1997). If all legislators vote sincerely, the estimated ideal points will reflect the true ideal points, and if all legislators vote strategically, the estimated ideal points will reflect legislators' preferences over the expected sophisticated equivalents. That said, determining whether votes are cast for sincere or strategic reasons is obviously critical for interpreting what the estimates mean.

Difficulties emerge, however, if the model of individual choice does not account for variation in legislators' behavior—e.g., some legislators vote sincerely based on the proposal being considered and others vote strategically based on expectations of what the success or failure of the vote will produce. Applying commonly used roll call estimators in such circumstances will produce erroneous estimates that reflect a mixture of sincere and strategic behavior. Roberts (2007), for

example, shows how the apparent extremity of party leaders' ideal points in the US Senate is affected by the fact that majority leaders sometimes strategically vote with the minority in order to offer a motion to recommit if the majority loses, and Spirling & McLean (2007) argue that the government-versus-opposition nature of the British Westminster system can create incentives for strategic voting.

The prevalence of strategic voting is a topic of continuing interest to scholars who analyze roll calls, but it is typically studied at the agenda stage rather than at the legislator level (e.g., Jenkins & Munger 2003, Groseclose & Milyo 2010). The prospect of strategic voting by legislators is obviously consequential because it means that scholars must again consider whether the sincere, strategic, or pooled ideal points are most relevant for the question and model of interest.

A secondary complication that emerges if legislators vote strategically is the fact that strategic voting will make it difficult to compare the estimates of seemingly identical proposals across time or contexts. Even if identically worded proposals are voted on in different times or institutions, the meaning of the estimates may differ because of changes in the strategic environment, and the influence of strategic considerations may therefore make it difficult to interpret the meaning of the estimated vote parameters from seemingly identical choices. This will create obvious difficulties if such votes are used to "bridge" estimates across time or institutions.

Random Abstention?

Existing roll call models also commonly assume that abstentions are random. This means that whether a legislator votes or not is unrelated to both the issue being voted on and the legislator's ideal point. If this is not the case, conventional roll call estimators may recover misleading estimates because of the legislator-level or vote-level selection models present in the data-generating process (Heckman 1979). To take an extreme example, if a legislator only votes on party-line votes, the observed pattern of votes will reveal an extreme ideal point because the member is only observed voting with the most extreme members of her party.

The assumption that abstentions can be treated as votes that are missing at random is contrary to some work that argues that legislators' absences from roll call votes are often intentional (e.g., Cohen & Noll 1991, Forgette & Sala 1999). The substantive consequences of ignoring the non-random selection depends on the prevalence of abstention. Rosas et al. (2011) document the level of abstention in legislatures around the world and show that if the model of abstention behavior is known, the model of individual choice that is assumed by the roll call estimator can be modified to account for nonrandom abstention. In the US Congress, the amount of abstention varies over time (e.g., Jones 2003), but the relatively low level of missingness means that ignoring the nonrandom abstention is likely inconsequential (Poole & Rosenthal 2007) for the estimates of all but a few legislators (e.g., Clinton et al. 2004b). That said, abstention is probably more consequential in other contexts, especially if the abstention is motivated by a desire to avoid taking positions.

The assumption that a legislator's votes are missing at random may also be consequential when the analyst is attempting to create comparable preference measures by "bridging" ideal points across institutions or time. Addressing questions involving the behavior of elites in different institutions or at different times often requires constructing a directly comparable measure of elite preferences. For example, analyzing the influence of multiple branches of the US government on law making requires comparing the preferences of Congress, the president, the courts, and executive agencies (e.g., Bailey 2007, Treier 2011).

To construct comparable measures of elite preferences, scholars sometimes treat the public positions of noncongressional actors on issues coming to a vote in Congress as equivalent to a roll call vote and estimate ideal points based on the positions that are taken. For example, scholars

have used presidential statements to estimate presidential ideal points (McCarty & Poole 1995) and nominees' testimony before Congress to measure the preferences of agency heads (Bertelli & Grose 2011). So doing assumes that the positions that are not taken can be treated as missing at random, even though it seems likely that there are strategic reasons why only some public positions are taken. If so, the model of individual selection may matter, and the resulting ideal points may provide a misleading characterization of elite preferences (e.g., Treier 2010). The possibility of nonrandom selection also almost certainly arises whenever the canonical roll call model is applied to discretionary behavior for which participation widely varies and there may be strategic incentives that affect the decision to participate or not (e.g., the decision to cosponsor legislation, speak on a topic, or contribute to a particular candidate).

Independent and Idiosyncratic Stochastic Shocks?

The final assumption I consider involves the voting error that is assumed to be present in roll call voting models. The presence of voting error is absolutely critical for roll call analyses because only ordinal rankings are identified with perfect (i.e., errorless) spatial voting (Poole 2000, 2005). Ordinal rankings of legislators are of limited use, however, because most formal models generate predictions involving either the distance or dispersion of elites' preferences. If the assumption of probabilistic voting error is violated in some legislatures—typically Westminster parliamentary systems—other methods may be required (e.g., Rosenthal & Voeten 2004).

Even if legislators vote probabilistically, however, how the voting errors occur in the model of individual choice may matter. Standard roll call estimators assume that legislators sometimes vote incorrectly despite perfectly knowing their ideal point and the location of the proposals being voted on. The assumed voting error is therefore not attributable to uncertainty about the proposals being voted on or the ideal point that the legislator is attempting to implement. Instead, each legislator experiences an independent and identically distributed exogenous shock to her calculation of which alternative is closer to her ideal point on every vote.

Alternative specifications of voting error are certainly possible, but the literature has only briefly explored them. Ladha (1991), for example, presents a probabilistic voting model where legislators idiosyncratically and imperfectly perceive the proposal being voted on, and Lauderdale (2010) allows for heteroskedastic voting error. Given the importance of voting error documented in the simulations below, better understanding the nature and effects of voting error is important for interpreting the meaning of the recovered estimates.

Because ideal points result from applying a statistical model to a model of individual choice, scholars who use the ideal points based on revealed behavior to measure elite preferences must consider the implications of the models used to extract ideal points. If the assumptions of the model being evaluated conflict with the models used to estimate ideal points, it is unclear what the latter can reveal about the former. Not every concern will be relevant for every use, but considering the connections and implications of the choices that are made to generate ideal point estimates like those I highlight above helps to ensure that the resulting estimates are appropriately used and interpreted.

THE IMPORTANCE OF THE NONRANDOM SELECTION OF ROLL CALL VOTES

Unlike the sample of a properly designed public opinion survey, roll calls are not a random sample of possible policy proposals (e.g., Van Doren 1990, Jackson & Kingdon 1992, Smith 2007). The issues that come to vote are intentionally chosen by the legislators themselves, and a concern therefore arises regarding whether the nonrandom selection of votes might affect the resulting

estimates. Does the nonrandom selection of votes affect the ability to measure the legislators' true preferences using estimated ideal points?

One prominent example involves the artificial extremism of interest group scores. Because interest groups strategically choose to score only a subset of observed votes, measures of preferences based on interest group scores appear more extreme than the true preferences would suggest (e.g., Snyder 1992a, Herron 1999). It is unclear whether similar concerns emerge when analyzing the roll calls that are produced by the political process. Roberts & Smith (2003), for example, show how allowing recorded votes in the Committee of the Whole of the US House affects the recovered ideal points by changing the composition of the issues being voted on. Hug (2010) uses unique circumstances in the Swiss lower house to show that votes that were thought by legislators to be confidential differ from those that were known to be public. Kestel & Lax (2008) use simulations to explore the effects of discretionary case selection on the estimation of judicial ideal points.

The endogeneity of the legislative agenda may not be relevant for every question (e.g., studies of why legislators behave as they do on the issues that come to vote), but it is most likely consequential when analyzing the production of policy. Formal models of the law-making process, for example, predict which outcomes should and should not be observed in equilibrium (e.g., Cox & McCubbins 2005). Even if we assume that every outcome receives a roll call vote (but see Clinton & Lapinski 2008 for evidence to the contrary), the decision of which issues to consider on the agenda is likely affected by elite preferences and the strategic context. Recent work has begun to model the incentive for recording a vote or not (e.g., Snyder & Ting 2003, 2005, 2011; Carruba et al. 2010), but more work is required to better understand the effect of the nonrandom selection of roll calls for the estimation and interpretation of ideal points.

An Example: Agenda-Setting Selection Models and Roll Calls

To demonstrate how selection models can affect whether estimated ideal point estimates reflect true preferences, consider the agenda setter game of Romer & Rosenthal (1978). Many prominent law-making models in political science can be interpreted as a variant of this game (see, e.g., Krehbiel 1996, Chiou & Rothenberg 2003, Cox & McCubbins 2005). Let S denote the legislator with the power to make take-it-or-leave-it proposals, and let the median legislator be known as M . Assume every legislator has a preferred outcome in a unidimensional policy space, and let X_S and X_M represent the ideal policy of the agenda setter and the median legislator, respectively, with $X_S > X_M$. Players have preferences over policy outcomes, and the utility of an outcome is a decreasing function of distance from a player's ideal policy. The sequence of the game is as follows: Nature chooses a status quo q for consideration, S proposes a piece of legislation p given q and knowledge of X_M , the legislature (represented by M 's decision) votes to pass p or keep q , and the game ends with M 's decision being realized.

Because only S can make a take-it-or-leave-it offer to M , the game is a model of positive agenda setting (PAS). If we change the game so that S decides whether to keep q or let M change q , the model is one of negative agenda setting (NAS). Changing the identity of S can therefore help illuminate the implications of agenda setting by congressional committees or majority parties, and adding additional veto players can help model the possible effects of supermajoritarian institutions such as vetoes and filibusters.

Given these models of negative and positive agenda setting, NAS and PAS (and the canonical median voter theorem), we can derive the equilibrium proposal p^* and the location of the cutpoint implementing the equilibrium proposal c^* —i.e., the location of the legislator indifferent between voting for p^* and voting for q —for every status quo policy q . **Table 1** presents the results.

PAS: positive agenda-setting model

NAS: negative agenda-setting model

Table 1 Predictions of various models of legislative activity

Location of status quo (q)	Median voter (MVT)		Positive agenda setter (PAS)		Negative agenda setter (NAS)	
	Equilibrium proposal (p^*)	Equilibrium cutpoint (c^*)	Equilibrium proposal (p^*)	Equilibrium cutpoint (c^*)	Equilibrium proposal (p^*)	Equilibrium cutpoint (c^*)
$< 2X_M - X_S$	X_M	$(X_M + q)/2$	X_S	$(X_S + q)/2$	X_M	$(X_M + q)/2$
$[2X_M - X_S, X_M]$	X_M	$(X_M + q)/2$	$2X_M - q$	X_M	X_M	$(X_M + q)/2$
$[X_M, X_S]$	X_M	$(X_M + q)/2$	none	none	none	none
$[X_S, 2X_S - X_M]$	X_M	$(X_M + q)/2$	X_S	$(X_S + q)/2$	none	none
$> 2X_S - X_M$	X_M	$(X_M + q)/2$	X_S	$(X_S + q)/2$	X_M	$(X_M + q)/2$

For status quos lying between the preferences of the median legislator and the agenda-setting legislator (i.e., for $q \in [X_M, X_S]$), the median voter theorem predicts a proposal at X_M because the median voter can propose her ideal point and have that proposal receive the support of a majority. The NAS and PAS models predict that policy change will be impossible in this circumstance because M and S will be unable to agree to any change. For status quos lying in the interval $[X_S, 2X_S - X_M]$, both the median voter model and the PAS model predict change—although they differ as to whether the outcome should reflect the ideal point of M or S. The NAS game predicts that change will not occur if q is in this interval because S will not allow M to change the policy and make S worse off.

Given these competing predictions, scholars often try to determine whether one of these three competing models can better account for the behavior we observe in a legislature of interest. Although no model can completely account for the observed activity, do observed outcomes rule out some predictions more than others? Most empirical investigations in political science investigate this question using the reduced-form approach summarized in **Figure 1**—ideal points from the “ideal point production” problem in the figure are used as dependent or independent variables in regressions investigating the comparative statics of formal models.

A prominent example of this type of analysis for the models listed in **Table 1** involves the measurement and analysis of “gridlock intervals” (Krehbiel 1998). Gridlock intervals define the regions in the policy space where status quos either cannot or will not be changed according to the model. For example, status quos in the interval $[X_M, X_S]$ are gridlocked under PAS models because any proposal made by S to amend such a q will be defeated. Status quos in the interval $[X_M, 2X_S - X_M]$ are also gridlocked in the NAS model because S will not allow such status quos to be considered because of the policy that M would enact.

Because the identities of the theoretically relevant actors M and S change over time, the gridlock intervals (G) expand and shrink over time, and the changes will presumably vary in different ways for the different models [i.e., $(G_{NAS,t} - G_{NAS,t+1}) \neq (G_{PAS,t} - G_{PAS,t+1})$]. If so, correlating measures of legislative performance (e.g., enactments) with gridlock intervals over time appears to reveal which gridlock interval best predicts the variation in observed activity (e.g., Binder 1999, Chiou & Rothenberg 2003).

An apparent complication with this approach is that the votes used to measure the ideal points used to measure gridlock intervals are themselves a result of the political process. Whereas the theoretical measure of the gridlock interval involves the legislators’ true preferences, analyzing roll calls can only recover the ideal points that rationalize the observed votes. As **Table 1** reveals, the formal models predict the votes that should and should not be observed if each theory is true.

Many scholars have raised the endogenous nature of roll calls (e.g., Van Doren 1990). Snyder (1992b, 1992c), for example, raises the issue with respect to the committee-outlier debate when

arguing that it may be problematic to use ideal points based on interest group scores to test competing theories of committee organization because the scored roll calls are presumably a consequence of both the committee process and the interest group's selection process. Herron (2001) extends the concerns of Snyder (1992a) to prove that using "artificially extreme" interest group scores in a second-stage regression produces inconsistent estimates because the interest group scores measure true preferences with error and the error depends on the interest groups' intentional choice of which votes to score. Clinton (2007) argues that computing gridlock intervals using ideal point estimates ignores the models' implications for which votes should be observed, and Hirsch (2011) demonstrates that the effect depends on the amount of probabilistic voting error.

Assessing Selection Effects Via Simulation

Because the analysis of roll calls uses a statistical model of an individual choice model to generate ideal points and test the theoretical predictions from possibly yet another set of models, a great deal is being assumed about the data-generating process. In such circumstances, simulations can help explore the possible consequences of the various decisions summarized in **Figure 1** because we can generate roll calls and estimate parameters for a known data-generating process. Simulation studies are being increasingly used for this reason (e.g., Krehbiel 2000, Kastellec & Lax 2008, Carroll et al. 2009, Stiglitz & Weingast 2010, Hirsch 2011), but valuable opportunities remain for exploring concerns like those expressed in the prior section.

As an illustration, I explore how the selection model and the amount of voting error affect the recovery of true preferences using roll calls. I use a simulation study to demonstrate the importance of the roll call selection model and I characterize how the effect depends on the amount of voting error that is assumed to be present in the data-generating process. To be clear, the investigation is intended only to illustrate how simulations might be used to better understand the properties of roll call estimators when applied to particular contexts.

Assume that 51 legislators have unidimensional ideal points with the preferences of the 22 leftmost members uniformly distributed over the interval $[-1, 0]$ and the preferences of the 29 rightmost individuals uniformly distributed over the interval $[0, 1]$. (This is the proportional size of Republican and Democratic parties in the 2011–2012 US House.) Further assume that 1000 status quos are uniformly distributed across $[-1, 1]$.² For each q , I calculate the equilibrium proposal reported in **Table 1** and the utility differential for every legislator on every vote using a quadratic loss function. To introduce voting error, I add an idiosyncratic, normally distributed error $N(0, \sigma^2)$ to each legislator's utility differential for each vote. (The error therefore affects the voting behavior but not the equilibrium proposals.)

The roll call matrix is constructed with vote $y_{it} = 1$ if the shocked difference in the utility of voting yes versus the utility of voting no for legislator i on vote t is greater than 0, and 0 otherwise. Because the amount of mean zero voting error is unknown, I consider three possible scenarios: $\sigma = 0.4$ (as in Hirsch 2011), which produces a 95% confidence interval for the introduced error of $[-0.78, 0.78]$ in the $[-1, 1]$ policy space; $\sigma = 0.10$, i.e., errors of $[-0.19, 0.19]$; and $\sigma = 0.01$, i.e., errors of $[-0.02, 0.02]$. Let the roll calls generated using the predictions of **Table 1** be Y_{PAS} , Y_{NAS} , and Y_{MVT} to indicate the selection model used to generate the roll calls—votes on proposals with status quos in the respective gridlock interval are not observed—and let x_{PAS} , x_{NAS} , and x_{MVT} denote the resulting ideal point estimates from analyzing the respective roll call matrix using the

²The ratio of votes to legislators is nearly 20 to 1. In contrast, the ratio in the US House for 2009–2010 was roughly 4 to 1 (1,647 votes taken by the 435 members).

Bayesian item response model of Clinton et al. (2004a). Because of agenda-setting incentives, only Y_{MVT} contains 1,000 votes; Y_{PAS} and Y_{NAS} have fewer votes because some status quos are not acted upon. For comparison, I also generate a roll call matrix using 1,000 uniformly distributed cutpoints Y_U to estimate ideal points that reflect the effects of voting error and statistical normalizations but not the nonrandom selection of votes.

To evaluate the effects of the selection model and voting error, I generate 50 roll call matrices for each selection model and choice of σ , and I consider the relationship between the ideal point means using the 50 posterior means for each combination.³

Figure 2 plots the true ideal points (y-axis) against the mean of the 50 ideal point means estimated from the uniform cutpoint model (Uniform), the PAS model, the NAS model, and the median voter theorem (MVT) for mean-zero normal voting errors with standard deviations of 0.40 (top row), 0.10 (middle row) and 0.01 (bottom row). Several conclusions emerge.

First, it is trivially true that the assumed statistical normalizations affect the scale of the recovered estimates. Because the statistical model assumes that ideal points are distributed with mean 0 and variance 1 to identify the scale of the policy space, the estimated ideal points span the interval $[-3, 3]$ even though the true ideal points responsible for generating the data only span $[-1, 1]$. However, the difference in scaling is entirely arbitrary and inconsequential.

More critical is the fact that the relationship between the true ideal points and the estimated ideal points differs across both the rows and columns of **Figure 2**. Despite the fact that all of the estimates correlate very highly with the true preferences (the PAS estimate with voting errors with $\sigma = 0.01$ is the least correlated at 0.96), there are differences between the ideal points and true preferences, and these differences depend on both the selection model responsible for generating the data and the amount of voting error. Recall that the true ideal points are uniformly distributed in $[-1, 0]$ for the 22 leftmost members and $[0, 1]$ for the 29 rightmost members in every case. Consistent with expectations, the Uniform selection model—which assumes the proposals being voted on are exogenously and uniformly distributed across the policy space—accurately recovers the uniformly distributed true preferences regardless of the amount of voting error [although some distortion (i.e., the s-shape) occurs when the standard deviation of the voting error is 0.01].

The three other selection models show varying effects on the ability to recover the true preferences of the data-generating process. As Hirsch (2011) shows, with sufficient voting error (e.g., $\sigma = 0.4$ in the top row of **Figure 2**), the correlation between the estimated and true preferences is strong regardless of whether the data is generated according to a PAS, NAS, or MVT model. However, the precision of the estimated ideal points is still somewhat affected by the selection model because there are fewer predicted policies with extreme cutpoints.

The effect of the selection models is largest when the amount of voting error is low and voting is near-perfect (e.g., $\sigma = 0.01$ as in the bottom row of **Figure 2**). In the PAS model, for example, there is a clustering of members with true ideal points located in the interval $[X_M, X_S]$ as well as among extremists because of the lack of cutpoints in either region. Only the assumption of probabilistic voting distinguishes between these members. Unlike the case when $\sigma = 0.4$, the lack of variation in the voting error results in an inability to discriminate between members. A similar indeterminacy occurs in the NAS model when there is low voting error. Increasing the error standard deviation from 0.01 to 0.10 smoothes out the clustering of ideal points in the interior of the policy space owing to gridlocked status quos, but clustering is still evident among the extremists.

³For each characterization, I estimate 50 sets of ideal point estimates, and I use the mean estimate of the 50 posterior means to characterize a legislator's ideal point. For each analysis, I use ideal 1.041 (Jackman 2011) with 10,000 iterations (thinning by 10) and a burn-in of 1,000. Ideal points are normalized to have a mean 0 and variance 1 in each instance.

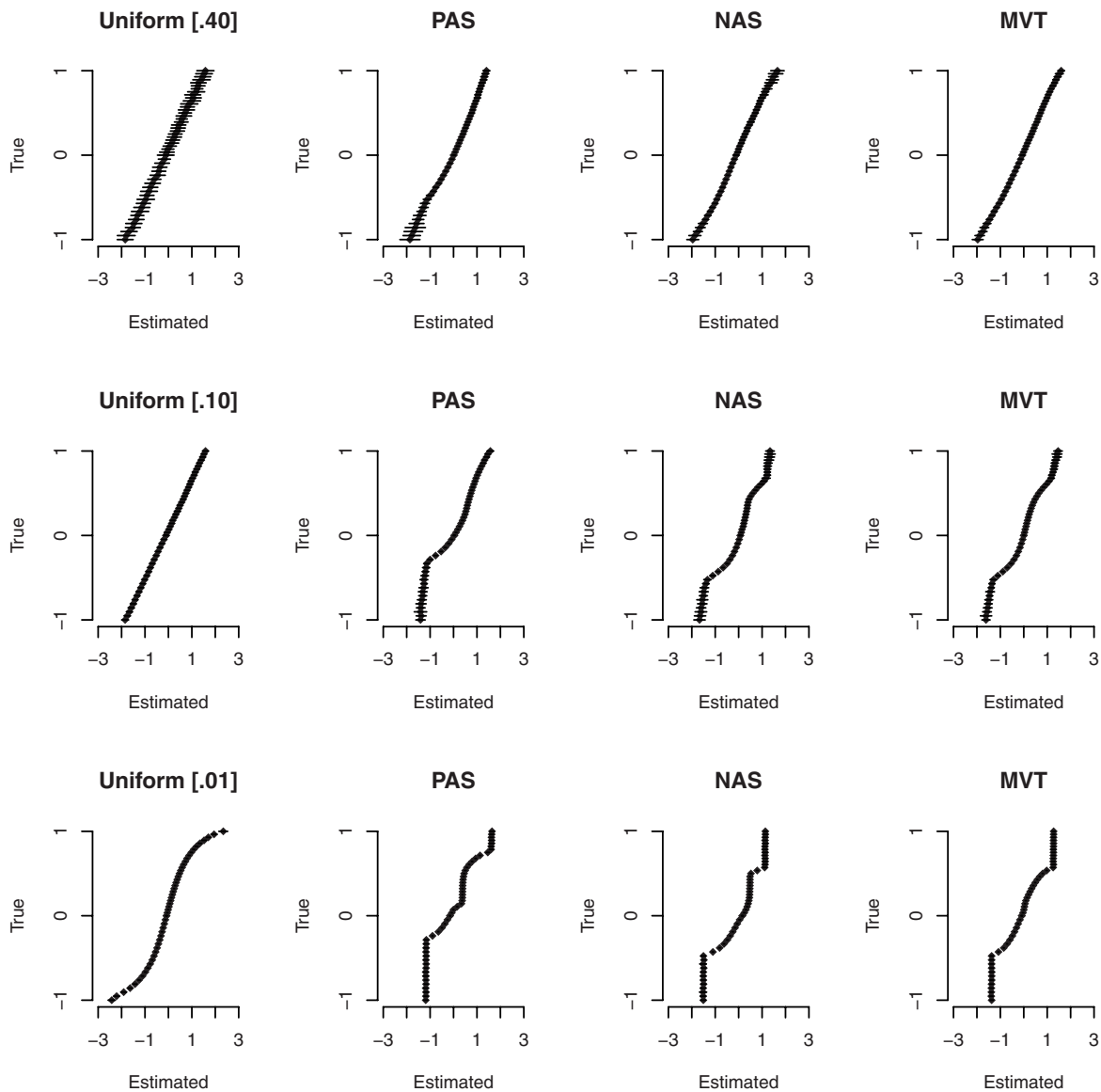


Figure 2

The effect of selection models and voting error on the relationship between estimated and true ideal points. The true ideal points, the distribution of status quos, and the normalization assumptions of the ideal point estimator are identical in all of the comparisons. The variance of the mean-zero voting error varies across rows, and the columns indicate the selection model used to generate the proposals: Uniform cutpoint, positive agenda setter (PAS), negative agenda setter (NAS), and median voter theorem (MVT).

Considering the case when $\sigma = 0.10$ illustrates how the clustering of ideal points may matter even when the estimates correlate with the true ideal points in excess of 0.98. For example, if we compute gridlock intervals, the estimated distance between X_M and X_S ranges from 0.37 (with a 95% Bayesian credible interval of [0.32, 0.47]) under the NAS model to 0.64 (with a 95% Bayesian credible interval of [0.55, 0.73]) under the uniform cutpoint model.

Overall, **Figure 2** reveals that if the amount of voting error is low, the nonrandom selection of votes can produce ideal points that fail to account for variation in the preferences used to generate the roll calls. As the level of voting error in the individual choice model increases, however, the clustering of ideal points evident in the bottom row of the figure due to the lack of cutpoints smooths out regardless of whether the roll calls are generated by the MVT, PAS, or NAS.

Although **Figure 2** reveals that the variation in ideal points due to the selection model depends on the amount of voting error, a difficulty is that we do not know either the amount of voting error or the true nature of the roll call selection when analyzing roll calls. Given the number of assumptions required to produce the estimates, not only may it be difficult to know if the amount of voting error is enough to overcome the potential selection issues, but it will also be difficult to make inferences about the process that generates the votes using the pattern of recovered estimates.

To further illustrate the effects of the selection model and voting error, the top row of **Figure 3** graphs the relationship between ideal points generated using a Uniform selection model and a NAS selection model, holding the amount of voting error fixed at the level noted above each column. Comparing X_{NAS} to X_U isolates the effect of the selection model on the recovered estimates by eliminating the effect of the voting error and statistical normalizations. As in **Figure 2**, the selection model can have strong effects depending on the amount of voting error—moving from left to right in the top row of **Figure 3** reveals that the relationship between the two sets of estimated ideal points varies if the standard deviation of the voting error is 0.10, 0.05, or 0.01. To give a sense of the magnitude of the voting errors, recall that a voting error with a standard deviation of 0.10 will produce utility shocks with a 95% confidence interval of $[-0.2, 0.2]$ in the $[-1, 1]$ policy space.

The bottom row of **Figure 3** reveals how the amount of variation in the voting error affects the estimated ideal points, holding the selection model fixed. Using only votes generated by the NAS selection model, the plots compare the average ideal points from the 50 simulations generated with varying amounts of mean-zero voting error. Comparing the leftmost relationship to the rightmost relationship reveals that as the variance of the voting error decreases, the clustering of legislator ideal points increases owing to the absence of cutpoints because of negative agenda setting.

Another potential complication involves the assumed distribution of status quos. Although the formal models of **Table 1** are static models in which Nature randomly chooses the status quo policy to be considered, the actual distribution of status quos is unknown and likely changes over time as status quo policies are actively modified by the political process (Krehbiel 1998). The simulations assume that the 1,000 status quos are uniformly distributed across the policy space, but the distribution of status quo policies can also obviously affect the estimated ideal points. Because the predicted equilibrium cutpoints in **Table 1** depend on both the equilibrium proposal p^* —i.e., the roll call selection model—and the location of the status quo q , changing either the selection model responsible for generating p^* or the distribution of q will change the distribution of cutpoints and, depending on the amount of voting error, the distribution of the estimated ideal points as well (Clinton 2007, Hirsch 2011, Woon & Cook 2011). To take an extreme example, if all status quos are located at a single point, the resulting roll calls will be unable to measure the distribution of underlying preferences because there will be no variation in proposals.

Some have tried to estimate the location of status quos. For example, Clinton (2012) uses assumptions about how legislators perceive the proposals being voted on the legislative agenda, Richman (2011) uses legislators' survey responses on the National Public Awareness Test to try to deduce where the status quo must be located given legislators' responses, and Peress (2011) uses data on cosponsorship behavior to infer where the status quo must be given the location of the legislators who are proposing changes. All three make strong assumptions, but knowing the location of the status quo is important because it would allow scholars to better estimate the

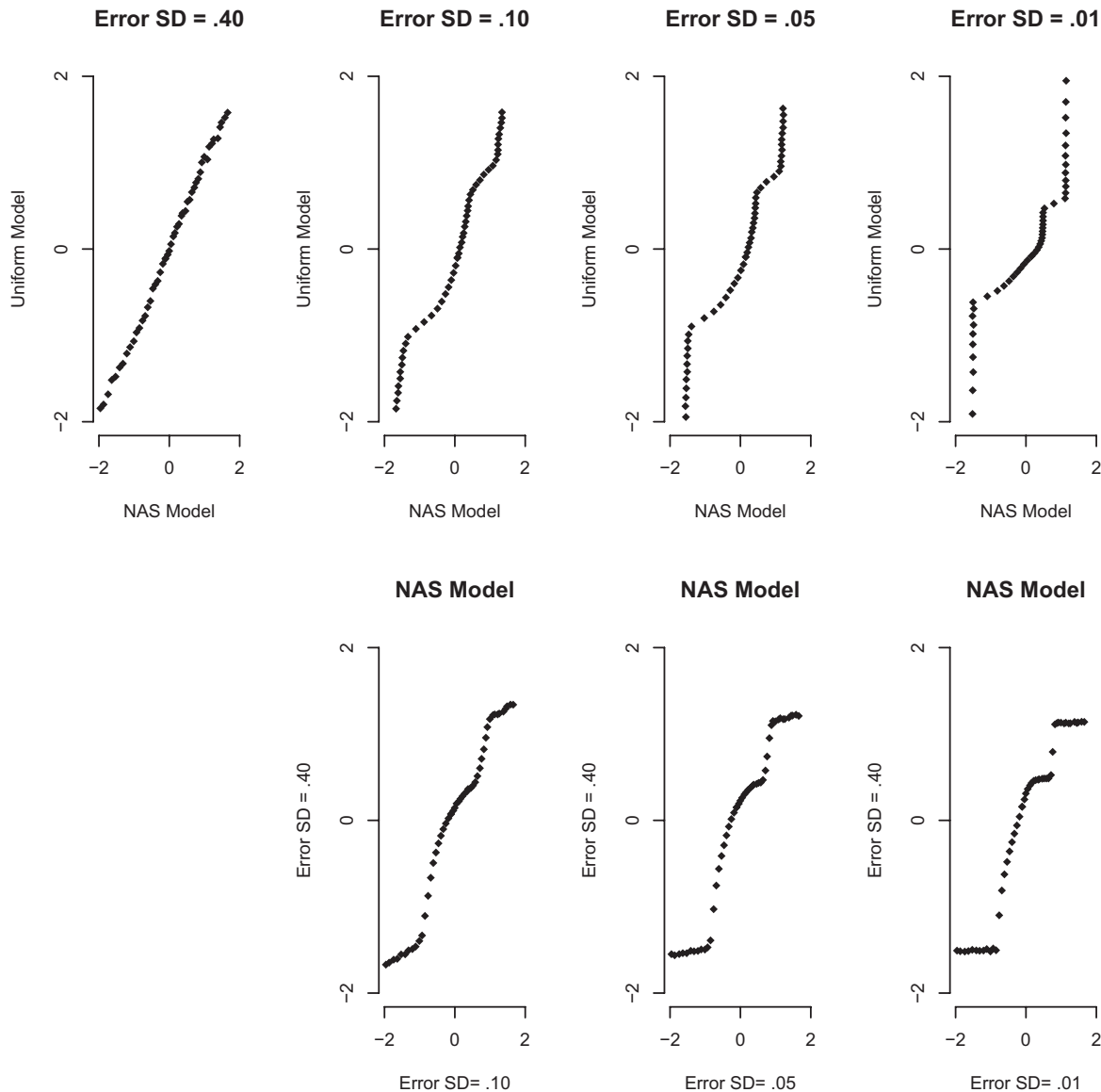


Figure 3

The effect of voting error on the estimated ideal point. The top row compares ideal points generated using the negative agenda-setter (NAS) selection model and the Uniform selection model for varying amounts of voting error. The bottom row compares ideal points generated using differing amounts of voting error using the selection model of the NAS. The true ideal points, the status quo distribution, and the normalization assumptions of the ideal point estimator are held fixed in each case. SD, standard deviation.

location of the proposal being voted on and test additional predictions given the q 's critical role in **Table 1**.

The general point is that even if the issues of the previous section are irrelevant, and formal and statistical models adopt identical models of individual choice, aspects such as the amount of voting error and the selection of votes may affect the extent to which the recovered estimates reflect the

true preferences. Simulations provide a way to characterize the possible consequences and illustrate how the relationship between the true and estimated ideal points can depend on the model of roll call selection, the amount of voting error, the assumed status quo distribution, and the statistical normalizations employed. Because the distortions evident in **Figures 2** and **3** are correlated with the true preferences—e.g., the clustering of ideal points in certain intervals in the policy space—if the estimates based on the observed voting behavior are used to characterize the true underlying preferences, the result can be inconsistent and misleading estimates (Herron 2001). The illustrated differences will certainly not always be consequential, but the results hopefully highlight the need to be sensitive to the conditions under which ideal points may not adequately measure policy preferences.

INTEGRATED APPROACHES TO TESTING FORMAL THEORIES USING IDEAL POINTS

Recent work attempts to better integrate the various models involved in the analysis of roll calls by deriving statistical models that directly estimate the parameters of theoretical interest, rather than first estimating ideal points and then using them in a secondary statistical analysis as is typically done in a reduced-form approach (e.g., Signorino 2003). The roll call estimators may be structural, in the sense of estimating parameters from theoretical models that are assumed to be true; examples include Reiss & Wolak's (2007) discussion of structural estimation in industrial organization and Morton's (1999) notion of a "complete model." Alternatively, the statistical model may be more closely tailored to the particular theoretical model of interest without basing the estimator directly on the assumed game form. Either way, the goal is to derive estimators of the data generating process based more closely on the model being tested.

One possibility is to use the game form to derive the likelihood function and estimate parameters of the model assuming that the equilibrium is true. Although there has not been much of this so-called structural estimation in political science, some recent work has begun to take this approach (e.g., Iaryczower & Shum 2012, Iaryczower et al. 2011). Given the nature of the estimation, theory testing in such circumstances involves evaluating the ability of the model to fit the data (typically using out-of-sample predictions).

More common in political science is the practice of using implications from formal models of either the individual or collective choice environment to provide additional structure when analyzing roll calls. Londregan (2000a,b), for example, extends the spatial voting model to include valence considerations in the individual voting model to analyze legislators' votes in Chile, and McCarty et al. (2001) modify the model of individual choice to allow for the possibility of party influence in the US House.

Clinton & Meirowitz (2001, 2003, 2004) do not modify the individual choice model assumed by common models of roll call voting. Instead, they argue that additional information can be incorporated in the statistical estimation using parameter constraints based on assumptions about how legislators perceive the proposals being voted on. If the agenda can be treated as fixed and legislators anticipate the effects of their decisions on subsequent outcomes, a set of equality constraints can be used to better identify the location of the policies being voted on. Scholars have used similar approaches to investigate the politics of particular incidents (e.g., Jeong 2008, Jeong et al. 2009, Pope & Treier 2011).

A third approach is to derive testable predictions directly in terms of the observed votes or estimated parameters (e.g., Bianco & Sened 2005, Krehbiel et al. 2005, Patty 2008, Stiglitz & Weingast 2010, Woon & Cook 2011, Clinton 2012). Such an approach avoids the potential complications posed by the nonrandom selection of roll calls because the analysis focuses directly

on the question of which votes are selected and the pattern of resulting estimates. Focusing on roll rates [i.e., the percentage of times that a majority of a party votes together on a bill and is defeated (Cox & McCubbins 2005)] is one variant of this research strategy—the question being asked is whether the votes that are selected for consideration are consistent with a particular model of politics.⁴ A difficulty with this approach is that the number of unknown parameters may make it hard to reach definitive conclusions. For example, given the similarity of ideal points plotted in **Figure 2** across columns, it may be difficult to determine which selection model is most likely responsible for generating a pattern of ideal points if the distribution of true preferences and status quos and the amount of voting error are all unknown.

Although issues remain, integrated approaches minimize the possibility that theoretical predictions are rejected because of differences in the models being assumed in the theoretical and empirical analyses. Such approaches allow us to test theoretical predictions under the most favorable conditions for theory, but an obvious downside is that the resulting analyses are limited to particular models or particular incidents. Conducting the analysis also requires working with both formal and statistical models to evaluate the consequences of the decisions made when estimating the parameters responsible for generating the observed behavior.

CONCLUSION

Votes in institutions across time and around the globe are routinely analyzed in the hope of gaining deeper insights into the nature of political processes and interactions. A vibrant community of scholars has collected a vast amount of data and developed many methods for analyzing observed roll call votes. Analyzing roll calls is important because of the potential to measure the elusive but important preferences of political elites and to use the measures to better understand the determinants and consequences of elite behavior. When used to characterize legislators' public positions over a series of votes, scholars face almost no barriers—the interpretation is relatively clean, and few statistical or computational issues remain.

Using ideal points to measure true preferences or assess the predictions of formal models is more complicated. Independent of the philosophical debate over the feasibility and value of testing formal models, difficulties can emerge because the statistical and theoretical models are simplified representations of the political process. Depending on the context and application, the models may be inconsistent in what they assume. Moreover, the ideal points estimated from observed roll calls cannot necessarily be treated as measuring legislators' true preferences because the estimated ideal points are the set of points that rationalize the revealed—presumably equilibrium—behavior of legislators given the assumed model of individual choice.

Not every issue I have discussed will be consequential for every use of ideal points, and some omitted aspects may also be significant, but at the very least it is important to continue to consider possible complications to increase the fidelity of empirical assessments of the predictions of formal models. If the models in the empirical and theoretical analysis conflict, it may be difficult to interpret evidence for or against theoretical predictions based on the statistical analysis.

I have focused on the analysis of theoretical predictions using roll call votes, but it is important to emphasize that similar concerns emerge whenever we analyze observed behavior in the hope of measuring true policy preferences. Statistical models of sponsorship and cosponsorship decisions, campaign contributions, and political speeches that aspire to recover the participants' underlying

⁴However, Krehbiel (2007) argues that the results can be explained by a contrary theory with observationally equivalent predictions.

policy preferences all attempt to interpret revealed behavior according to an assumed model of individual choice (e.g., Woon 2008). As is the case when analyzing roll calls, we need to understand how the assumed model of individual choice and possible selection issues relates to the models generating the theoretical predictions of interest. Without that understanding, it will be difficult to interpret the resulting estimates.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I thank the participants at the Empirical Implications of Theoretical Models conference at the Harris School of Public Policy, University of Chicago, for providing a venue for a public airing of portions of this review.

LITERATURE CITED

- Ansolabehere SD, Snyder JM, Stewart C. 2001. Candidate positioning in US House elections. *Am. J. Polit. Sci.* 45:136–59
- Bafumi J, Gelman A, Park D, Kaplan N. 2005. Practical issues in implementing and understanding Bayesian ideal point estimation. *Polit. Anal.* 13:171–87
- Bafumi J, Herron MD. 2010. Leapfrog representation and extremism: a study of American voters and their members in Congress. *Am. Polit. Sci. Rev.* 104:519–42
- Bailey MA. 2007. Comparable preference estimates across time and institutions for the court, Congress and presidency. *Am. J. Polit. Sci.* 51:433–48
- Bartels LB. 1991. Constituency opinion and congressional policy making: the Reagan defense buildup. *Am. Polit. Sci. Rev.* 85:457–74
- Bertelli A, Grose CR. 2011. The lengthened shadow of another institution? Ideal point estimates for the executive branch and Congress. *Am. J. Polit. Sci.* 55:767–81
- Bianco WT, Jeliaskow I, Sened I. 2004. The uncovered set and the limits of legislative action. *Polit. Anal.* 12:256–76
- Bianco WT, Sened I. 2005. Uncovering evidence of conditional party government: reassessing majority party influence in Congress and state legislatures. *Am. Polit. Sci. Rev.* 99:361–71
- Binder SA. 1999. The dynamics of legislative gridlock, 1947–96. *Am. Polit. Sci. Rev.* 93:519–33
- Bonica A. 2011. *Ideology and interests in the political marketplace*. Work. pap., Dep. Polit. Sci., Stanford Univ.
- Carroll R, Lewis JB, Lo J, Poole KT, Rosenthal H. 2009. Comparing NOMINATE and IDEAL: points of difference and Monte Carlo tests. *Legis. Stud. Q.* 34:555–91
- Chen J, Johnson T. 2011. *Federal unionization and presidential control of the bureaucracy*. Work. pap., Dep. Polit. Sci., Univ. Mich.
- Chiou FY, Rothenberg L. 2003. When pivotal politics meets with partisan politics. *Am. J. Polit. Sci.* 47:503–22
- Clarke KA, Primo DM. 2007. Modernizing political science: a model-based approach. *Perspect. Polit.* 5:741–53
- Clinton JD. 2007. Lawmaking and roll calls. *J. Polit.* 69:455–67
- Clinton JD. 2012. Congress, lawmaking and the Fair Labor Standards Act, 1971–2000. *Am. J. Polit. Sci.* In press
- Clinton JD, Jackman S. 2009. To simulate or NOMINATE? *Legis. Stud. Q.* 34:593–622
- Clinton JD, Jackman S, Rivers RD. 2004a. The statistical analysis of roll call voting: a unified approach. *Am. Polit. Sci. Rev.* 98:355–70
- Clinton JD, Jackman S, Rivers RD. 2004b. The most liberal senator? Analyzing and interpreting congressional roll calls. *PS: Polit. Sci. Polit.* 37:805–11

- Clinton JD, Lapinski J. 2008. Laws and roll calls in the US Congress, 1889–1994. *Legis. Stud. Q.* 33:511–542
- Clinton JD, Meirowitz A. 2001. Agenda constrained legislator ideal points and the spatial voting model. *Polit. Anal.* 9:242–60
- Clinton JD, Meirowitz A. 2003. Integrating voting theory and roll call analysis: a framework. *Polit. Anal.* 11:381–96
- Clinton JD, Meirowitz M. 2004. Testing accounts of legislative strategic voting: the compromise of 1790. *Am. J. Polit. Sci.* 48:675–89
- Cohen LR, Noll RG. 1991. How to vote, whether to vote: strategies for voting and abstaining on congressional roll calls. *Polit. Behav.* 13:97–127
- Cox GW, McCubbins MD. 1993. *Legislative Leviathan: Party Government in the House*. Berkeley: Univ. Calif. Press. 324 pp.
- Cox GW, McCubbins MD. 2005. *Setting the Agenda: Responsible Party Government in the US House of Representatives*. New York: Cambridge Univ. Press. 326 pp.
- Cox GW, Poole KT. 2002. On measuring partisanship in roll-call voting: the US House of Representatives, 1977–1999. *Am. J. Polit. Sci.* 46:477–89
- Desposato SW, Kearney MC, Crisp BF. 2011. Using cosponsorship to estimate ideal points. *Legis. Stud. Q.* 36:531–65
- Diermeier D, Krehbiel K. 2003. Institutionalism as a methodology. *J. Theor. Polit.* 15:123–44
- Enelow JM, Hinich MJ. 1984. *The Spatial Theory of Voting: An Introduction*. Cambridge, UK: Cambridge Univ. Press. 238 pp.
- Evans CL. 2008. *The House Whip process and the textbook Congress*. Work. pap., Dep. Gov., College of William and Mary
- Forgette R, Sala B. 1999. Conditional party government and member turnout on Senate recorded votes, 1873–1935. *J. Polit.* 61:467–84
- Glazer A, Griffin R, Grofman B, Wattenberg M. 1995. Strategic vote delay in the US House of Representatives. *Legis. Stud. Q.* 20:37–45
- Granato J, Lo M, Sunny Wong MC. 2010. A framework for unifying formal and empirical analysis. *Am. J. Polit. Sci.* 54:783–97
- Groseclose T, Milyo J. 2010. Sincere versus sophisticated voting in Congress: theory and evidence. *J. Polit.* 72:1–14
- Heckman JJ. 1979. Sample selection bias as a specification error. *Econometrica* 47:153–61
- Heckman JJ, Snyder JM Jr. 1997. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *Rand J. Econ.* 28:142–89
- Herron MC. 1999. Artificial extremism in interest group ratings and the preferences versus party debate. *Legis. Stud. Q.* 24:525–42
- Herron MC. 2001. Interest group ratings and regression inconsistency. *Polit. Anal.* 9: 260–74
- Hirsch A. 2011. Theory driven bias in ideal point estimates—a Monte Carlo study. *Polit. Anal.* 19:87–102
- Hug S. 2010. Selection effects in roll call votes. *Br. J. Polit. Sci.* 40:225–35
- Iaryczower M, Lewis G, Shum M. 2011. *To elect or to appoint? Bias, information, and responsiveness of bureaucrats and politicians*. Work. pap., Dep. Polit., Princeton Univ.
- Iaryczower M, Shum M. 2012. The value of information in the court. Get it right, keep it tight. *Am. Econ. Rev.* In press
- Jackman S. 2011. *pscl: classes and methods for R developed in the Political Science Computational Laboratory, Stanford University*. Dep. Polit. Sci., Stanford Univ. Version 1.04.1.
- Jackson JE, Kingdon JW. 1992. Ideology, interest group scores, and legislative votes. *Am. J. Polit. Sci.* 36:805–23
- Jenkins JA, Munger MC. 2003. Investigating the incidence of killer amendments in Congress. *J. Polit.* 65:498–517
- Jeong GH. 2008. Testing the predictions of the multidimensional spatial voting model with roll call data. *Polit. Anal.* 16:179–96
- Jeong GH, Miller G, Sened I. 2009. Closing the deal: negotiating civil rights legislation. *Am. Polit. Sci. Rev.* 103:588–606

- Jones DR. 2003. Position taking and position avoidance in the US Senate. *J. Polit.* 65:851–63
- Kalandrakis T. 2010. Rationalizable voting. *Theor. Econ.* 5:93–125
- Kastellec JP, Lax JR. 2008. Case selection and the study of judicial politics. *J. Empir. Legal Stud.* 5:407–46
- Krehbiel K. 1996. Institutional and partisan sources of gridlock: a theory of divided and unified government. *J. Theor. Polit.* 8:7–40
- Krehbiel K. 1998. *Pivotal Politics: A Theory of US Lawmaking*. Chicago: Univ. Chicago Press. 274 pp.
- Krehbiel K. 2000. Party discipline and measures of partisanship. *Am. J. Polit. Sci.* 44:206–21
- Krehbiel K. 2007. Partisan roll rates in a non-partisan legislature. *J. Law Econ. Organ.* 23:1–23
- Krehbiel K, Meirowitz A, Woon J. 2005. Testing theories of lawmaking. In *Social Choices and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, ed. D Austen-Smith, J Duggan, pp. 249–68. New York: Springer
- Ladha KK. 1991. A spatial model of legislative voting with perceptual error. *Public Choice* 68:151–74
- Lauderdale BE. 2010. Unpredictable votes in ideal point estimation. *Polit. Anal.* 18:151–71
- Laver M, Benoit K, Garry J. 2003. Extracting policy positions from political texts using words as data. *Am. Polit. Sci. Rev.* 97:311–31
- Londregan JB. 2000a. Estimating legislators' preferred points. *Polit. Anal.* 8:35–56
- Londregan JB. 2000b. *Legislative Institutions and Ideology in Chile*. New York: Cambridge Univ. Press. 280 pp.
- Lowell AL. 1902. The influence of party upon legislation in England and America. *Annu. Rep. Am. Hist. Assoc.* 1:321–542
- Martin AD, Quinn KM. 2002. Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999. *Polit. Anal.* 10:134–53
- McCarty NM, Meirowitz A. 2007. *Political Game Theory: An Introduction*. New York: Cambridge Univ. Press. 431 pp.
- McCarty NM, Poole KT. 1995. Veto power and legislation: an empirical analysis of executive and legislative bargaining from 1961–1986. *J. Law Econ. Organ.* 11:282–312
- McCarty NM, Poole KT, Rosenthal H. 2001. The hunt for party discipline in Congress. *Am. Polit. Sci. Rev.* 95:673–87
- McCarty NM, Poole KT, Rosenthal H. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: MIT Press. 256 pp.
- Morton RB. 1999. *Methods and Models: A Guide to the Empirical Analysis of Formal Models in Political Science*. Cambridge, UK: Cambridge Univ. Press. 340 pp.
- Patty JW. 2008. Equilibrium party government. *Am. J. Polit. Sci.* 52:636–55
- Peress M. 2011. *Estimating proposal and status quo locations using voting and cosponsorship data*. Work. pap., Dep. Polit. Sci., Univ. Rochester
- Poole KT. 2000. Non-parametric unfolding of binary choice data. *Polit. Anal.* 8:211–37
- Poole KT. 2005. *Spatial Models of Parliamentary Voting*. New York: Cambridge Univ. Press. 230 pp.
- Poole KT, Rosenthal H. 1985. A spatial model for legislative roll call analysis. *Am. J. Polit. Sci.* 29:357–84
- Poole KT, Rosenthal H. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford Univ. Press. 297 pp.
- Poole KT, Rosenthal H. 2007. *Ideology and Congress*. New Brunswick, NJ: Transaction. 344 pp.
- Pope J, Treier S. 2011. Reconsidering the Great Compromise at the Federal Convention of 1787: deliberation and agenda effects on the Senate and slavery. *Am. J. Polit. Sci.* 55:289–306
- Popper K. 1961. *The Logic of Scientific Discovery*. New York: Sci. Ed. 544 pp.
- Reiss PC, Wolak FA. 2007. Structural econometric modeling: rationales and examples from industrial organization. In *Handbook of Econometrics*, 6A:4280–415. Amsterdam: North-Holland
- Richman J. 2011. Parties, pivots, and policy: the status quo test. *Am. Polit. Sci. Rev.* 105:151–65
- Riker WH. 1980. Implications from the disequilibrium of majority rule for the study of institutions. *Am. Polit. Sci. Rev.* 74:432–46
- Roberts JR. 2007. The statistical analysis of roll call data: a cautionary tale. *Legis. Stud. Q.* 32:341–60
- Roberts JR, Bell LC. 2008. Scoring the Senate: scorecards, parties and roll-call votes. In *Why Not Parties? Party Effects in the United States Senate*, ed. NW Monroe, JM Roberts, DW Rohde, pp. 52–70. Chicago: Univ. Chicago Press
- Roberts JR, Smith SS. 2003. Procedural contexts, party strategy, and conditional party voting in the US House of Representatives, 1971–2000. *Am. J. Polit. Sci.* 47:305–17

- Romer T, Rosenthal H. 1978. Political resource allocation, controlled agendas and the status quo. *Public Choice* 33:27–44
- Rosenthal H, Voeten E. 2004. Analyzing roll calls with perfect spatial voting. *Am. J. Polit. Sci.* 48:620–32
- Signorino CS. 2003. Structure and uncertainty in discrete choice models. *Polit. Anal.* 11:316–44
- Smith SS. 2007. *Party Influence in Congress*. New York: Cambridge Univ. Press. 254 pp.
- Snyder JM Jr. 1991. On buying legislatures. *Econ. Polit.* 3:93–109
- Snyder JM Jr. 1992a. Artificial extremism in interest group ratings. *Legis. Stud. Q.* 17:319–45
- Snyder JM Jr. 1992b. Committee power, structure-induced equilibria, and roll call votes. *Am. J. Polit. Sci.* 36:1–30
- Snyder JM Jr. 1992c. Gatekeeping or not, sample selection in the roll call agenda matters. *Am. J. Polit. Sci.* 36:36–39
- Snyder JM Jr, Groseclose T. 2000. Estimating party influence in congressional roll-call voting. *Am. J. Polit. Sci.* 44(2):193–211
- Snyder JM Jr, Ting M. 2003. Roll calls, party labels, and elections. *Polit. Anal.* 11:419–44
- Snyder JM Jr, Ting M. 2005. Why roll calls? A model of position-taking in legislative voting and elections. *J. Law Econ. Organ.* 21:153–78
- Snyder JM Jr, Ting M. 2011. Roll calls, party labels, and elections. *Polit. Anal.* 11:419–44
- Spirling A, McLean I. 2007. UK OC OK? Interpreting optimal classification scores for the UK House of Commons. *Polit. Anal.* 51:85–96
- Stiglitz EH, Weingast BR. 2010. Agenda control in Congress: evidence from cutpoint estimates and ideal point uncertainty. *Legis. Stud. Q.* 35:157–85
- Treier S. 2010. Where does the president stand? Measuring presidential ideology. *Polit. Anal.* 18:124–36
- Treier S. 2011. Comparing ideal points across institutions and time. *Am. Polit. Res.* 39:804–31
- Van Doren PM. 1990. Can we learn the causes of congressional decisions from roll call data? *Legis. Stud. Q.* 15:311–40
- Voeten E. 2000. Clashes in the assembly. *Int. Organ.* 54:185–215
- Wiseman AE, Herron MC. 2008. Gerrymanders and theories of law making: a study of legislative redistricting in Illinois. *J. Polit.* 70:151–67
- Wiseman AE, Wright JR. 2008. The legislative median and partisan policy. *J. Theor. Polit.* 20:5–29
- Woon J. 2008. Bill sponsorship in Congress: the moderating effect of agenda positions on legislative proposals. *J. Polit.* 70:201–16
- Woon J, Cook IP. 2011. *Competing gridlock models, status quo assumptions, and legislative productivity*. Work. pap., Dep. Polit. Sci., Univ. Pittsburg