Testing Lawmaking Theories with (Endogenous) Roll Calls, 90th - 106th U.S. House *

Joshua D. Clinton
Department of Politics
Princeton University
Princeton, NJ 08544-1012
clinton@princeton.edu

Version 2.5
February 15, 2006

Abstract

Theories of lawmaking generate predictions about legislative behavior and policy outcomes. Tests of theoretical predictions often use roll calls and roll call based measures such as ideal point estimates. This paper argues that the endogeneity of roll calls poses problems for testing theories of lawmaking because predictions regarding equilibrium outcomes necessarily yield predictions about the agenda used to enact the outcomes. Using two prominent lawmaking theories, I provide a solution for the problem posed by the endogeneity of roll calls and I illustrate how the predictions from the theories can be embedded within statistical models used to estimate ideal points (for a subset of the roll call record). Accounting for the endogeneity of the roll call record in the contemporary U.S. House (90th - 106th) reveals lackluster support for either theory.

*The author thanks Keith Krehbiel, David Lewis, John Londregan, Nolan McCarty, Adam Meirowitz, Keith Poole, Eric Schickler and participants at Princeton’s Center for the Study of Democratic Politics for helpful comments.
Estimates derived from roll call behavior are prominently used to test theories of lawmaking. Interest group scores and ideal point estimates (hereafter ideal points) such as NOMINATE (Poole and Rosenthal 1997) appear to offer the ability to achieve such diverse and critically important tasks as: measuring the width of variously defined “gridlock” intervals (e.g., Segal 1997; Krehbiel 1998; Coleman 1999; Epstein and O’Halloran 1999; Chiu and Rothenberg 2003; Covington and Barger 2004; Krehbiel, Meirowitz and Woon 2005; Cox and McCubbins 2005), presidential veto intervals (McCarty and Poole 1995), constraints on congressional decision-making imposed by the Supreme Court (Martin 2001), party heterogeneity (e.g., Binder 1999; 2003, and Schickler 2000) and the extremity of congressional committees (e.g., Kiewiet and McCubbins 1991; Cox and McCubbins 1993; Londregan and Snyder 1994; Dion and Huber 1996). As this impressive list suggests, many prominent tests of lawmaking rely on roll call estimates to measure theoretically implied concepts.

A fundamental problem in using roll calls to test theories of lawmaking is that roll calls – and therefore measures derived from roll calls such as ideal point estimates – are not exogenous to the theories they are commonly used to test. Almost every existing use of roll calls to test lawmaking theories fails to account for the fact that in generating predictions about congressional lawmaking, the theories necessarily generate predictions about member behavior on the roll calls enacting the predicted policies. The empirical implications of theoretical models of lawmaking are more profound than currently realized and divorcing the statistical and theoretical models risks the potential for incorrect characterizations and inferences.

Although the endogeneity of the observed roll call record has been noted previously (e.g., Snyder 1992; 1992b, Shepsle and Weingast 1994), existing work fails to fully account for the consequences when using roll call votes and ideal point estimates. Prior work may recognize the potential problems caused by the endogeneity of the roll call record, but a solution has yet to emerge. Consequently, some have argued that roll call measures cannot sensibly be used to test theories of lawmaking – Shepsle and Weingast (1994) note that ideal points “provide an inappropriate basis for testing the hypothesis of party differentiation...[because]..these scores are...endogenous to the legislative context” (pg. 173). Others proceed without addressing the problem – regressing measures of legislative accomplishment, for example, on measures constructed from ideal points.

\footnote{For example, Snyder (1992; 1992b) raises the question of whether ideal points are largely one dimensional because of committee gatekeeping – “Whatever forces shape the floor agenda, however, this agenda is likely to be skewed in favor of some types of bills and issues and against others, and as a results, roll call voting may not accurately reveal legislators’ true preferences” (p. 37).}
The endogeneity of roll calls does not mean that roll calls and ideal point estimates cannot be used for theory testing. Instead, it suggests that their proper use requires accounting for the politics by which the roll call agenda is realized. Problems resulting from the fact that ideal points are not exogenous can be solved by integrating the theoretical predication within the statistical voting model.\textsuperscript{2} For (at least) two prominent classes of lawmaking theories frequently subjected to tests using ideal points, accounting for the endogeneity of roll calls yields testable predictions directly in terms of the estimated ideal points.

Considering a prominent debate within the legislative organization literature shows how formal models of gatekeeping with majority-party gatekeepers (Cox and McCubbins 1993; 2005) or committee gatekeepers (Shepsle 1979; Shepsle and Weingast 1987; Dion and Huber 1996; Cox 2001) and models with veto players such as the president (e.g., McCarty and Poole 1995) or pivotal actors in Congress (Krehbiel 1998) yield very simple tests for determining the validity of the theory once the connections between the statistical model used to analyze roll call behavior and lawmaking theories are realized. The ability to calculate a “gridlock interval” using ideal points constitutes evidence either that the theoretical prediction is false or else that measures are based upon roll calls outside the scope of the theory and therefore irrelevant for assessing the theory’s validity.

Although I implement the proposed solution and assess the support for two prominent theories of lawmaking in the 90th - 106th U.S. Houses, the relevance of the paper’s argument of the paper extends beyond the congressional politics literature. The conclusions apply to the tests of any related theories conducted on a parliament, court or deliberative body that implements outcomes via recorded votes (e.g., Schulz and Konig’s (2000) investigation of gridlock in the European Union, Londregan’s (2000) investigation of presidential agenda control in Chile or Segal’s (1997) exploration of the impact of the separation of powers on judicial decision-making).

The paper proceeds as follows. Section 1 briefly discusses the differences between preferences as used in formal models and ideal points from roll call analysis and Section 2 outlines two prominent theoretical models that ideal points are frequently used to test. Section 3 reviews a statistical model of voting behavior and derives predictions about distribution of estimated ideal points for the theories discussed in section 2. Section 4 implements the tests derived in Section 3 using the

\textsuperscript{2}Attempts have recently been made to integrate theoretical predictions and statistical roll call models in other applications. For example, Groseclose and Snyder (2000) and McCarty, Poole and Rosenthal (2001) examine party influence, Londregan (2000; 2000b) provides an application to proposal behavior, and Clinton and Meirowitz (2003; 2004) assess claims regarding strategic voting.
contemporary U.S. House and discusses unresolved issues. Section 5 discusses possible objections and section 6 concludes.

1 Preferences and Ideal Points

Most formal theories of Congressional action assume members possess single-peaked (typically Euclidean) preferences (e.g., Enelow and Hinich (1984)). These preferences may be explicitly defined over policy outcomes or they may include both outcome and position-taking considerations; they may reflect purely personal preferences or they may represent re-election constituency preferences. Preferences in the context of formal models are primitives independent of actions.

Because members’ preferences are unobserved, scholars are forced to try to measure preferences using observed actions. Roll call votes are one of the most common (and longest used) measures in congressional scholarship; Rice proposed using roll calls in political science as early as 1925 and examining subsequent scholarship reveals their prominent and ubiquitous use (see, for example, Poole’s (2005) recent treatise). In exchange for making some assumptions about member behavior, ideal point estimates offer the possibility of using roll calls to measure members’ preferences.

A serious problem in using ideal points is that they are endogenous to lawmaking theories. In generating predictions about equilibrium policy outcomes, lawmaking theories also predict which roll calls should be observed to enact those outcomes. Because lawmaking theories generate predictions about the roll call agenda and member behavior on that agenda, using voting behavior to generate preference estimates to test lawmaking theories without accounting for the politics by which the roll calls are generated may lead to erroneous substantive conclusions.

The proper use of ideal point estimates depends critically on whether the theory being studied has implications for the set of roll call votes that should be observed. Accounting for the endogeneity of roll calls for lawmaking theories requires dramatically different (and simpler) tests than are currently used. In fact, lawmaking theories typically tested with ideal points yield testable implications directly in terms of the estimated ideal points.

An important caveat is that these concerns are inconsequential for some uses of ideal points. Most notably, if behavior on roll call votes is of interest rather than the latent preference structuring a member’s roll call behavior – as is the case for testing theories about voting behavior (e.g., Snyder and Ting (2003) and the literature on shirking (see, for example, Rothenberg and Sanders (2000)
for a recent contribution) then neither concern is problematic. Ideal points are arguably exactly the right measure to assess accounts of member position-taking because the measure of interest is observed voting behavior (Mayhew 1974; Arnold 1990).

2 Selected Theories of Lawmaking

A debate of some ferocity in the Congressional scholarship concerns the nature of lawmaking. A prevalent argument (see, for example Rohde (1991), Aldrich (1995), Cox and McCubbins (1993; 2005)) is that the majority party leadership in the contemporary House exercises substantial influence through its ability to control the agenda. Although the details vary, Cox and McCubbins (1993; 2005) provide one of the most complete articulations. The argument is that members have correlated electoral fates by virtue of their party affiliation and that these correlated fates offer the possibility of collective electoral gain (or loss). To solve the resulting collective action problem, party members endow a party leader with the ability to punish members who act contrary to party interests through committee assignments and other institutional prerogatives. The party leader undertakes costly monitoring and organizational activities because of the rents resulting from agenda control; in return for solving the collective action problem the leadership can determine the agenda.

A competing claim is that the gains from collective action are insufficient to motivate party members to maintain cohesion (see, for example Mayhew 1974). In combination with the fact that congressional institutions themselves are adopted by majority rule and therefore endogenous (Krehbiel’s (1992) Majoritarian Postulate) - the result is a more individualistic environment. Policymaking depends on the preferences of individual members and institutional features such as the possibility of a filibuster and a presidential veto (Krehbiel 1998; Brady and Volden 1998).

A fair characterization of the empirical literature is that a consensus has yet to emerge as to which account, if either, is best supported (Smith (2000) provides a nice summary). Although by no means exhaustive, Cox and McCubbins (1993;2005), Cox and Poole (2002) and Lawrence, Maltzman and Smith (2006) find support for the partisan theory, Krehbiel (1998), Schickler (2000), Krehbiel and Wiseman (2001), Wawro and Schickler (2004) find support for the majoritarian theory, and Krehbiel, Meirowitz and Woon (2005) find little support for either. As none of this work fully accounts for the problems created by endogeneous roll calls, it is unknown whether the conflicting
results are due to methodological rather than theoretical errors.\(^3\)

To make the comparison between the theoretical predictions and the analysis of roll call behavior explicit, I consider simple versions of the two accounts. The theoretical results are not novel, but it is useful to present the results to illuminate their connection to ideal point estimates.

### 2.1 Theoretical Predictions: Gate-Keeping Theories

Assume a unidimensional policy space \(X\). Members have single-peaked preferences in \(X\) that are defined in terms of spatial proximity and member \(i\)'s most preferred outcome/position is \(x_i \in X\). Consider gatekeeper \(g\) with ideal point \(x_g\) and the pivot required for passage – the median member in the case of a majority voting rule – \(m\) with ideal point \(x_m\). For simplicity, suppose \(x_m < x_g\).

The theoretical result is well-known and it has been widely adopted in the congressional scholarship to model both committee (see, for example, Shepsle 1979; Denzau and Mackay 1983; Shepsle and Weingast 1987; Snyder 1992a; Crombez, Groseclose and Krehbiel 2005) and majority party (see, for example, Cox and McCubbins (1993; 2005); Aldrich 1995) agenda control. The play of the game is as follows:

- Nature draws a status quo policy \(q \in X\) that is perfectly observed by all.
- Gatekeeper \(g\) decides to allow or disallow action on \(q\). If action is allowed, play continues. If action is disallowed, status quo policy \(q\) is realized.
- If gatekeeper \(g\) allows action, the chamber median \(m\) makes chooses the proposal \(p \in X\) to enact into law.

The Nash Equilibria to this game consists of a set of decisions by the gatekeeper and a set of proposals by the chamber median. The latter is trivial – under majority rule the chamber median is decisive and proposes her ideal point \(p = x_m\) in every instance. The set of equilibrium actions by the gatekeeper is likewise trivial. Since any status quo for which action is allowed results in the policy \(p = x_m\), the gatekeeper only permits actions for which \(x_m\) is closer to \(x_g\) than \(q\). Consequently, if \(x_m < x_g\), inaction should be observed for any realization of \(q \in [x_m, 2x_g - x_m]\). The gatekeeper allows action on all other realizations. Figure 1A denotes the set of unique equilibrium policy

\(^3\)Ideal point estimates are clearly not the only, or even the most important measure used to test lawmaking theories. For example, Cox and McCubbins (1993) randomly sample 100 reported bills from the 82nd, 83rd, 92nd and 97th and determine whether a majority party member sponsored the bill (p. 260). Krehbiel (1998) uses coalition sizes, and Cox and McCubbins (2005) use roll rates and direction of policy change. However, they are certainly prevalently used measures.
outcomes $p^*$ for each realization of the status quo.\textsuperscript{4}

Given an equilibrium prediction $p^*$ and a status quo $q$, the cutpoint for the enacting vote – defined as the point at which a member is indifferent to the two outcomes – is $|p^* + q|/2$.\textsuperscript{5} The set of equilibrium cutpoints for this game is given in Figure 1B. Cutpoints are undefined in the region of $[x_m, 2x_g - x_m]$ because gatekeeping prevents votes being taken on these status quos.

\subsection*{2.2 Theoretical Predictions: Veto Theories}

Models of gatekeeping are typically talked about in opposition to majoritarian theories. Such theories posit that restrictions to lawmaking occur not because proposals are kept off the agenda, but rather because any successful policy must circumvent a series of veto players (e.g., the president (e.g., McCarty and Poole (1995) and Cameron (2000) or the president and the filibuster pivot Krehbiel (1998)). For exposition, suppose the chamber median $m$ has an ideal policy outcome denoted by $x_m$ and that there are two veto players denoted $l$ and $r$ with most preferred policies $x_l \in X$ and $x_r \in X$ such that $x_l < x_m < x_r$.

With the same assumptions as above, typical play in a veto game is as follows:\textsuperscript{6}

- Nature draws a status quo policy $q \in X$ that is perfectly observed by all.
- Chamber median $m$ decides whether to make a proposal $p \in X$. If so, she reports $p$, if not, the game ends and $q$ is realized.
- Veto player $l$ decides whether to veto $p$. If so the game ends and $q$ is realized. If not, play proceeds.
- Veto player $r$ decides whether to veto $p$. If so the game ends and $q$ is realized. If not $p$ is realized.

The equilibrium prediction from such a game is a unique mapping from the set of status quo realization $q$ to a set of policy outcomes. Figure 2 presents the equilibrium outcomes and cutpoints.

\textsuperscript{4}Figure 1 and 2 summarize the results presented in Krehbiel, Meierowitz and Woon (2004).
\textsuperscript{5}Assuming both $p^*$ and $q$ are greater than (or less than) 0.
\textsuperscript{6}Krehbiel’s (1998) pivotal politics version of this game adds the possibility of overriding the decision of one of the veto-players with a super-majority vote.
3 A Statistical Voting Model

Testing the theoretical predictions of the prior section typically uses a statistical model to measure member preferences. Such statistical models presume that members vote for the policy-outcome or position closest to their most-preferred policy (position) with error. To explicate the link between the theoretical and statistical models it is useful to consider the statistical model in some depth. For exposition, I use the model of Clinton, Jackman and Rivers (2004). Although the details slightly differ from other statistical models of roll call behavior (see, for example, Heckman and Snyder (1997), Poole and Rosenthal (1997) and Poole (2000)), the conclusions of the paper are not sensitive to the particulars of the statistical model employed.

Associated with each roll call vote \( t \) \((t = 1 \ldots T)\) is a pair of locations in the policy space – one associated with a successful roll call vote \( (\theta_y(t)) \), and one associated with an unsuccessful roll call vote \( (\theta_n(t)) \). Given that the theories suppose a unidimensional policy space, I restrict attention to a unidimensional statistical model. Assume that the utility for representative \( i \) \((i = 1 \ldots L)\) with ideal point \( x_i \) is given by:

\[
U_i(\theta_y(t)) = -||x_i - \theta_y(t)||^2 + \eta_{it}
\]

\[
U_i(\theta_n(t)) = -||x_i - \theta_n(t)||^2 + \nu_{it}
\]

(1)

where \( \eta_{it} \) and \( \nu_{it} \) represent the stochastic portion of the utility function and \( || \bullet || \) represents the Euclidean norm.

The latent utility differential for representative \( i \) in roll call \( t \) is therefore:

\[
y_{it}^* = U_i(\theta_y(t)) - U_i(\theta_n(t))
= -||x_i - \theta_y(t)||^2 + \eta_{it} - (-||x_i - \theta_n(t)||^2 + \nu_{it})
= -||x_i - \theta_y(t)||^2 + ||x_i - \theta_n(t)||^2 + \epsilon_{it}
= 2x(\theta_y(t) - \theta_n(t)) - \theta_y^2(t) + \theta_n^2(t) + \epsilon_{it}
\]

(2)

where \( \epsilon_{it} = (\eta_{it} - \nu_{it}) \).

Assuming \( \epsilon_{it} \) is iid \( N(0, \sigma_t^2) \), and letting \( \Phi(\cdot) \) denote the standard normal CDF yields:

\[
Pr_i(y_{it}^* = 1|x_i, \theta_y(t), \theta_y(t)) = Pr(y_{it}^* > 0)
= \Phi \left( \frac{1}{\sigma_t} \left( 2x(\theta_y(t) - \theta_n(t)) - \theta_y^2(t) + \theta_n^2(t) \right) \right).
\]

(3)
Conditional on \( x \) and the \( T \)-length vectors of proposal locations \( \theta_{y(t)} \) and \( \theta_{n(t)} \), assuming that representatives vote independently with respect to both indexes yields the probability of observing the \( L \times T \) matrix of roll call votes \( Y \):

\[
L(x, \theta) = \prod_{i=1}^{L} \prod_{t=1}^{T} \Phi \left( \sigma_t^{-1} \left( 2x(\theta_{y(t)} - \theta_{n(t)}) - \theta_{y(t)}^2 + \theta_{n(t)}^2 \right) \right)^{Y_{it}} \times \left( 1 - \Phi \left( \sigma_t^{-1} \left( 2x(\theta_{y(t)} - \theta_{n(t)}) - \theta_{y(t)}^2 + \theta_{n(t)}^2 \right) \right)^{1-Y_{it}} \right)
\]

where the only observable portion of the likelihood function is \( Y \). To draw a similarity to item response models and facilitate estimation, it is possible to derive an expression for the item discrimination parameter as \( \beta_t = 2\sigma_t^{-1}(\theta_{y(t)} - \theta_{n(t)}) \) and the item difficulty parameter as \( \alpha_t = \sigma_t^{-1}(\theta_{n(t)}^2 - \theta_{y(t)}^2) \).\(^7\) This yields the likelihood of:

\[
L(x, \theta) = \prod_{i=1}^{L} \prod_{t=1}^{T} \Phi(\beta_t x - \alpha_t)^{Y_{it}} \times (1 - \Phi(\beta_t x - \alpha_t)^{1-Y_{it}}).
\]

A choice of priors and a set of identifying restrictions completes the specification in the Bayesian context (see Clinton, Jackman and Rivers (2004) for details).\(^8\)

The cutpoint associated with each vote is the ideal point at which the probability of voting “yea” is the same as voting “nay” (i.e., the position of an indifferent member on vote \( t \)). In terms of the model, this implies \( \Phi(\beta_t x - \alpha_t) = .5 \), or \( \beta_t x - \alpha_t = 0 \). Solving for \( x \) yields: \( \alpha_t / \beta_t \).

Having outlined the basic statistical model used to analyze roll call behavior, I now demonstrate how the predictions of formal theories of lawmaking regarding the impact of gatekeepers and veto-players can be integrated, and therefore tested, within this statistical model.

### 3.1 Theoretical Implications for Roll Call Estimates

As roll calls are the mechanism by which policy change occurs, the theoretical predictions regarding equilibrium outcomes – and therefore proposals – have direct consequences for observed roll call

\(^7\)An alternative factorization yields the expression employed by Bafumi et al. (2005).

\(^8\)A Bayesian model enables the computation of any statistic of the posterior and yields an immediate assessment of the precision of each estimate. This offers two benefits. First, ideal point standard errors are estimated directly and it is straightforward to determine whether ideal point differences are non-zero (see Lewis and Poole (2004) for a means of generating bootstrapped standard error estimates for NOMINATE ideal point estimates). Second, an estimate and standard error for any quantity of theoretical interest such as the chamber median and the majority party median can be computed while accounting for any uncertainty about the identity of each. In contrast, ideal point estimators such as NOMINATE do not estimate the required order statistics (i.e., chamber median and median of majority party members) and scholars are consequently forced to assume that the identify of the chamber and party median can be perfectly determined (i.e., the median member’s ideal point is the chamber median).
behavior and ideal points estimated from that roll call behavior. Table 1 recounts the equilibrium policy prediction for each realization of the status quo $q$ and the implied cutpoint for each theory.\(^9\)

<table>
<thead>
<tr>
<th>Status Quo Location ($q$)</th>
<th>Proposal Attempted</th>
<th>Eqm. Outcome</th>
<th>Eqm. Cutpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q &lt; x_m$</td>
<td>yes</td>
<td>$x_m$</td>
<td>$q + \left(\frac{x_m - q}{2}\right)$</td>
</tr>
<tr>
<td>$q \in [x_m, 2x_g - x_m]$</td>
<td>no</td>
<td>$q$</td>
<td>None</td>
</tr>
<tr>
<td>$q &gt; 2x_g - x_m$</td>
<td>yes</td>
<td>$x_m$</td>
<td>$q - \left(\frac{q - x_m}{2}\right)$</td>
</tr>
<tr>
<td>$q &lt; 2x_l - x_m$</td>
<td>yes</td>
<td>$x_m$</td>
<td>$q + \left(\frac{2x_l - q}{2}\right) = x_l$</td>
</tr>
<tr>
<td>$q \in [2x_l - x_m, x_l]$</td>
<td>yes</td>
<td>$2x_l - q$</td>
<td>None</td>
</tr>
<tr>
<td>$q \in (x_l, x_r)$</td>
<td>no</td>
<td>$q$</td>
<td>None</td>
</tr>
<tr>
<td>$q \in [x_r, 2x_r - x_m]$</td>
<td>yes</td>
<td>$2x_r - q$</td>
<td>$q + \left(\frac{q - (2x_r - q)}{2}\right) = x_r$</td>
</tr>
<tr>
<td>$q &gt; 2x_r - q$</td>
<td>yes</td>
<td>$x_m$</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 1: Theoretical Predictions: Gatekeeping (top) and Veto (bottom) Models.

Cutpoints describe the ability to distinguish between member preferences using roll calls with spatial voting. Because cutpoints denote the location of a member indifferent between the status quo $q$ and the equilibrium proposal $p^*$, if two members’ preferences are separated by a cutpoint then the members should vote against one another on that vote. As Table 1 makes clear, predictions about equilibrium proposals also predict which cutpoints should be observed and therefore which members should vote together. In other words, because ideal points are a function of member induced preferences and the observed agenda, the predictions about equilibrium outcomes from gatekeeping and veto theories have implications about the pattern of estimated ideal points.

Gatekeeping theories predict we should never observe a cutpoint in the interior of $[x_m, x_g]$ resulting in policy change.\(^{10}\) Consequently, $g$ and $m$ should never be observed voting differently on votes implementing policy change because $g$ would never allow an issue that divides them on the agenda. Because no cutpoint that divides the two are ever observed in equilibrium, this is equivalent

---

\(^9\)Strictly speaking, the predictions assume that members are (only) policy oriented and that proposals are costly. This rules out the possibility of non-successful proposals (as might happen if members are motivated by position-taking considerations).

\(^{10}\)The theories also clearly generate predictions about the location of successful final passage cutpoints. However, testing whether the cutpoints lie in theoretically impossible locations of the policy space is an inferior test because testing whether the distribution of recovered cutpoints are theoretically consistent requires the ability to estimate member preferences in the space. Krehbiel, Meirowitz and Woon (2005) use ideal points based on the set of all votes for this purpose, but nothing guarantees that the policy preferences relevant for lawmaking also structure behavior on procedural and amendment votes (see, for example, work by Smith and Roberts (2003) suggesting otherwise). As it is unclear whether the ideal point estimates derived using the set of all votes represent policy outcome preferences, interpreting the relationship between the estimated cutpoints and the estimated ideal points becomes difficult. The test I propose avoids such difficulties.
to the prediction that the ideal points of the gatekeeper and chamber median are indistinguishable. The width of the gatekeeper interval calculated using ideal points should be zero. A nonzero interval indicates the presence of roll calls in which the party and chamber medians vote differently from one another exist – an impossibility if the theory is true.

These predictions are true only so long as we restrict the set of roll calls being analyzed to final passage votes. On non-final passage votes – and even on “insignificant” final passage votes (see, for example, Cox and McCubbins (2005, Chapter 6)) – this relationship need not hold because the theory is silent on such votes. For example, neither theory precludes unsuccessful amendments from being offered and voted upon – amendments whose cutting line might lie between the pivots. Consequently, if the entire set of roll calls is analyzed, neither theory necessarily predicts a non-zero interval. Because some final passage votes may be “insignificant,” considering the set of successful final passage votes arguably provides the closest connection between theory and data.

Denoting the ideal point on the relevant observed roll call agenda as \( \hat{x} \) yields the prediction:

**Gatekeeping Theory:** \( \hat{x}_g = \hat{x}_m \).

Testing gatekeeping theories therefore requires identifying the appropriate set of votes and determining whether the ideal points of the gatekeeper and chamber median estimated using those votes are distinguishable. To illustrate the test for gatekeeping models, I use the predictions of the party cartel model (Cox and McCubbins 1993; 2005) which supposes gatekeeping by the majority party caucus (an analogous test could be used to test committee gatekeeping). The relevant test involves estimating ideal points using all successful final passage roll calls and determining whether the ideal points of the chamber median and majority party median are statistically indistinguishable.

There are arguably two ways of conducting the test, and, for robustness, I do both. First, one could estimate and rank order member ideal points, determining whether the median member’s ideal point differs from the ideal point of the party median member. This is how existing calculations of the width of the gridlock interval proceed. A complication is that this calculation assumes that the identities of the medians are known with certainty. A second procedure, and one that escapes this

---

11 Because the theories are silent with respect to behavior on non-final passage votes (or final passage votes that are not vital to a party’s programme), the predicted size of the interval using additional votes is non-obvious. For example, if legislator activity on amending activity is governed by position-taking it is not obvious how such behavior relates to outcome motivated behavior, nor how position-taking preferences relate to policy outcome preferences. Using such votes therefore requires imposing additional assumptions regarding legislator behavior that are not implied by the theories and the possibility that the results depend on these “extra-theoretical” assumptions.
assumption, requires estimating the median ideal point (as opposed to the median member’s ideal point) and the ideal point whose rank ordering corresponds to the location of the party median. Estimating these two order statistics provides the benefit of accounting for uncertainty regarding the identity of the party and chamber median when estimating the statistics of interest.

An analogous argument yields the prediction for veto theories.

**Veto Theory:** \( \hat{x}_l = \hat{x}_m = \hat{x}_r \).

Testing the predictions of a veto model, of which Krehbiel’s (1998) pivotal politics is perhaps the most widely analyzed, requires estimating ideal points on the set of successful final passage votes and determining whether it is possible to distinguish between the ideal points of the set of members whose policy preferences are in the interior of the gridlock interval.\(^\text{12}\) Because status quos in the gridlock interval cannot be changed, veto players should never vote against one another on successful enactments because such measures will never be proposed. Consequently, the ideal points of pivotal members should be indistinguishable.

A complication results from the apparent need to compute gridlock intervals using members from different chambers and branches (see Bailey (2005)). Calculating the gridlock interval appears to require directly comparing Senate, House and Presidential preferences to one another which entails assuming either that members voting in both chambers have identical induced preferences (as Poole’s (1998) Common Space scores assume), or else that legislation considered in both chambers have identical status quos. It is possible to circumvent these complications by examining whether it is possible to distinguish between members assuredly in the interior of the gridlock interval. If so, this casts doubt on the theory because it indicates that the pivotal members do not vote identically.

To test whether the veto players’ ideal points are indistinguishable I determine the largest set of legislators centered around the chamber median whose ideal points are indistinguishable and assess the plausibility that this set includes the veto players. Given that many of the veto players of interest in the pivotal politics model involve particular quantiles of the chamber (e.g., the 41st (60th) most liberal Senator needed to invoke cloture and stop a conservative (liberal) filibuster, or the 1/3rd (2/3rd) most liberal House member needed to override a Republican (Democratic) President), this determination is informative with respect to theoretical predictions. Consequently,\(^\text{12}\) The theory makes slightly different predictions for budget-related legislation since these have statutory debate limits that preclude filibusters.
I test the pivotal politics model by determining the percentage of the chamber whose ideal points are indistinguishable from the chamber median.

Having demonstrated that the theories of most interest to contemporary scholars have testable implications in terms of the distribution of ideal points estimated on the set of relevant roll calls, I now consider the inadequacy of existing tests that ignore the endogeneity of the roll call record and I implement the proposed tests for the 90th - 106th U.S. House.

4 Empirical Tests of Theoretical Predictions

To test the theoretical predictions of section 3, scholars frequently use estimates from statistical models of roll call behavior. A prevalent application is to use ideal point estimates $\hat{x}$ to calculate the partisan (i.e., $|\hat{x}_m - \hat{x}_g|$) and non-partisan (i.e., $|\hat{x}_l - \hat{x}_r|$) gridlock intervals for use as independent variables in a subsequent analysis. Measures of legislative accomplishment such as number of enacted important statutes or descriptions of the legislative environment such as roll rates are then regressed on these, or related, measures to determine which quantity best explains the legislator’s observed behavior. Despite widespread use, at least two severe problems result.

First, to the extent that ideal-point based measures are used in a regression analysis (e.g., predicting the incidence of legislation as a function of the width of a theoretically implied gridlock interval), the fact that the theory being tested has implications for both the dependent variable and the measures of the gridlock interval means that such an analysis is plagued by endogeneity concerns. Because the tested theory has direct implications for the types of roll calls that should be observed, roll call based measures are not exogenous.

Ideal points are noisy measures of true policy preferences and the error associated with that noise is correlated with the dependent variable (e.g., number of significant enactments, roll rates) because of the endogeneity of roll calls. Despite claims that sometime appear in the literature to the contrary, Achen (1983) shows that the consequences are severe – resulting in inconsistent coefficient estimates of unknown magnitude and direction. It is not the case that using endogeneous ideal point estimates to test lawmaking theories results in attenuated coefficient estimates for the ideal point measure. Although the inconsistency of unknown direction and magnitude of regression

---


14 Although the proof is omitted, it is available from the author. The results follows directly from any textbook.
estimates resulting from endogenous regressors is sufficiently severe, the problem is actually worse.

A second, deeper point is that the entire endeavor of calculating gatekeeping intervals for use in a regression analysis is misguided. As section 3 makes clear, because the theory being tested yields predictions about the distribution of recovered ideal points directly, the existence of a non-zero interval indicates: 1) the theory is false, or 2) the theory is true but the votes used to generate ideal points are irrelevant to the theory. It is therefore unclear what introducing the width of a gridlock interval in regression analyses (for example) accomplishes – the width of the gridlock interval should be zero if the correct votes and estimator are used and the theory is true. Finding a non-zero interval using the set of final passage votes is either: 1) the only evidence required to test a theory or 2) evidence that the votes and method are insufficient to test the theory.

4.1 Integrating Theory and Method: The 90th-106th U.S. Houses

To illustrate the integration of estimation and theory-testing I examine behavior in the 90th - 106th Houses (1967 - 2000). The point of the analysis is not to resolve the ongoing “pivots-vs-party” debate, but rather to suggest that at least some of the debate is being waged on the wrong terrain because of problematic measures.

The first task requires identifying the set of non-unanimous roll calls that are most theoretically relevant – the set of successful final passage votes. Successful final passage votes is theoretically appropriate so long as members believe that the legislation was likely to become law (or relevant for the majority party “brand”). I use the determination of Rohde (2004) to identify the set of final passage votes, and I determine whether the vote passes (accounting for super-majoritarian
requirements such as votes considered under Suspension of the Rules).  

Using only a subset of the roll call record is unproblematic for three reasons. First, the subset of examined votes are those that are most obviously related to the theories of interest. Although theoretically relevant information may indeed be contained in other votes, using these additional votes (e.g., votes on procedural issues or amendments) requires reaching beyond the theoretical models and imposing additional assumptions on member behavior. Imposing these assumptions is undesirable because it introduces another possible error in the test — the possibility that the results are a consequence of mistaken assumptions regarding member behavior on votes whose relationship to the theories is uncertain. Second, even if plausible assumptions exist, increasing the number of analyzed roll calls only increases the precision of the estimated ideal points and only strengthens the substantive conclusion of the following section. It is not the case that the reported results are a consequence of small sample size; if anything, increasing the sample will only strengthen the results. Using fewer votes makes the estimates more imprecise and increases the likelihood that the ideal points cannot be distinguished (i.e., we cannot reject the theory).

After identifying the theoretically relevant roll calls, testing the predictions involves estimating two statistical ideal point models — one that is unconstrained and one that imposes the theoretically implied constraints — and testing whether the constraints are valid (Clinton and Meirowitz 2003). This appears difficult for the predictions of section 3 because the identities of the legislators whose ideal points should be constrained are unknown. An alternative, but statistically equivalent approach involves estimating an unconstrained model and testing whether the relevant recovered ideal points are statistically distinguishable.

Given the simplicity of the test, the analysis is straightforward. For each set of successful final passage votes, I estimate members’ unidimensional ideal points using the statistical model of section 2 while imposing the identification restriction that the ideal point estimates have mean zero and unit variance (Rivers 2003). Results using NOMINATE estimates are substantively identical.

\[17\text{Using Rohde’s (2004) classification of House roll calls for the 83rd through 106th Congresses, I use roll calls classified as (the number of such roll calls is listed in parentheses): Final Passage/Adoption of a Bill, Final Passage/Adoption of a Conference Report, Final Passage/Adoption of a Resolution, Passage/Adoption of a Bill under Suspension of the Rules, or Passage/Adoption of a Resolution under Suspension of the Rules.}

\[18\text{In fact, the inability to determine the identity of the relevant members (i.e., } g, m, l \text{ and } r) \text{ is why ideal points are often used to test the theory. Although it might appear possible to use a computationally simpler method such as counting the number of times that the relevant actors vote differently, without external information as to which members we should compare this calculation is uninformative. Not only do ideal points account for differences in voting behavior by construction, but, in so doing they also provide an ability to identify the identity of the theoretically critical members using information in the observed voting patterns.}

\[19\text{In each case, I generate 1,000,000 samples thinning by 2,000 using IDEAL 0.5 (Jackman 2004). No evidence of}
For purposes of illustration, Figure 3 presents the results graphically for the 106th House based on the 219 successful final passage votes.

[Insert Figure 3 About Here]

The results are immediately evident and discouraging for both theories. In terms of the party gatekeeping model, there is no doubt that the estimated chamber median (m) and majority party median (g) are disjoint. Because the illuminated 95% regions of highest posterior density do not overlap we can conclude that 219 observed successful final passage votes contain enough information to distinguish between the chamber and majority party median. The estimated difference between the chamber and Republican party median is .56 with a 95% confidence interval of [.49,.62]. Although unsurprising given this difference, examining the distribution of estimated cutpoints (plotted along the x-axis) reveals that 12% have posterior medians that lie in the interval where none should occur if the theory is true.20

The results for the pivot theory are no more encouraging. Figure 3 immediately suggests the result, as the distribution of ideal points is far from contained in the 95% region of highest posterior density for the chamber median. In fact, only 8% of the chamber vote identically with the chamber median (m) using a 95% threshold – a quantity far less than the predicted supermajority.21

Having illustrated the analysis using the 106th House, I now examine the robustness of the substantive conclusion across time (and differently sized roll call matrices) by replicating the analysis for the 90th through the 106th Houses. Table 2 reports the results.

For each House, Table 2 reports: the number of successful final passage votes used to estimate member ideal point estimates, the distance (and 95% Highest Posterior Density estimates for that distance) between the chamber median and the majority party median, the number of votes whose estimated cutpoints lie in this interval with greater than 95% posterior probability, and the percentage of the chamber that is indistinguishable from the chamber median.22

---

2022% of the cutpoints have 95% confidence intervals with support in the interval.

21One response is that the low percentage is the possibility that it is due to incorrectly assuming a unidimensional space in the statistical model. I address this possibility in the next section.

22For robustness, the distance was also calculated between the chamber median’s ideal point and the party median’s ideal point. Although this more closely replicates existing uses, it fails to account for the fact that these identities are not known with certainty. The posterior differences (and standard error) for the 90th through 106th Houses are: .71 (.23), .44 (.26), .48 (.19), .63 (.18), .54 (.14), .47 (.14), .49 (.16), .50 (.23), .47 (.22), .60 (.26), .60 (.23), .55 (.22), .49 (.20), .47 (.18), .47 (.17), .61 (.19), .53 (.12). The standard errors for these differences are larger than those reported in Table 2 because the posterior median is more precisely estimated than the median of the posterior means.
Congress | Roll Calls Analyzed | Distance [95% HPD] | Pct. Invalid Cutpoints | % Identical Members
--- | --- | --- | --- | ---
90 (1967-1969) | 188 | .72 [.64,.80] | 36.8 % | 16.5 %
91 (1969-1971) | 174 | .47 [.39,.55] | 27.7 % | 29.2 %
92 (1971-1973) | 226 | .48 [.42,.54] | 17.1 % | 22.1 %
93 (1973-1975) | 364 | .61 [.55,.68] | 15.1 % | 18.6 %
94 (1975-1977) | 423 | .52 [.47,.57] | 12.8 % | 12.1 %
95 (1977-1979) | 413 | .46 [.41,.51] | 12.7 % | 15.6 %
96 (1979-1981) | 288 | .47 [.38,.54] | 30.8 % | 19.6 %
97 (1981-1983) | 163 | .51 [.44,.57] | 26.5 % | 24.5 %
98 (1983-1985) | 205 | .47 [.40,.53] | 30.6 % | 17.7 %
99 (1985-1987) | 135 | .56 [.48,.62] | 36.8 % | 19.0 %
100 (1987-1989) | 201 | .55 [.49,.62] | 37.9 % | 22.3 %
101 (1989-1991) | 172 | .56 [.49,.62] | 37.6 % | 19.0 %
102 (1991-1993) | 182 | .45 [.39,.52] | 36.9 % | 16.9 %
103 (1993-1995) | 188 | .45 [.40,.51] | 18.1 % | 16.2 %
104 (1995-1997) | 197 | .45 [.40,.50] | 34.1 % | 14.4 %
105 (1997-1999) | 192 | .61 [.55,.67] | 37.6 % | 9.3 %
106 (1999-2001) | 219 | .56 [.49,.62] | 22.0 % | 8.0 %

Table 2: All Successful Final Passage Votes

The results are immediately and consistently evident across every House I examine. It is never the case that the chamber and party median are indistinguishable. Although the scales are not comparable across Houses, the non-zero differences are sizable – ranging from .45 (103rd and 104th) to .72 (90th). As is to be expected given these differences, the percentage of successful final passage votes with cutpoints estimated to lie in the party gridlock interval ranges from 12.8 % to 37.9 %. The critical result is not the slight variation in differences evident across Houses, but rather the non-zero differences clearly evident in every case.

The predictions of the pivot theory fare no better. Only between 8 % and 29.2 % of the House are indistinguishable from the chamber median – far less than the predicted supermajority. So long as the preferences of the President and Senate veto pivots are distinguishable from the House median (a likely possibility), these findings are problematic for the pivot theory.23

To examine the robustness of the results I replicate the analysis in two ways. First, because there is ambiguity over what constitutes a final passage vote, I replicate the analysis for selective Houses using an alternative coding scheme. Second, to account for the possibility that the evident differences are due to incorrectly assuming that votes are structured by a single underlying dimension in the statistical voting model, I conduct the tests using only successful final passage votes

---

23Replicating the analysis for selected Senates (unreported) suggests that this possibility is indeed unlikely.
involving domestic, non-appropriation votes.

4.2 Checking Robustness: Measuring “Final Passage”

To ensure that the substantive conclusions are not a consequence of incorrectly identifying final passage votes, I replicate the analysis for selective Houses using the final passage measures of Krehbiel, Meirowitz and Woon (2005). The determinations differ, in part, because Krehbiel, Meirowitz and Woon (2005) count votes on House Joint Resolutions (typically special rules) as final passage votes. Although I remain agnostic as to whether such votes are theoretically relevant, these differences represent a useful robustness check from a theory-testing perspective.

The Houses selected for examination include those with the largest (90th) and smallest (103rd) differences and the smallest number of roll calls (99th) based on the results reported in Table 2. Table 3 presents the results.

<table>
<thead>
<tr>
<th>House</th>
<th>KMW Roll Calls Analyzed</th>
<th>Distance [95% HPD]</th>
<th>Pct. Invalid Cutpoints</th>
<th>% Identical Members</th>
</tr>
</thead>
<tbody>
<tr>
<td>90 (1967-9)</td>
<td>136</td>
<td>.76 [.67,.84]</td>
<td>37.6 %</td>
<td>18.8 %</td>
</tr>
<tr>
<td>99 (1985-7)</td>
<td>160</td>
<td>.54 [.48,.60]</td>
<td>38.6 %</td>
<td>15.3 %</td>
</tr>
<tr>
<td>103 (1987-9)</td>
<td>99</td>
<td>.39 [.34,.45]</td>
<td>9.1 %</td>
<td>18.7 %</td>
</tr>
</tbody>
</table>

Table 3: Replication Using Alternative Final Passage Measure

Table 3 reveals identical results. Because the substantive conclusions are robust to differing interpretations of how observed roll calls relate to the production of policy outcomes (i.e., what constitutes final passage), our confidence in the conclusion is enhanced.

4.3 Checking Robustness: Multiple Dimensions

A second, and arguably more consequential issue relates to the assumption that all of the roll calls are determined by the same underlying preference dimension. In other words, a potential objection is that the detected differences between the chamber and majority party medians’ ideal points (and between the chamber median’s and members’ ideal points) constitutes evidence that the empirical model is insufficient rather than the inadequacy of the theoretical predictions. Because the statistical model allows for errors, the observed difference between the chamber median and the gatekeeper may be the result of a higher dimensional space being projected into a lower dimension.
representation with error.\textsuperscript{24} If so, the problem lies not with the theories, but rather with the estimation procedure. If actual voting behavior is structured by multiple dimensions, incorrectly estimating a single dimension may create the observed differences.

Although it has been argued that contemporary voting is largely structured by a single dimension (see, for example, Poole and Rosenthal 1987), the determination of “how many dimensions” is arguably a substantive rather than statistical question (see Jackman (2001) and Poole (2005)). To examine whether the results are likely due to mistakenly estimating the dimensionality of the policy space, I replicate the analysis using only successful final passage votes on non-appropriations bills that primarily address domestic issues.\textsuperscript{25} Restricting the analysis to this subset effectively imposes the weaker assumption that member policy preferences on domestic, non-appropriation legislation (as opposed to all legislation) are structured by a single underlying dimension.

Although further subsets of the policy space could be investigated using more nuanced substantive classifications of the roll calls, the more restrictive the definition of the issue becomes, the fewer votes there are to be analyzed. With fewer votes, the resulting estimates become increasingly imprecise and it becomes more difficult to reject the theoretical predictions. A trade-off between substantive specificity and the ability to recover estimates of sufficiently useful precision exists.

Table 4 presents the results.

As Table 4 makes clear, none of the substantive conclusions are affected when only successful final passage votes on domestic, non-appropriations legislation are used. Using fewer votes, but votes for which we have stronger reasons to believe that a common underlying preference might exist, reveals no difference. Consequently, the results are not an artifice of differing member policy preferences on domestic and foreign issues being projected onto a common dimension.

The reported results are not intended to definitively resolve whether lawmaking is shaped by party institutions – such a resolution would require moving beyond roll calls, replicating the above analysis across longer periods of congressional history, reconciling discrepancies in what constitutes “final passage,” and potentially restricting the analysis to further subsets of the roll call record (e.g., classifying votes according to the issues they raise or by the “importance” of the legislation being considered) – but the results suggest that neither performs particularly well. However,\textsuperscript{24} A statistical model that allows for probabilistic voting recovers a much smaller dimensional space than one which assumes perfect spatial voting. The latter requires an additional dimension for every unique voting profile.\textsuperscript{25} I use the determinations of Rohde (2005) – specifically, whether the ISSUE addresses: Economy, Taxes and Budget Issues (500’s), Energy and Environment (600’s), Government Operations, Civil Rights, and Justice (700’s), Welfare and Human Services (800’s), or Miscellaneous Domestic (900’s).
they do definitively resolve whether the conflicting substantive conclusions of prior roll call based studies were potentially affected by the endogeneity of the votes. Eliminating the possible problems introduced by endogenous measures reveals weak support for either theory. For every subset of votes I consider in the House using either measure of final passage, the chamber median and party median can be distinguished. Furthermore, the percentage of the House whose voting profile (and therefore ideal point) is indistinguishable from the chamber median is consistently small.

5 Problems of Theory or Problems of Estimation and Data?

An objection to the test I propose to address the endogeneity of roll calls might be that the estimator is incorrect even though substantive information is used to restrict the set of analyzed roll calls to domestic issues that do not involve appropriations. It may still be the case that the estimated difference is due to incorrectly projecting a multi-dimensional preference space (with each vote being decided by a single dimension) into a single dimension. Although it is impossible to conclusively dismiss this possibility, reasonable reactions exist.

If the observed differences in the tests I propose are due to incorrectly estimating the relevant policy space, then existing uses of ideal points to understand lawmaking must be similarly invalid for
the same reason. The fact that we do not know the dimensionality of policy preferences underlying
the recovered roll call record may make the implementation of the test I propose difficult, but it
does not invalidate the legitimacy of the tests. The tests still apply to the properly restricted
set of votes (e.g., “important” legislation or votes within certain issue areas). The difficulty of
implementing the test does not diminish its superiority over existing – and similarly flawed – uses
of ideal points.

If identifying the relevant set of votes is difficult, this provides reason for caution in using ideal
points to test theories of lawmaking; statistical models must be properly restricted to the relevant
set of votes. Although this paper attempts to do so by examining the robustness of the results to
alternative definitions of final passage and restricting the set of votes to votes pertaining to similar
substance, the roll call record can obviously be parsed in many alternative ways.

In light of this possibility, one “solution” would be to abandon the use of ideal points and
test theories using implications regarding the roll call record directly. For example, because the
party cartel model predicts that the chamber median and majority party median should never
vote differently on final passage votes, we can examine whether a majority of the majority party
votes together. Testing the pivot theory similarly reduces to examining whether have enacting
coalitions on successful final passage votes are larger than 66% of the participating chamber. Such
investigations do not assume anything about the dimensionality of member policy preferences –
in fact, they assume that voting is perfect and that every unique voting profile (i.e., supporting
coalition) indicates an additional dimension in the policy space.

Moving away from roll call based estimates avoids the problems introduced by using estimates
based on the wrong dimension, but the cost of doing so is high. The tests become so crude and
incorporate so little of the theoretical explanation so as to be almost uninformative with respect to
the suggested causal mechanism.\textsuperscript{26} Is it really surprising that a majority of the majority party votes
together on successful final passage votes? Although consistent with the predictions resulting from
party agenda control, this fact says nothing about the underlying mechanism (e.g., which members
vote together). Likewise, is the fact that 72% of votes in the 103rd House have super-majority
enacting coalitions strong or weak evidence for a self-described simple theory based on veto pivots?

\textsuperscript{26}Absent an ability to identify the identity of the members in each coalition, it may also become difficult to
distinguish between the theories. Most notably, although the pivot theory precludes an enacting coalition smaller
than a super-majority, observing a super-majority coalition does not invalidate the predictions of the party theory
– evidence supportive of the pivot theory is also supportive of the party gatekeeping theory. The only way to reject
the party theory is to observe an enacting majority that does not contain a majority of the majority party.
Furthermore, any performance will pale in comparison to the analogous predictions of a median voter model because a majority votes will always vote together on successful final passage votes.\footnote{Of course, the median voter model is also impossible to reject using ideal point and cutpoint estimates.}

These reactions are not intended to dismiss the importance of alternative investigations and results, but rather to suggest that testing theories using measures only weakly related to postulated causal mechanisms results in weak conclusions and subsequent debates about the interpretation of the results (see, for example, Krehbiel 2000). In contrast, tests that utilize theoretically implied concepts are more powerful for evaluating claims concerning the causal mechanism. In the event that our ability to measure theoretically implied covariates introduces error as a consequence of making measurement assumptions that cruder tests avoid, both examinations are useful. If so, the arguments and conclusions of this paper are important and informative.

6 Conclusion

The ability to generate theories of lawmaking has not been matched by the ability to conclusively evaluate theoretical predictions. One difficulty in achieving scientific consensus is the absence of theoretically consistent measures and disagreement about the evidence for and against theories of interest. The lack of consensus is extremely consequential because the accumulation of knowledge and scientific progress depends on the ability to identify measures whose relationship to theories of lawmaking is clear and uncontested (Kramer 1986).\footnote{The importance of this task is reinforced by the National Science Foundation’s funding for multi-year workshops (Empirical Implications of Theoretical Models) focused on precisely this task.}

Proper tests of theoretical predictions require that measures used to test the predictions are generated using models that, at the very least, are not inconsistent with the theory being tested. This point has important implications for theory-testing using roll call data because the theories being tested typically have direct implications for the roll calls that should be observed. Treating ideal points as if they are exogenous to theories of lawmaking is improper and likely misleading.

Because the legislative agenda is endogenous to the party gatekeeping and pivotal politics theories for theoretically relevant votes, a non-zero gridlock interval actually indicates either: 1) the theory is false or 2) the theory is true but the votes used to generate ideal points are irrelevant to the theory. In either case it is unclear what the prevalent practice of using non-zero gridlock measures in regression analyses accomplishes. Accounting for the endogeneity of roll calls in the
90th through 106th U.S. Houses reveals that neither theory performs particularly well.

Assessing the support for theoretical predictions using measures that are questionably related to theoretically implied concepts results in ineffective tests. However, achieving consensus on measures is difficult in political science because theoretically relevant quantities are almost never observed and the data that is observed is often faintly related to theoretical concepts. Similar to Poole’s (2005) caution that “anyone can construct a spatial map...but the maps are worthless unless the user understands both the spatial theory that the computer program embodies and the politics of the legislature that produced the roll calls” (p. 209), the arguments of this paper demonstrate that recognizing and accounting for the connections between the roll calls and theories of lawmaking is essential for the proper use of ideal points to test lawmaking theories.
References


Figure 1: Equilibrium Predictions from Gatekeeping Model.
Figure 2: Equilibrium Predictions from Veto Model.
Figure 3: Theoretically Implied Estimates for the 106th House: The figure graphs the density of Democratic (dashed) and Republican (solid) ideal points, the 95% quantiles for the chamber (m) and majority party (g) medians, the distribution of estimated cutpoints and the percentage located in each partition.