Testing Theories of Lawmaking

Keith Krehbiel[1], Adam Meirowitz[2], and Jonathan Woon[3]

[1] Stanford University, Stanford, CA krehbiel@stanford.edu
[2] Princeton University, Princeton, NJ ameirowi@princeton.edu
[3] Stanford University, Stanford, CA jwoon@stanford.edu

Explaining the policies that result from collective choice in legislatures has been a prominent objective of positive political theory. Two paths of complementary research have been traversed. Social choice theory attempts to identify stable policies by focusing on characteristics of legislative preferences and the choice space. Noncooperative game theory more explicitly relies upon characterizations of institutions to identify regularities in behavior. Jeffrey Banks made fundamental contributions to both important strands of theory.

From the social choice perspective, a few seminal papers have clarified the mathematical structure of policies that are supported by stable coalitions. Plott [28] demonstrated that the existence of stable policies is critically dependent on symmetry conditions. McKelvey and Schofield [27] established that these conditions generically do not hold as long as the policy space is of sufficiently high dimensionality. Banks [6] provided a more precise treatment of the problem, correcting the results from earlier papers, and Saari [32] closed the question by providing tight dimensional bounds.

Institutionalism and noncooperative game theoretic approaches began to flourish in the late 1970s and early 1980s, and Banks, even in the infancy of his career, was an immediate, accelerating force. By recognizing and exploiting the fact that there is a modicum of structure even in environments where social choice theory implies the absence of stable policies, Banks [5] characterized the set of policies that can emerge as undominated equilibrium outcomes from amendment agendas. In characterizing what has become known as the Banks Set, the article "Sophisticated voting outcomes and agenda control" illuminated the interplay between sophisticated behavior and a fixed agenda structure, thereby paving the way for the subsequent focus on characterizing the types of agendas developed endogenously. Thereafter, Austen-Smith [4] considered sophisticated voting and optimal agenda construction and showed that for an empirically plausible type of agenda formation, sincere voting

[0] Forthcoming in David Austen-Smith and John Duggan (eds.) *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*

and sophisticated voting are observationally equivalent. Similarly, Banks and Gasmi [9] shed light on both strategic proposing and voting. Most recently, Banks and Duggan [7] pushed on to the frontier while also circling back to the past. They consider a large class of endogenous agenda problems and find important relationships between solution concepts in social choice and equilibria in bargaining games.

Concomitant with these developments, a related but more applied literature progressed in the area of legislative studies. Scholars intent on understanding the role of parties in Congress developed and occasionally tested models of endogenous agenda formation. Aldrich [1, 2], Aldrich and Rohde [3], Shepsle and Weingast [34], Weingast [38], and Krehbiel and Meirowitz [25], for example, all analyzed models similar to that of Banks and Gasmi but with an explicit status quo. Within this set of paradigmatically similar works, the range of implications and testable hypotheses regarding rights and power within legislatures is diverse. For example, one particularly lively debate concerns whether political parties are first- or second-order forces in the shaping of law. One set of researchers argues that the majority party in Congress exerts disproportionate control over legislative choices (see, for example, [14, 15, 31, 35]). An opposing stance is taken by an assortment of researchers whose theoretical and empirical work suggests that the evidence for party strength in Congress is overstated and that nonpartisan theories provide better first-order accounts than partisan theories of systematic patterns in the data (e.g., [11, 22, 33]).

Only with rare and noteworthy exceptions, such as Banks and Duggan [8], do researchers appreciate the intricate and crosscutting relationships between social choice, noncooperative game theory, institutionalism, nonpartisan and partisan models, and jointly theoretical and empirical research. The connections can be portrayed roughly as follows. In the debate within legislative studies on the role of parties, the two sides are, in effect, competing endogenous agenda models. The models, furthermore, are unidimensional special cases within the class of models that Banks and Duggan consider. From a theoretical perspective, the question of whether legislative policy making is better characterized by strong or weak parties is a question about the type of aggregation rule(s) and the distribution and sequence of proposal rights in the bargaining problem. In the *party cartel* model, Cox and McCubbins contend that the majority party leadership can veto any and all changes from the status quo that the party dislikes. In the *pivotal politics* model, Krehbiel contends that either of two different agents – a veto pivot or a filibuster pivot – has a key, constraining influence on outcomes. Both models implicitly assume that amendment rules are sufficiently open that core convergence is possible. However, because of supermajority requirements and more extreme veto rights, core convergence is not assured.

Adjudicating between the two competing models is methodologically difficult for several related reasons. Chief among these is the challenge of distinguishing preference effects from party effects [20, 26, 37]. Other problems

involve difficulties in measuring preferences [12, 13, 17, 18, 29], difficulties in interpreting results based on conventional measures, such as "party voting" and "party cohesion" [23], and, not least, the near observational equivalence of the theories in question [15, 24].

Eventually, tests of these simple models may also need to confront an even more serious challenge: how meaningful are comparisons between specific models – say partisan versus nonpartisan – if neither such model explains a great deal of legislative behavior? One reason for this may be that a single dimension is insufficient for characterizing the lawmaking environment, whereas both pivot and cartel theories explicitly rely on this assumption.

Of course, views on the dimensionality of the congressional policy space and how it should be portrayed in models are unsettled, too. Social choice theoretic results clearly demonstrate the analytical significance of dimensionality. Practical significance may be another matter, however. For example, recent applied research in legislative studies has tended to rely on unidimensional models, in part because Poole and Rosenthal find that a large quantity of roll call voting in Congress can be reduced statistically to one primary dimension of conflict. This finding is not without qualification or controversy either (see [12, 18, 36] among others), yet most empirically testable models of legislative choice remain unidimensional. Furthermore, such models are tested under the implicit assumption that legislative choice is also unidimensional. The danger of a heavy reliance on unidimensional models is that, even if preferences across dimensions are highly correlated, proposals and voting behavior in a higher dimensional world may be dramatically different than proposals and voting behavior in a one-dimensional world.

This suggests that most direct horse-race tests between two theories that have not been shown to depart from a credible null model run the risk of producing a Pyrrhic victor. Exerting great effort to select the better of theories A and B when neither is demonstrably superior to null hypothesis C is comparable to an optometrist being compulsive about the astigmatism numbers in a prescription for Stevie Wonders' sunglasses.

Mindful of the relationship between the competing specific models of legislative policy making and the much larger subsuming class of endogenous agenda models, we approach the associated obstacles and caveats as follows. We not only extract falsifiable predictions from the cartel and pivot models, as others have done, but also test them against an even simpler unidimensional model of endogenous agenda formation—a chamber operating under an open rule agenda with majority rule. This latter model predicts that the *median* legislator's ideal point will be enacted.

This approach leads to some modest but unique results. First, while the cartel, pivot, and median theories have proven observationally similar, tracing out predictions about the types of cutpoints that should be observed leads to a straightforward, potentially-discriminating test of the theories. The theoretical overview in the next section identifies the situations in which the predictions diverge. Thereafter, use of DW-Nominate cutpoints and ideal points

allow for a seemingly straightforward test between three competing theories: median voter (the baseline model), party cartel, and pivotal politics.

In light of the discussion above and some additional subtleties introduced below, however, the test is not adequate for making confident inferences. Specifically, we note that this test may not place the theories on an equal footing and also that their predictive abilities seem neither very impressive nor very different from one another. As a result, we devise another test in which we compare their predictions to those of a naïve null model. Again, the findings are inconclusive. None of the unidimensional models perform particularly well compared to the naïve random null model based on prediction rates. Nonetheless, visual inspection of the distribution of cutpoints reveals more clearly that gridlock-interval theories have an empirical basis. Deviations from normally distributed cutpoints occur, as hypothesized, in both theories' gridlock intervals. A concluding discussion attempts to add coherence to the admittedly mixed findings and to forecast future work in this area with a sense of optimism befitting of the honoree of this volume.

1 Models

Broadly construed, endogenous agenda models may be considered bargaining games in which agents select a policy from some set of alternatives. An exogenous rule assigns individuals the right to make proposals in particular periods. In each period, before a decision is reached, either a proposal is made or a binary vote is taken. The game form specifies the voting rule that is used for each decision. Banks and Duggan [8] present a class of models in which the reversion policy may be strictly preferred to some of the feasible alternatives and demonstrate that the pivot model is a member of this class. Clearly, the cartel model is as well.

The cartel and pivot models have much more in common analytically than their substantive conclusions suggest. As noted above, both models are unidimensional, and each model has a specified pair of actors whose ideal points define a gridlock interval inside which the status quo policy is predicted to remain unchanged. Therefore, each theory partitions the policy space into different intervals that nevertheless have common qualitative properties. For status quo points sufficiently extreme on either side of the policy spectrum, in equilibrium, the new policy gravitates to the ideal point of the median voter and is, therefore, centrist. However, for status quo points that are centrally located, gridlock occurs.

In contrast, the differences between the theories are fewer in number, more subtle, yet of great substantive significance. In a nutshell, the party-centered cartel theory predicts that, at any given time, public policy can be significantly out-of-step with what the median legislator wants, while in a world of pivotal politics, relatively moderate (hence, representative) policies will prevail.

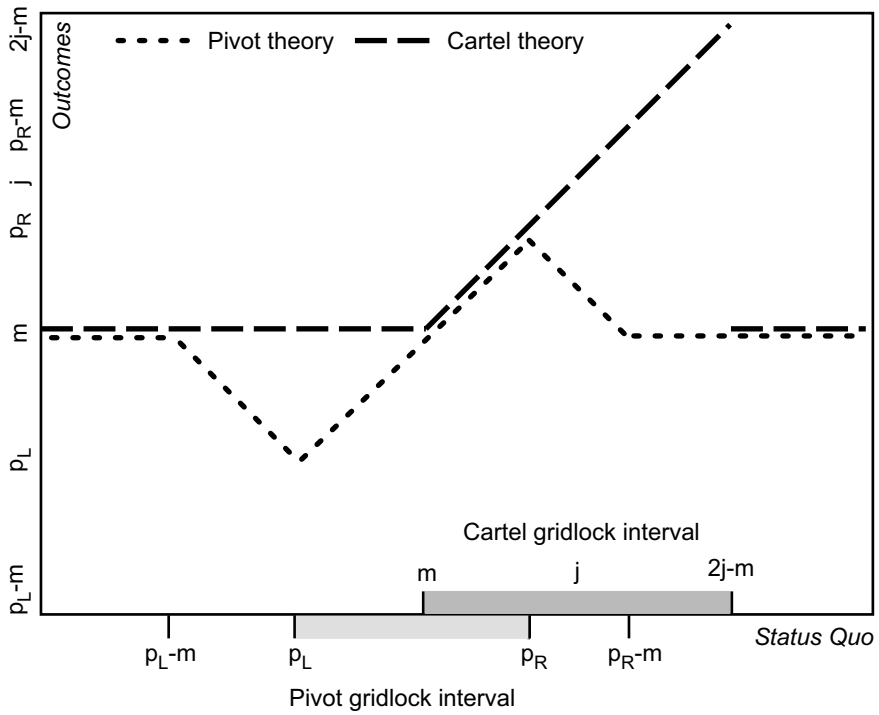

Fig. 1. Equilibrium outcomes in cartel and pivot theories

Figure 1 illustrates these expectations. The horizontal axis represents the policy space and underscores the fundamental dependence of outcomes on the status quo. The top part of the horizontal axis is the characterization of the gridlock interval under cartel theory, while the bottom portion represents the pivot theory's gridlock interval.

Consider cartel theory first. There are two stages of the game, as explicated by Cox and McCubbins [15]. First, an agent of the majority party decides whether to report a proposal to the entire legislature, or whether instead to exercise "negative agenda power" (a.k.a. "gatekeeping"). Second, if the party leader makes a proposal, then the entire legislature chooses a final policy under an open rule. Accordingly, let m represent the median voter and j represent the party median agent. (For purposes of exposition, suppose $m < j$.) Then $2j - m$ is the reflection point of m with respect to j. With this notation, the interval $(m, 2j - m)$ exhaustively defines the status quo policies that a majority of the majority party prefers to the median's ideal point. Therefore, any such status quo elicits majority-party gatekeeping, hence gridlock. On the other hand, for any status quo point outside this cartel gridlock interval, the majority party chooses to report a proposal knowing that it will be amended under an open rule and result in a policy at the chamber median's ideal point.

The vertical axis in Fig. 1 is interpreted as the outcome of the legislative process. We can then represent the cartel theory as a mapping from status quo points, q, into outcomes, x. The graph of the outcomes for the cartel theory begins on the far left side at the median voter's ideal point, m, and is flat while q increases, all the way to the beginning of the cartel gridlock interval. The graph of the outcomes then turns upward 45 degrees and continues to climb throughout the gridlock interval. At the right boundary of the interval $(2j - m)$, the majority party median is indifferent between the status quo and the chamber median. At all subsequent points, however, the party-median prefers the chamber median outcome and, as in the case of the status quo points to the left of the gridlock interval, sends a proposal to the floor, which is then amended until m is passed.

The pivot theory is only slightly more complex [21, 22]. Like the cartel theory, it has a gridlock interval, although it is more centrist than the cartel gridlock interval. Let p_L be a generic left-of-center pivot (e.g., the ideal point of the 41st senator from the left, whose vote is necessary and sufficient for invoking cloture) and let p_R be a generic right of center pivot (e.g., the 67th senator from the left, whose vote is necessary to override a presidential veto). The gridlock interval is then (p_L, p_R). For status quo points within this interval, the equilibrium policy is the status quo policy.

Outside of the pivot theory's gridlock interval, however, two separate types of behavior may occur, depending upon the extremity of the status quo point relative to the median voter's ideal point. For the most extreme status quo points, convergence to the median voter's ideal point is complete, just as it was in the cartel theory. For two intervals adjacent to the gridlock interval, however, this convergence is limited by the equilibrium requirement that the optimal proposal leaves the pivotal voter indifferent between the status quo point and the proposal.

The mapping of outcomes for the pivot theory is also depicted in Fig. 1. It has a trough in one area in which the cartel theory is flat, and a peak in the adjacent area within the cartel theory's gridlock interval.

In addition to the pivot and cartel theories, we consider two baseline or null models. The *median* model is characterized by an open amendment rule and majority rule voting. It is well known that the equilibrium outcome from this process is the median legislator's ideal point [10]. Theorems 5 and 6 of Banks and Duggan [7] establish that all no-delay stationary equilibria in unidimensional bargaining models involve selection of the median voter's ideal point. The other baseline model is *random* and assumes that agendas are generated by an exogenous stochastic process. Specifically, it assumes that each yea and nay location pair is drawn from a bivariate normal distribution.

The median-voter and random-alternatives models are sensible baselines for somewhat different reasons. The defense for the median-voter-theoretic baseline is that most research in legislative studies argues (or assumes) that institutional features are rich and textured and that such features have significant effects on choice behavior. By implication, institutionally sparse models

are inadequate in accounting for variation. It stands to reason that models with institutional features should predict better than relatively institution-free counterparts. The defense for the random null is similar but somewhat more extreme. The median baseline, while sparse, is nevertheless grounded in collective choice theory and has a modicum of credibility as a predictive theory. In contrast, the random null model has no grounding in explicit institutions and behavior. It, therefore, provides a more fundamental opportunity to determine whether the signals extracted from knowingly noisy data are consistent with any model.

2 Cutpoints

While Fig. 1 clarifies the analytic differences between cartel and pivot theories, it is not data-ready, because status quo points are not observable and are difficult to estimate. This obstacle is somewhat less debilitating in the presence of cutpoint estimates – essentially roll-call-specific midpoints between the estimated bill and status quo locations. Consider the case of a vote on final passage of a bill b implicitly paired against the status quo q. If legislators vote sincerely (as theory suggests they will at the final stage) and have symmetric single-peaked preferences, then the cutpoint for such a vote is $c = (b + q)/2$. Everyone on the bill side of this cutpoint votes yea and everyone on the status quo side votes nay. This simple observation provides considerable leverage on the problem, because it is now straightforward to supplement the graphs of outcomes in Fig. 1 with a graph of implied cutpoints. Figure 2 does this.

Two observations reveal the logic of cutpoint analysis. First, for any interval of the status quo, q, define $b^*(q)$ to be the equilibrium proposal. Then, the cutpoint implied by the theory on the vote on final passage is simply $c(b^*(q), q) = (b^*(q) + q)/2$. Thus, the theory implies a mapping from status quos to cutpoints that circumvents the need to identify status quos directly. Second, any theory with a gridlock interval is falsifiable via some cutpoints within its gridlock interval. For example, any cutpoint within $(p_L, p_R$ in the pivot theory implies that either the bill location is out of equilibrium or the status quo is in the gridlock interval in which case we ought not to observe a proposal. Similarly, any cutpoint in (m, j) within cartel theory implies, contrary to the theory, that either the bill does not lie at the chamber median or that majority party leaders ought not to have let the bill be considered.

Within this cutpoint framework, the implications of the two theories jointly partition the policy space as shown on the vertical axis of Fig. 2. In the two extreme intervals, I and V, the existence of cutpoints is consistent with both theories, and, so, too, are the large coalition sizes that form when the status quo is extreme relative to the proposed, moderating bill, $b^*(q)$. Similarly, in the center and moderate interval, III, no discrimination between the theories is possible. Here, however, the reason is that both theories predict no cutpoints. Finally, in intervals II and IV, a cutpoint-based test can discriminate

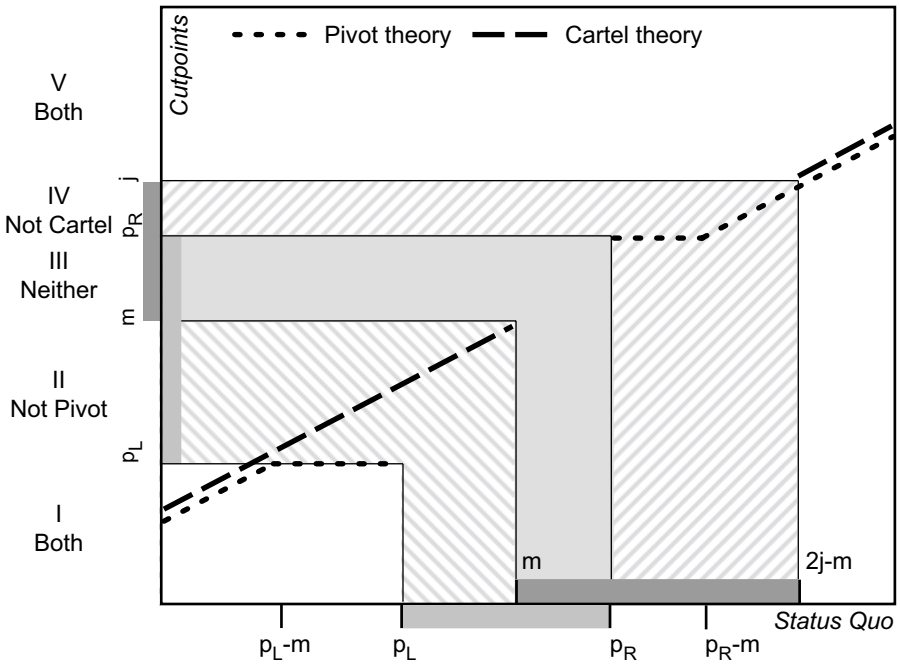

Fig. 2. Cutpoint consistency in the cartel and pivot theories

cleanly between the two theories. Interval II defines cutpoints that are inconsistent with pivot theory but consistent with cartel theory. Substantively, these cutpoints are on the minority party side of the median voter but less extreme than the veto or filibuster pivots.[4] Conversely, interval IV defines cutpoints that are inconsistent with cartel theory but consistent with pivot theory. These cutpoints are on the majority party side of the median and more *extreme* than that side's veto or filibuster pivot but less extreme than the filibuster or veto pivot. This interval is typically very small.[5]

[4] The relevant pivot depends on the location of the president. In unified government the filibuster is the binding constraint; in divided government, the veto pivot is.

[5] The median and random models can generate any possible cutpoint. In the median model every final-passage yea location is, by construction, at the median legislator's ideal point, but since the status quo or nay locations are unconstrained any cutpoint is possible. In the random model, cutpoints are normally distributed because they are simply midpoints between two normally distributed random variables.

3 Data

Within this theoretical framework, a direct test of the theories with gridlock intervals via cutpoints is possible. The ideal point estimates allow us to construct Senate specific intervals that correspond to intervals I-V on the vertical axis of Fig. 2. In the simplest and most direct test, then, we can simply compare theories based on their respective frequencies of inconsistent (theory-contradicting) cutpoint estimates. The master dataset includes all roll-call votes in the U.S. Senate for the 80th through the 107th Congresses. For 86 percent of these roll calls, DW-Nominate ideal point and cutpoint estimates were calculated.[6]

It is not abundantly clear whether all or a subset of votes should be analyzed. An emerging convention in roll call analysis is to focus on final passage votes. Theoretically, one reason is that only once voting reaches its last stage does it simplify to a binary choice, when we can then be certain that sophisticated voting does not occur. If the voting on an individual policy were not sincere, then the estimated cutpoint may not accurately measure what we think it does. On the other hand, the argument for excluding votes upon which voting may be strategic exposes a limitation of using estimates from Nominate or any other scaling procedure that uses votes other than final passage types. If the sophisticated-voting-based defense for selecting only final passage votes is compelling, then it must be that sophisticated voting occurs with sufficient regularity that one should exclude pre-passage votes. However, if sophisticated voting passes this threshold, then tests using Nominate (or other large-N) ratings must be treated with suspicion, too, because our usual interpretations of the recovered ideal points and cutpoints rest on an assumption that the inputs for the rating scheme are sincere votes.

Even if all voting is sincere, however, another reason to focus on final passage votes is that the theories describe the policy alternatives that ultimately emerge from the legislative process and are silent about the amendments that would be observed on the equilibrium path. Moreover, pre-passage votes would typically not be between the equilibrium proposal $b^*(q)$ and the status quo q. In that case, we would not necessarily be able to cleanly partition the policy

[6] We use the estimates provided by Keith Poole and available at the web site: http://k7moa.uh.edu/. Even though use of the DW-Nominate yea and nay location estimates would yield a more direct test of the mappings in Fig. 2, there is a compelling reason to use the cutpoint estimates instead of the yea and nay location estimates. The classification procedure that recovers the ideal points also recovers the bill cutpoints (and yea and nay locations). It is well known, however, that yea and nay locations are not well-identified. Procedures like DW-Nominate recover these parameters by leveraging off of the postulated functional forms of the loss function and error distributions. In contrast the cutpoints are the dual of the ideal points and are as reliably estimated as the ideal points. Accordingly, the cutpoint estimates are much more robust than the yea and nay location estimates. See the appendix of [30] for more information about DW-Nominate.

space as in Fig. 2. For example, in a pre-passage vote, two alternatives (perhaps an amendment and a substitute amendment) may produce a cutpoint within the gridlock interval but the alternative that survives this vote would later lose to $b^*(q)$.

A related complication is the absence of a clear consensus about how "votes on final passage" should be defined and operationalized. We will consider three definitions: (A) expansive, (B) moderate, and (C) restrictive. The expansive definition (A) considers any of the following a final-passage vote: relatively unambiguous motions (such as "to pass S.1234"), a vote on a veto override, a vote on a motion to suspend the rules in order to pass a measure (uncommon in the Senate), final votes on resolutions (even though most of these are measures of strictly intra-Senate relevance), final votes on procedural initiatives (e.g., to initiate investigations, to discharge a committee), final votes on nonbinding measures such as resolutions of disapproval), votes on engrossment and third reading of bills, votes on conference reports, and on amendments between the chambers. The moderate definition (B) tightens up these criteria (and reduces the number of false positives) by not counting as final passage votes: resolutions, nonbinding measures, procedural votes, and amendments between the chambers. The most restrictive definition (C) does not count any of the above-listed as final passage votes but rather demands that legislation have the possibility of becoming Public Law. The trade-offs are more or less obvious: the more expansive is the definition, the more likely it is that not-really-final-passage votes slip in; the more restrictive is the definition, the more likely it is that truly substantive votes that are de facto the last action taken by the chamber on significant legislation are omitted.[7]

We address these complications by reporting various combinations of votes, ranging from all roll calls on the inclusive end of the spectrum ($N = 15,943$) to narrowly defined final passage votes on the exclusive end of the spectrum ($N = 1,912$).

4 A Direct Test

All cutpoints are, of course, consistent with the median and random models, because neither of these has a gridlock interval, which is a necessary condition for refutation using the direct cutpoint approach. The analysis discussed in this section, accordingly, focuses on the refutable implications of only the cartel and pivot theories. Table 1 shows the distribution of roll calls across the cutpoint intervals as defined on the vertical axis of Fig. 2. While potentially significant differences between all roll calls and final-passage roll calls are evident, it seems not to matter whether expansive or restrictive definitions of final passage votes are assessed. Unless differences are striking, we henceforth

[7] Selection of roll calls was implemented via Microsoft Access queries of ICPSR and CQ Washington Alert roll call records.

Table 1. Cutpoint-inconsistency of two theories

Inconsistent with:	I Neither	II Pivot	III Both	IV Cartel	V Neither	N
All roll calls	5,940	4,067	2,439	573	2924	15,943
	37%	26%	15%	3.6%	18%	
Final passage A	1,611	511	252	47	619	3,040
	53%	17%	8.3%	1.5%	20%	
Final passage B	1,395	381	190	37	510	2,513
	56%	15%	7.6%	1.5%	20 %	
Final passage C	1,076	266	151	29	390	1,912
	56%	14%	7.9%	1.5%	20%	

report results only for the moderate definition (B). Columns I and V of Table 1 show clearly that most cutpoints are extreme relative to pivotal Senators' ideal points. Even though fewer than 30 percent of cutpoints lie in the interior segments, however, these observations have discriminatory potential. In this phase of the analysis, the findings seem clearly to favor the cartel over the pivot theory. Cutpoints that are inconsistent with both theories (interval III observations) occur roughly ten percent of the time, but the remaining two intervals (II and IV), in which only one theory is rejected, are consistently in favor of the partisan model.

As noted above, intervals II and IV are necessarily on the opposite sides of the median voter in the legislature, with the rejection interval IV for the cartel theory residing considerably farther from the median than the median-adjacent interval II. Furthermore, interval II is on the minority party side of the median, while interval IV is on the majority party side. Accordingly, we can view the comparative frequencies in these two intervals as a record of which party tends to win the figurative policy tug-of-war. This interpretation is similar to the increasingly prominent research methodology focusing on "roll rates:" the proportion of times that one party defeats another as indicated by intra-party homogeneity or cohesiveness and inter-party opposition. Upon first glance, the two methods – roll rate analysis and cutpoint analysis – seem to validate one another. Here, as elsewhere, the ratios range roughly from 4:1 to 8:1 (see also [15] and [16]).

5 Strong Consistency

Although the direct horse-race setup is intuitive and convenient, things are not as simple as it seems upon first glance. Three caveats need to be addressed. First, the data in Table 1 may include a large number of false positives: namely, roll calls in which the cutpoint may be consistent with the theory but in which the bill and status quo locations are not consistent with the theory. Second, the

inclusive data set may be distorted by the presence of motions that do not pass, because the theories are silent about policy-inconsequential proposals and voting. Third, and most challenging, the two theories differ substantially in the degree to which they are exposed to refutation. Specifically, cutpoints throughout the entire gridlock interval in the pivot theory are inconsistent with the theory, while cutpoints throughout only half of the gridlock interval in the cartel theory are inconsistent with it. As an empirical matter, the average ratio of pivot-to-cartel widths in terms of rejection intervals is 1.66 to 1. A further and even more subtle bias exists in the likely case that cutpoints are likely to be approximately normal, in which case the centrality of the pivot rejection interval relative to the cartel rejection interval exacerbates the asymmetry in exposure of the theories. Fortunately, additional analysis can address these obstacles, at least in part.

Table 2 addresses the false positives obstacle by adding two filters to the location-based definition of consistency employed in Table 1. Specifically, three percentages are reported for each of three theories: median, cartel, and pivot. The first such number is the consistency rate based solely on whether the estimated cutpoint lies outside of the respective theory's rejection interval. So, for example, because the median voter theory has no rejection interval, its consistency rate is necessarily 100 percent. For the cartel theory, the location-based consistency rate is the sum of percentages from Table 1, row 1, columns I, II, and V. For the pivot theory, the rate is the sum from columns I, IV, and V. For both sets of votes – all roll calls or type B final passage votes – the ordering is necessarily the same as in Table 1 because nothing has changed about the data. Specifically, the irrefutable median voter theory necessarily performs best, and the cartel theory outperforms the pivot theory.

The filters are applied in the second and third columns (for each of two sets of roll calls). For all three theories, when votes occur, the consequence should be convergence toward the median voter's ideal point on the primary dimension. Although we share Poole and Rosenthal's skepticism about the point precision of their estimates of yea and nay locations, it seems clear we may assume that the estimates are quite accurate in assigning the direction of the yea alternative. Therefore, in the columns titled Direction, we tabulate the percentage of roll calls both whose cutpoints are theory-consistent and whose yea alternative lies closer to the median than the nay location. That is, the conditions are each necessary and cumulatively sufficient.

Table 2. Cutpoint-consistency under increasingly stringent definitions

	All roll calls			Final passage roll calls		
	Location	Direction	Success	Location	Direction	Success
Median voter theory	100%	59%	50%	100%	86%	81%
Cartel theory	81%	49%	44%	91%	81%	77%
Pivotal politics theory	59%	37%	33%	77%	70%	68%

The result is a shocking drop-off in the success rates of all three theories, particularly in the larger data set. The corresponding drop-offs from the direction criterion in final passage votes is much lower – only about 10 percent. This suggests that focusing on final passage votes may, indeed, yield somewhat more confident inferences than the much more inclusive sample.

Finally, the third and sixth columns add the requirement that the motion obtain the requisite number of votes to carry. The drop-off here is lower than when the direction criterion was applied, but the final result is discouraging in the case of the larger data set. The median voter theory does no better than a flip of a fair coin in predicting whether motions were successful and moved the status quo closer to the median voter. The cartel and pivot theories perform even worse. More precisely, at least half and as many as two thirds of roll calls in the Senate failed to meet the most stringent cumulative criterion based on cutpoint locations, direction of change, and success. The situation is much better, however, for final passage votes. Here, the drop-off is minor by comparison, and the most-stringent-criterion findings are more sanguine. The range of prediction rates across theories is 68 percent for pivot theory to 81 percent for median voter theory. In summary, it seems that this class of models does not perform well over the entire set of roll calls, but it does perform well on final passage votes.

6 Exposure-adjusted Consistency

Although the more stringent criteria for selecting observations tend to bring together the three models in terms of predictive performance on final passage votes, the models are still not quite on even footing with the previous forms of analysis. Specifically, we have still not addressed the fact that, on average, the pivot rejection region is 1.6 times the size of the cartel rejection region. An implication of this fact is that if the cutpoints were uniformly generated, we would expect the pivot model to perform only 62.5 percent as well as the cartel model. In fact, the number of cutpoints that are inconsistent with pivot is not far from 1.6 times the number of cutpoints that are inconsistent with cartel theory. Twenty percent of the cutpoints are inconsistent with the cartel theory, and 30 percent are inconsistent with the pivot theory, so the empirical ratio is 1.5.

The problem of differential exposure is the primary reason for introducing the random model. The approach is to compare three cutpoint-generating processes: cartel-theoretic, pivot-theoretic, and an atheoretic random draw from a normal distribution. Because the null model assigns more cutpoints to large and central rejection intervals than to small and one-sided intervals, a comparison of the actual percentage of cutpoints in a model's rejection region to the expected percentage of cutpoints that the null model would put in the model's rejection region makes a neutral test possible. Though not a formal statistical test, this procedure is conceptually similar. It captures the idea that,

Table 3. Exposure-adjusted consistency

	Pivot	Cartel
Null	73%	86%
Observed	82%	94%
Improvement	12%	10%

if a theory does not explain the data appreciably better than a simple naïve model, then we can conclude that the model's fit is underwhelming. In the spirit of comparing logit regression models to a null of tossing an unfair coin, we compare pivot and cartel to a similarly exposure-calibrated null model.

We utilize the normal distribution because of an a priori affinity which is reinforced by the actual distribution of cutpoints. We rely on the data to choose the mean and variance parameters for each Congress's distribution. For each Congress, therefore, the null model is that the cutpoints are drawn from a normal distribution where the first and second moments correspond to the estimated sample moments for that Congress's cutpoints. For this analysis we focus on the votes satisfying the moderate (B) final passage criterion.

Table 3 presents the first of two sets of results that address the differential exposure problem. The first row gives the percentage of roll calls whose cutpoints would be consistent with the respective theory, if cutpoints were distributed normally.[8] As asserted earlier, the cartel theory's expected percentage (86) is greater than the pivot theory's (73), because the former's exposure is not as great as the latter's. The second row gives the actual observed percentages for the two theories. Finally, the third row calculates the percentage improvements in prediction over the null or baseline percentage. The difference between theories, while small (2 percent), is in the opposite direction of the direct-test finding. In other words, relative to expectations, the pivot theory slightly outperforms the cartel theory.

A second, closely related method for assessing the exposure-adjusted strengths or limitations of the theories is based on Congress-level observa-

[8] Table 3 uses weighted averages across Congresses. The results are only very slightly different than the unweighted averages. We again require that consistent votes move in the right direction throughout this section. While this may seem unfair to the cartel and pivot models because the null model faces no such test, in fact this concern is unwarranted. Recall, that the real null model is that the yea and nay locations are themselves normally distributed. As a consequence the cutpoints are normal (as the weighted average of two normal variables). In addition, by altering the means and variances of the yea and nay location distributions, it is possible to alter the expected percentage of yea and nay locations which would satisfy the directional tests. Accordingly, for any desired normal distribution of cutpoints, we can construct distributions for yea and nay locations that give us the appropriate distribution of cutpoints and would do exactly as well as the actual data on the directional tests.

tions. Specifically, for each Congress we calculate a proportional improvement measure of each theory over its baseline expectation, given the width and location of its gridlock interval. The scatterplot in Fig. 3 displays this information. The horizontal axis represents pivot-theoretic improvement over the null, while the vertical axis measures the corresponding values for the cartel model. Therefore, observations that are to the northeast of the point $(0, 0)$ represent Congresses in which both models improve on the random null. Observations below (above) the 45 degree line represent congresses in which the pivot's relative performance is better (worse) than that of the cartel model's performance. Consistent with Table 3, Fig. 3 indicates that the pivot theory outperforms the cartel theory in most Congresses – 19 of 27 to be precise. However, in four Congresses the performance of the pivot theory is abysmal, and, when standard deviations are calculated for the two variables, it is clear that one cannot reject the hypothesis that the theories' average proportional-improvement rates are equal.

The results as a whole, then, are mixed and inconclusive regarding cartel and pivot theories, but either of these seems to perform better than the median voter theory. If there is a surprise performance in these exercises, it would seem to be the random model. Can we confidently say that the two gridlock theories are responsible for observed deviations from normality in the distributions of cutpoints? Although a rigorous statistical test is beyond the scope of the

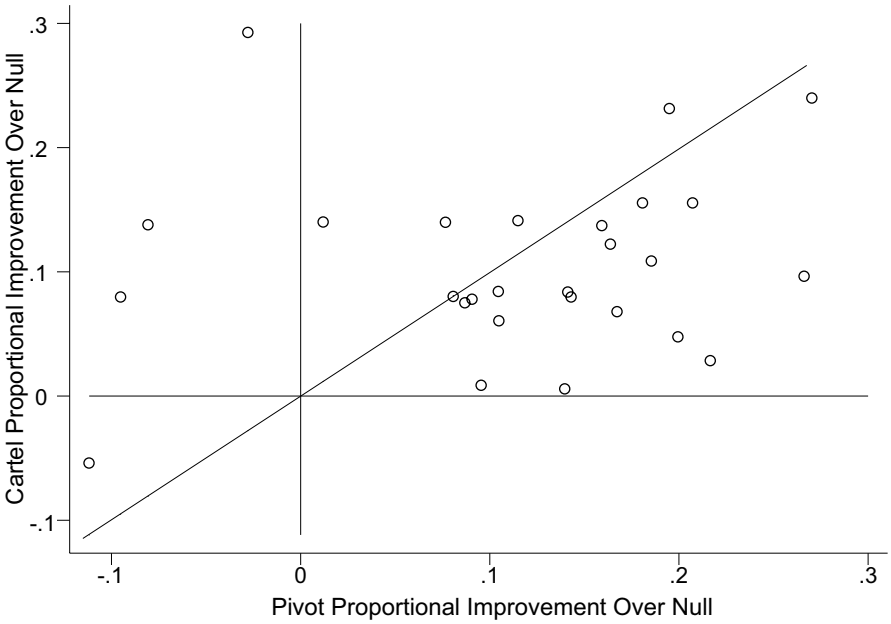

Fig. 3. Comparative improvements in predictions over a naïve null model

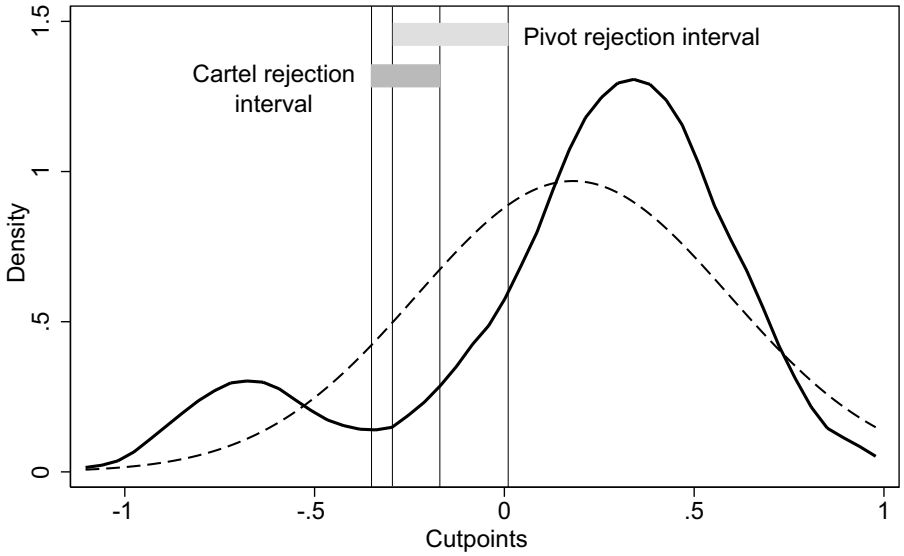

Fig. 4. Kernel and normal distributions of cutpoints, 94th Congress

analysis, we can get some insight into this question simply by observing kernel density plots of the cutpoint distributions.

Presenting all such distributions would unnecessarily clutter the presentation, so we select one congress as an example (noting that nearly all of the congresses are qualitatively quite similar). The kernel density plot of the final passage cutpoints for the 94th Congress is shown in Fig. 4.[9] The corresponding normal distribution for this set of cutpoints is superimposed on the figure, and the vertical lines correspond to the intervals labeled I-V in Fig. 2. Neither models' rejection region consistently matches up with the troughs in the data. Nevertheless, there is striking evidence that troughs are present and qualitatively they are in the correct part of the policy space. Overall, this statement applies to 23 out of 27 congresses which seems to be genuine support for some type of gridlock theory and, likewise, more-than-coincidental evidence against the baseline random cutpoint model.

7 Discussion

Our conclusions are partly methodological and partly substantive.

As a methodological exercise, the study has been humbling, both in comparison with other similar research and throughout the various phases of analysis in this project. Prior research, of course, has studied partisan and nonpartisan theories of legislative choice in considerable detail. Researchers have

[9] Again, these cutpoints satisfy type B restrictiveness and the direction criterion.

tended to come down in favor of one theory over the other with considerable confidence. Often, however, the reason for their confidence is that "the other" theory is either ignored, or not put on equal footing with the focal theory. The intended neutral stance taken here is based on the premise that, if discrimination is to occur convincingly, the high degree of observational equivalence of pivot and cartel theories requires joint, neutral tests. In this project, we have pursued joint tests of two sorts: direct and indirect. The direct test – a comparison of consistency (or inconsistency) of cutpoints with the respective theories – suggests quite strongly that the cartel theory is better than the pivot theory. However, the direct test can be questioned in terms of the equal-footing criterion. Furthermore, it is unclear whether either gridlock-based theory is significantly better than the simple median-voter baseline.

A series of indirect tests was then devised, and the results were painfully inconclusive. The pivotal politics theory fares better than the cartel theory relative to their respective baseline expectations. However, the variance in performance levels is sufficiently large that we cannot be confident that the difference is real. While this analysis suggests somewhat more convincingly that gridlock-based theories provide marginal value over median voter theory, their marginal value over the random-normal null model is questionable. In short, the visuals are suggestive, but the data remains ambiguous. In light of all of the above, the study at least underscores the value of multiple methodological approaches prior to making definitive conclusions. Kramer [19] reminds us that well formulated theoretical propositions are fragile. We have focused on well-formulated theoretical propositions and illustrated their fragility.

As a substantive exercise, the study fails in the obvious respect that discrimination among the four models proves to be much more difficult than anticipated. The interplay between theory, method, and data, however, is such that we can offer some remarks about why our findings are inconclusive. The two primary possibilities are problematic data or problematic theory.

To dwell on the first of these is admittedly cowardly insofar as we ought not to have conducted the tests had we not thought the data were up to the task. Still, the possibility that we were mistaken deserves some retrospection.

The nature of the exercise – here and elsewhere – is to assume not only that there is a primary dimension of conflict on all issues but also that legislators are arrayed in the same way on a primary dimension on all issues. Clearly, this stretches credulity. One important avenue of future methodological inquiry is to assess rigorously the degree and nature of bias on various tests of errors or imperfections in the array of preferences projected from a multidimensional onto a unidimensional space. Although this may seem like a theory problem rather than a data problem, there is a subtle reason that it is more of the latter than the former. Unidimensional theories assume only that preferences exist and are well behaved on a single dimension *in any given choice situation*. They do not assume, as does the implementation of tests of the sort we conducted, that well-defined preferences of Senators maintain the same orderings and locations on the *same* primary dimension *across all roll calls*. In other words,

unidimensional theories can, in principle, predict perfectly well even in a truly multidimensional world.

The second possibility – inadequate theory – brings us back to the concerns and contributions of our late colleague Jeffrey Banks. Banks and Duggan [8, p. 26] conclude:

> While we have considered mainly foundational issues here, we have proposed a general framework in which more substantive questions, about the nature of public goods provided or the coalitions that form to pass proposals, for example, can be taken up in special cases with more structure.

The pivot and cartel models are exactly that: special cases with structures that are justified by observations about the nature of legislative institutions. Our examination of these examples leads to the recommendation that additional structures and/or special cases ought to be investigated. We share Banks's hope and expectation that future models will improve upon those considered here and shed a brighter light on legislative policy-making.

References

1. Aldrich J (1994) A model of a legislature with two parties and a committee system. Legislative Studies Quarterly 19:313–340
2. Aldrich J (1995) Why parties? The origin and transformation of party politics in America. University of Chicago Press, Chicago
3. Aldrich J, Rohde D (1998) Measuring conditional party government. Presented at the Annual Meeting of the Midwest Political Science Association, Chicago
4. Austen-Smith D (1987) Sophisticated sincerity: voting over endogenous agendas. American Political Science Review 81:1323–1330
5. Banks JS (1985) Sophisticated voting outcomes and agenda control. Social Choice and Welfare 1:295–306
6. Banks JS (1995) Singularity theory and core existence in the spatial model. Journal of Mathematical Economics 24:523–36
7. Banks JS, Duggan J (2000) A bargaining model of collective choice. American Political Science Review 94:73–88
8. Banks JS, Duggan J (2003) A bargaining model of legislative policy-making. Typescript, University of Rochester
9. Banks J, Gasmi F (1987) Endogenous agenda formation in three person committees. Social Choice and Welfare 4:133–152
10. Black D (1958) The theory of committees and elections. Cambridge University Press, London
11. Brady DW, Volden C (1998) Revolving gridlock: politics and policy from Carter to Clinton. Westview Press, New York
12. Clinton JD, Meirowitz A (2001) Agenda constrained legislator ideal points and the spatial voting model. Political Analysis 9:242–259
13. Clinton JD, Meirowitz A (2003) Integrating voting theory and roll call analysis: a framework. Political Analysis 11:381–396

14. Cox GW, McCubbins MD (1993) Legislative leviathan: party government in the House. University of California Press, Berkeley
15. Cox GW, McCubbins MD (2002) Agenda power in the US House of Representatives, 1877 to 1986. In: Brady DW, McCubbins MD (eds) Party, process and political change in Congress: new perspectives on the history of Congress. Stanford University Press, Stanford
16. Groseclose T, Snyder JM (2003) Interpreting the coefficient of party influence: comment on Krehbiel. Political Analysis 11:104–107
17. Groseclose T, Levitt S, Snyder JM (1999) Comparing interest group scores across time and chambers: adjusted ADA scores for the U.S. Congress. American Political Science Review 93:33–50
18. Heckman J, Snyder JM (1997) Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. RAND Journal of Economics 28:142–189
19. Kramer GH (1986) Political science as science. In: Weisberg H (ed) Political science: the science of politics. Agathon Press, New York
20. Krehbiel K (1993) Where's the party? British Journal of Political Science 23:235–266
21. Krehbiel K (1996) Institutional and partisan sources of gridlock: a theory of divided and unified government. Journal of Theoretical Politics 8:7–40
22. Krehbiel K (1998) Pivotal politics: a theory of U.S. lawmaking. University of Chicago Press, Chicago
23. Krehbiel K (2000) Party discipline and measures of partisanship. American Journal of Political Science 44:206-221
24. Krehbiel K (2003) Cartels and pivots reconsidered. In: Adler ES, Lapinski J (eds) Macropolitics of Congress
25. Krehbiel K, Meirowitz A (2002) Minority rights and majority power: theoretical consequences of the motion to recommit. Legislative Studies Quarterly 27:191–217
26. McCarty N, Poole KT, Rosenthal H (2001) The hunt for party influence in Congress. American Political Science Review 95:673–693
27. McKelvey R, Schofield N (1987) Generalized symmetry conditions at a core point. Econometrica 55:923–934
28. Plott C (1967) A notion of equilibrium and its possibility under majority rule. American Economic Review 57:787–806
29. Poole KT, Rosenthal H (1985) A spatial model for legislative roll call analysis. American Journal of Political Science 29:357–384
30. Poole KT, Rosenthal H (1997) Congress: a political-economic history of roll call voting. Oxford University Press, Oxford
31. Rohde DW (1991) Parties and leaders in the postreform House. University of Chicago Press, Chicago
32. Saari D (1997) The generic existence of a core for q-rules. Economic Theory 9:219–260
33. Schickler E (2000) Institutional change in the House of Representatives, 1867-1998: a test of partisan and ideological power balance models. American Political Science Review 94:269–289
34. Shepsle K, Weingast B (1987) The institutional foundations of committee power. American Political Science Review 81:85–104
35. Sinclair B (1997) Unorthodox lawmaking: new legislative processes in the U.S. Congress. Congressional Quarterly Press, Washington, D.C.

36. Snyder JM (1992) Committee power, structure induced equilibria, and roll call votes. American Journal of Political Science 36:1–30
37. Snyder JM, Groseclose T (2000) Party influence and congressional roll-call voting. American Journal of Political Science 44:193–211
38. Weingast BR (1989) Floor behavior in the U.S. Congress: committee power under the open rule. American Political Science Review 83:795–816