

# Integrating Voting Theory and Roll Call Analysis: A Framework

Joshua D. Clinton and Adam Meirowitz

Department of Politics,  
Princeton University, Princeton, NJ 08544  
e-mail: [clinton@princeton.edu](mailto:clinton@princeton.edu)  
e-mail: [ameirowi@princeton.edu](mailto:ameirowi@princeton.edu)

Scholars of legislative studies typically use ideal point estimates from scaling procedures to test theories of legislative politics. We contend that theory and methods may be better integrated by directly incorporating *maintained* and *to be tested hypotheses* in the statistical model used to estimate legislator preferences. In this view of theory and estimation, formal modeling (1) provides auxiliary assumptions that serve as constraints in the estimation process, and (2) generates testable predictions. The estimation and hypothesis testing procedure uses roll call data to evaluate the validity of theoretically derived *to be tested hypotheses* in a world where *maintained hypotheses* are presumed true. We articulate the approach using the language of statistical inference (both frequentist and Bayesian). The approach is demonstrated in analyses of the well-studied Powell amendment to the federal aid-to-education bill in the 84th House and the Compromise of 1790 in the 1st House.

## 1 Introduction

The typical empirical study of legislative behavior uses ideal point estimates (usually interest group or NOMINATE scores) as either independent or dependent variables in a regression equation. While much can be learned from a two-stage approach that first generates estimates of legislator-induced preferences and then uses the estimates to address substantive or theoretical questions, several shortcomings result. First, as Snyder (1992), Poole and Rosenthal (1997), Herron (1999), and Clinton et al. (2003) argue, legislator preference measures resulting from any statistical (or counting) procedure are imperfect and “second-stage” models must account for this error for valid inferences. A second, and less-remarked upon issue is that the statistical model generating the ideal point estimates may not be neutral with respect to the hypotheses being tested in the second-stage model. For example, one might worry about the use of “off-the-shelf” estimates of legislator voting behavior to test theories of partisan influence in legislatures given that the error structure implicit in scaling procedures excludes the possibility of vote buying.

Our thesis is that theory and methods pertaining to legislative voting behavior can be better integrated by directly incorporating *maintained* and *to be tested hypotheses* in the statistical model used to estimate induced legislator preferences. In this view of theory and estimation, formal modeling (1) provides auxiliary assumptions that serve as constraints in the estimation process, and (2) generates testable predictions. Estimation and hypothesis

testing use data to evaluate the validity of *to be tested hypotheses* in the world where *maintained hypotheses* are presumed true.

In this article we develop a straightforward framework that highlights the relationship between theory and estimation. The framework demonstrates how particular accounts and/or theories of legislative voting behavior can be tested using estimates that are congruent with the model being tested and it clarifies the role of auxiliary assumptions, maintained hypotheses, testable predictions, and estimated parameters. The approach can be used to evaluate competing explanations of legislative behavior (e.g., strategic vs. sincere voting) while incorporating auxiliary assumptions common to the explanations. Also, the additional information in theoretically motivated constraints may assist in the identification of the statistical model.<sup>1</sup>

The power of our approach comes at a cost; scholars must translate legislative agendas and theoretical accounts of legislative politics into parameter constraints on bill locations, ideal points, and/or moments of error terms. Furthermore, statistical procedures must be customized instead of using “off-the-shelf” estimates/techniques. However, McKelvey’s comments in the introduction outline the necessity of incurring the cost associated with the “applied theory” we advocate in this article. The basic approach involves (1) constructing an account of the legislative history that isolates two sets of hypotheses: those that are *to be maintained* and those that are *to be tested*; (2) translating hypotheses into constraints on parameters in the statistical model; (3) estimating the model assuming that the identifying constraints implied by the *maintained hypotheses* are true; (4) and testing the constraints corresponding to the *to be tested hypotheses*.

The outline of the article is as follows. In the next subsection an alternative approach to combining theory and data is reviewed. Section 2 presents the framework both informally and formally in the language of frequentist and Bayesian inference. Section 3 uses the approach to investigate strategic voting on (1) the Powell amendment to the aid-to-education bill in the 84th House as well as (2) the Compromise of 1790 in the 1st House. Section 5 concludes.

### 1.1 *Related Approaches to Combining Theory and Data*

Recent advances in combining game theoretic models and econometric procedures seem quite promising, but they have not yet been applied to studies of legislative voting. Quantal response equilibria (hereafter QRE) (McKelvey and Palfrey 1995, 1996, 1998), for example, connects predictive theory and statistical estimation. The basic approach involves specifying a game theoretic model of group interaction that constrains agents to select probabilistic strategies. As agents select better strategies with higher probability, QRE converge to Nash Equilibria as the required noise in agent strategies vanishes. In contrast to simply augmenting deterministic equilibria with noise, in QRE agents anticipate that other players’ actions are also noisy. This innovation is attractive because a QRE for a given game generates a likelihood function for observable data. Maximum likelihood estimation of the parameters permits inference about explicit components of the game.

QRE has been applied by Signorino (1999) to study a large number of country pairings in security studies. In experimental settings, insights about the nature of strategic interaction

<sup>1</sup>For example, Londregan (1999) utilizes the additional information contained in the identity of proposers and Clinton and Meirowitz (2001) note the availability of information contained in the relationship between proposals being voted upon. Neither work integrates theory (predictions or assumptions) into the estimation of legislator preferences.

involving questions central to voting theory have been recovered (e.g., Guarnaschelli et al. 2000). Quantal response works well in settings where many observations of a game involving a small number of players are observed.

Unfortunately, procedures like QRE are not likely to serve scholars of legislative voting behavior despite a voluminous and growing formal literature because of the nature of the data-generating process. In such settings, one typically confronts a matrix of roll call votes representing the observable behavior in response to a given legislative agenda. If voting on the agenda is viewed as a game, it is immediately clear that only one realization is observed. While scholars of security studies can argue that all dyads (or triads) interact according to the same extensive form game, and while experimental game theorists can replicate a game as many times as funding will allow, scholars of legislative voting behavior are constrained to observe a single repetition of a particular agenda (or game). In short, the problem faced is much different than those previously handled using QRE.

## 2 A Framework

Theories of legislative voting and roll call analysis procedures (generally) involve a finite set of legislators  $L$ . Each legislator  $l$  has preferences over the policy space  $X \subset \mathbb{R}^n$  represented by the utility function  $u(\cdot; x_l)$ , where  $x_l$  is a vector of legislator-specific parameters (usually the legislator's ideal point in  $X$ ). A sequence of  $T$  votes occurs with each legislator casting a binary ballot in each vote.<sup>2</sup> The model is closed by associating each vote option with a policy location. In theoretical work, these locations are often fixed (Plott 1967; McKelvey 1976; Enelow and Hinich 1984; McKelvey and Schofield 1987), but they may be endogenous—resulting from strategic proposal making (Austen-Smith 1987; Banks and Gasmi 1987; Banks and Duggan 2000; Krehbiel and Meirowitz 2002). In standard roll call estimation models, the yea and nay locations (denoted  $y_t \in X$ ,  $n_t \in X$  respectively) associated with each roll call vote  $t$  are “free” parameters, and either the parameters or functions of the parameters (e.g., midpoints/cutting hyperplanes) are estimated. The only other difference between theoretical and statistical spatial models is that the latter introduces a stochastic shock/error to agents' utility functions. Estimation proceeds by specifying a latent utility model and estimating agent-specific ideal points and vote-specific parameters.

Existing statistical models typically assume that yea and nay locations are unrelated across votes and that individual voting errors are i.i.d. with zero mean.<sup>3</sup> In contrast, many theories about legislative voting behavior either include assumptions or generate predictions that violate these assumptions. For example, strategic voting implies strong constraints on perceived bill locations. The idea of sophisticated equivalents (McKelvey and Niemi 1978) implies that legislators perceive a relationship between the world resulting from a decision on one vote and the world resulting from a subsequent decision on a subsequent vote when making the initial decision.<sup>4</sup>

To clarify the argument, Fig. 1 presents a simple agenda tree. Suppose only three votes occur and only the second vote fails (i.e., votes 1 and 3 pass). In estimating these three roll call votes one would typically assume that there are three yea locations ( $y_1, y_2, y_3$ ) and three nay locations ( $n_1, n_2, n_3$ ) to be estimated in addition to the most-preferred (“ideal”) point  $x_i$  of each legislator. However, legislative voting theory says that if agents are operating

<sup>2</sup>Abstentions and missing data occur but are not considered here.

<sup>3</sup>But see Poole 2000.

<sup>4</sup>Note that this point is conceptually distinct from the claim by Clinton and Meirowitz (2001) that the legislative agenda contains useful information. The point here is that theories provide information *in addition to* information contained in the legislative agenda.

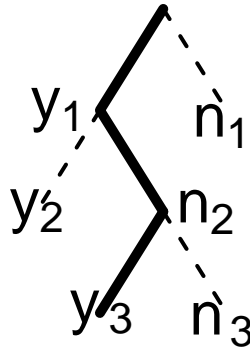


Fig. 1 Illustration of sophisticated equivalents.

under full information and behaving strategically, then the utility associated with passing the first vote is equivalent to the utility associated with the failure of the second motion (following passage of the first), and both of these utilities coincide with the utility resulting from the passage of the first motion, failure of the second, and passage of the third. In other words, the perceived bill locations  $y_1$ ,  $n_2$ ,  $y_3$  must coincide. Theory has less to say about the remaining locations as they are off the equilibrium path.

Combining theory-driven observations with scaling procedures leads to the conclusion that if one wants to analyze voting behavior on these three votes in a manner consistent with strategic voting and full information, then the estimates should be constrained to satisfy the *maintained hypotheses* of the theoretical account (i.e.,  $y_1 = n_2 = y_3$ ). Of course, one may not be content with the idea that voting is strategic or that legislators face little uncertainty (e.g., Denzau et al. 1985). In this case, these questions may be addressed by viewing these constraints as *testable hypotheses*. In fact, much of what we have to say about connections between theory and estimation boils down to the following prescription: in calibrating/testing theory, the estimation/model specification should satisfy as many relevant *maintained hypotheses* as possible, and hypothesis tests involve recovering either a distribution over parameter values or the probability that a *to be tested hypothesis* is true.

A second possible problem that may result from failing to consider theoretical implications when estimating relates to the error structure in latent utility models. Partisan influence or “vote-buying” can produce behavioral similarity in agents despite differences in policy preferences. Models of vote buying (Snyder 1991; Groseclose and Snyder 1996) make predictions about the relationship between legislator preferences and external pressure (possibly in the form of interest group contributions). These predictions can be interpreted as additive shocks to the latent utility for certain legislators on certain votes. More precisely, while standard estimation models postulate that the probability of a yea vote for legislator  $l$  on roll call  $t$  is  $F(u(y_t; x_l) - u(n_t; x_l))$ , where  $F(\cdot)$  is the cumulative distribution for the  $lt$  error, vote-buying has been interpreted to mean that the probability of a yea vote for legislator  $l$  on roll call  $t$  is  $F(b_{lt} + u(y_t; x_l) - u(n_t; x_l))$ , where  $b_{lt} \in \mathbb{R}^1$  is legislator  $l$ ’s “bribe” for voting yea on roll call  $t$ . To test the hypothesis that an interest group influenced a particular set of legislators  $C \subset L$  on a particular set of votes (say  $t, t+1, t+2, t+3$ ), one could augment the standard estimation model. Defining and estimating a term  $b$  which is 0 for roll calls not involving legislators in  $C$  or votes  $t$  through  $t+3$  and endogenous for the remaining votes is analogous to estimating the common mean of the shocks for the votes by legislators in  $C$  on roll calls  $t$  through  $t+3$ . A statistical test of whether  $b$  is

significantly different from 0 can be interpreted as a test of the hypothesis that the group influenced the legislators in  $C$  on the votes  $t$  through  $t + 3$ . Recent articles on measuring party influence on nonlopsided legislation (Snyder and Groseclose 2000; McCarty et al. 2001) utilize variations of this theme. In these instances the set  $C$  defines a party and the votes of interest are those that are not lopsided.

While this discussion of strategic voting and vote-buying is much simplified, it suggests a means of relating theories of legislative voting behavior to roll call scaling procedures. Specifically, the roll call estimation model is augmented with two types of constraints. *Maintained hypotheses* are constraints resulting from knowledge about the legislative agenda governing a sequence of roll calls, the substantive nature of the policies, or information about the preferences of the legislators (for example). *To be tested hypotheses* are constraints derived from a hypothetical account of legislator behavior on the sequence of votes. *Maintained hypotheses* are used to specify a likelihood function that depends on parameters (some of which might be the subject of *to be tested hypotheses*). Estimation determines the probability that the *to be tested hypotheses* are true given the data and *maintained hypotheses*. Although this paradigm is identical to the familiar Neyman–Pearson framework, our contribution is suggesting how legislative theory and roll call analysis can be more tightly embedded in this framework. We now formalize the logic.

## 2.1 The Statistical Model

The approach consists of specifying a statistical model of roll call voting behavior that requires parameters to satisfy the *maintained hypotheses*, and performing hypothesis tests on the relationships implied by the *to be tested hypotheses*. Formally, let  $\Theta$  denote the parameter space of the empirical model and let  $M$  denote the set of possible observable data sets (typically  $m \in M$  is a roll call matrix with element  $m_{it}$  taking on the value 1(0) if  $l$  voted yea (nay) on  $t$ ).  $L(\theta; m): \Theta \times M \rightarrow \mathbb{R}^1$  denotes the likelihood function. The vector  $\theta$  includes all estimated parameters: ideal points, policy locations, and parameters for error term distributions. In the standard ideal point problem

$$L(\theta; m) = \prod_{l \in L} \prod_{t \in T} [F(u(y_t; x_l) - u(n_t; x_l))]^{m_{lt}} [1 - F(u(y_t; x_l) - u(n_t; x_l))]^{1-m_{lt}}, \quad (1)$$

although the framework here is generalized to allow for alternative functional forms and additional parameters. By  $A \subset \Theta$  we denote the set of parameters that satisfy the *maintained hypotheses* (thinking of  $A$  as the admissible set of parameters) and by  $H \subset A$  we denote the set of parameters that also satisfy the *to be tested hypotheses*. If the *maintained hypotheses* consist of a linear system relating certain bill locations to other bill locations (as in Clinton and Meirowitz 2001) then  $A$  is a linear subspace of  $\Theta$ . If the *to be tested hypotheses* consist of a linear system relating certain bill locations (say for example  $y_t = y_{t'}$ ) then  $H$  is a linear subspace of  $A$ . We use the notation  $\dim A$  and  $\dim H$  to denote the dimensionality of these spaces.

In the frequentist maximum likelihood approach, the likelihood ratio test involves estimating  $L_u = \max_{\theta \in A} L(\theta; m)$  and  $L_c = \max_{\theta \in H} L(\theta; m)$ , and calculating the test statistic  $2(\ln L_c - \ln L_u)$ , which is distributed  $\chi^2_{\dim A - \dim H}$ . Other approaches (such as the Wald test) are also appropriate.

In the Bayesian paradigm, the model is augmented to include a prior  $\rho(\theta)$  that has support  $A$  and a prior  $\mu(\theta)$  that has support  $H$ . Finally a prior  $q$  over whether  $\theta \in H$  is assumed. The integrals  $b_u = \int L(\theta; m) d\rho(\theta)$  and  $b_c = \int L(\theta; m) d\mu(\theta)$  are evaluated and the Bayes'

factor

$$B = \frac{qb_c}{(1-q)b_u} \quad (2)$$

represents an assessment of whether the data support the conclusion that the *to be tested hypotheses* are true conditional on the assumption that the *maintained hypotheses* are true.

Aside from testing the *to be tested hypotheses*, the approach also yields estimates of  $\theta$  congruent with the *maintained hypotheses*. If the scholar is interested in calibrating a formal model that satisfies the *to be tested hypotheses*, then the estimated parameters  $\theta_c \in \arg \max_{\theta \in H} L(\theta; m)$  represent reasonable values. Although the hypothesis test outlined above reveals whether the constraints imposed by the *to be tested hypotheses* are reasonable, even if the analyst rejects the *to be tested hypotheses*, then as long as the constrained model makes sense the parameters  $\theta_u \in \arg \max_{\theta \in A} L(\theta; m)$  represent reasonable values. Of course, any aspect of the theory can be tested by defining appropriate *to be tested hypotheses* in an iterative approach. In the Bayesian approach the posteriors are proportional to

$$P_u(\theta) = L(\theta; m)\rho(\theta) \quad (3)$$

$$P_c(\theta) = L(\theta; m)\mu(\theta). \quad (4)$$

The framework presents a way to estimate parameters that may or may not satisfy the *to be tested hypothesis* while satisfying the *maintained hypotheses*. While this approach to hypothesis testing is certainly not new, a casual review of the ideal point estimation literature reveals that very little hypothesis testing uses the actual roll call data. The prevalent approach is to instead treat the recovery of parameters associated with roll call voting as a computation problem (i.e., find reasonable estimates) and then incorporate the estimates into a second-level analysis; hypothesis testing occurs in the second stage and not when the actual roll calls are analyzed. The above demonstrates that the procedure that generates ideal point estimates can also serve as a mechanism for inference.

## 2.2 The Role of Theory

The above statistical exposition summarizes the standard Neyman–Pearson and Bayesian perspectives. While  $m \in M$  (typically roll call matrices) is explicitly called and recognizable as data, a second source of information has been subtly introduced. The sets  $A, H \in \Theta$  also represent data in the statistical procedure, as these sets contain information supplied by the social scientist. Sources of information for *maintained hypotheses* include explicit assumptions of formal models, substantive knowledge about the pieces of legislation or agenda being voted upon, information about agents, and any broader theoretical forecasts that one is willing to take as given. Sources of information for the *to be tested hypotheses* are limited to propositions whose validity one wishes to investigate in the context of the maintained assumptions.

Our view is that explicit legislative theory and careful social scientific reasoning enable the selection of appropriate sets  $A$  and  $H$ . For example, a scholar interested in assessing whether strategic voting occurred when the Powell amendment to the aid-to-education bill was considered during the 84th House might use the theory of strategic voting as well as knowledge of the actual legislative agenda to formulate an agenda tree and uncover sophisticated equivalents. The relationship between sophisticated equivalents would form

the set  $H$  and parameter vectors in which sophisticated equivalents were not equivalent would not be in  $H$ .

In forming the agenda tree, relationships between other votes might be uncovered. If these relationships and votes are separate from the question of interest, it is possible to use this information to impose constraints by specifying that the set  $A$  excludes certain parameter profiles. The hypothesis test of whether  $\theta \in H$  then becomes a conditional test of whether strategic voting occurs *given* that the legislative politics satisfy a set of assumptions that must be made to determine what strategic voting means.

Theoretical models permit us to specify worlds in which it makes sense for hypotheses about legislative behavior to be true. An advantage of this perspective is that hypothesis tests are assured to be internally consistent. If one tests whether interest groups influence legislative voting behavior by regressing NOMINATE scores on interest group contributions then a puzzle might surface. If interest groups do contribute a large amount of resources to a subset of legislators and successfully pressure them on a subset of the votes, then NOMINATE (or any other estimator that assumes i.i.d. shocks) is not an appropriate estimator of legislator preferences. Will such an approach uncover strong effects? Probably. Is this the best approach? Assuredly not, as more nuanced inferences might be permitted if the potential influence of contributions is directly incorporated in the statistical model of legislator preferences.

To illustrate the applicability of this procedure, in the following section we consider the claim that strategic voting produced a “killer amendment” for the aid-to-education bill in the 84th House and the possibility that the location of the capitol and the question of Revolutionary War debt assumption were resolved via a log roll during the 1st House.

### 3 An Application: Strategic Voting

An important question to students of legislative politics is the extent to which legislator voting behavior is strategic (e.g., Enelow and Koehler 1980; Enelow 1981; Riker 1982; Denzau et al. 1985; Mouw and Mackuen 1992; Stratmann 1992; Calvert and Fenno 1994; Poole and Rosenthal 1997; Volden 1998). One often studied example is the “killer amendment” (e.g., Wilkerson 1999; Jenkins and Munger 2003), and scholars have investigated legislative histories in search of evidence. In general, scholars remain divided as to the relative incidence of strategic behavior (e.g., Denzau et al. 1985 argue that there is little evidence for the claim whereas Enelow and Koehler 1980, Stratmann 1992, Calvert and Fenno 1994, and Volden 1998 suggest otherwise). In this section we illustrate how the methodology described in Section 2 can be used to test for strategic voting behavior. We show how scholars interested in strategic voting can impose constraints and estimate models of strategic behavior that are congruent with a posited theoretical model.

To distinguish what is unique about the approach we adopt, it is useful to sketch the existing method of testing for strategic accounts of legislative voting behavior. Most work relies on the methodology outlined in Enelow and Koehler (1980) and Enelow (1981)<sup>5</sup>: (1) write down an agenda tree representation of the legislative agenda; (2) derive all possible preference profiles; (3) derive the voting profile implied by sincere voting for each

<sup>5</sup>Stratmann (1992) presents an alternative approach that uses a simultaneous-equation probit model controlling for constituency characteristics, ideological interests, campaign contributions, and party affiliation to estimate a sequence of amendment votes and then inspects the correlation of the errors to determine the presence of log rolling.

preference profile; (4) assume the collective preference profile over the alternatives (i.e., which alternatives will win in every possible pairwise comparison) and derive the voting profile implied by sophisticated voting; (5) using an external measure of legislator preferences, assign legislators to each of the possible preference profiles; and (6) determine the relative incidence of the sophisticated and sincere voting profiles within each of the preference profiles where the implied voting profiles differ.

Despite the application of this methodology in almost every investigation of strategic voting, some difficulties surface. One problem is that the scholar is forced to attribute a preference profile to every legislator even though this is unobserved. Consequently, summaries of prior voting behavior such as interest group scores or ideal point estimates are frequently employed. As argued above, this assumes that the employed estimates are neutral (or at least consistent) with respect to the hypotheses being tested. Second, the mapping from such voting summary measures into the relevant preference profiles must also be assumed. It is not entirely clear how externally generated measures of legislator preferences relate to the particular preference profiles of interest. Third, voting on the bills of interest is assumed to be deterministic and nonstochastic, which results in the exclusion of legislators whose voting profiles are unexplainable (although this is typically a very small number).

Explicitly embedding a theoretical account (i.e., agenda tree representation) within the roll call estimation procedure provides several benefits. First, we are not required to identify the preference profile held by each legislator from external information. Instead, we estimate measures of induced preferences using the included votes. This automatically accounts for estimation uncertainty and ensures that the preference estimates are compatible with the hypothesized instances of strategic behavior. We are not forced to rely upon measures of legislator preferences (such as interest group scores) that are only tangentially applicable to a specific problem and/or biased for or against the hypotheses being tested.

Second, changing the focus from voting profiles to perceived locations (and potential sophisticated equivalents) changes the focus of analysis from uniquely observed voting profiles to estimates of perceived proposal locations. As proposal location parameters enter into the likelihood function as many times as there are votes cast by legislators on the proposal (i.e., the parameter estimate is a function of legislator voting behavior and the spatial location of the voting legislators), there is reason to suspect that these parameter estimates are less sensitive to error than measures used to quantify the single realization of a legislator's observed voting profile.

Finally, it is possible to assess the extent to which proposals are similar because the procedure allows for hypothesis tests. The "output" of the methodology we present includes information that can be used to assess the nature of the politics on the issue (e.g., how the legislators perceived competing proposals). The formation of future theories may be informed by the analysis. We now demonstrate the approach in two applications.

### 3.1 *The Powell Amendment*

Existing research on strategic voting focuses on precisely the kinds of situations where scholars possess the detailed substantive knowledge necessary to implement the approach we advocate. For example, investigations have focused on reparations to William & Mary College in 1872–1873 (Jenkins and Munger 2003), amnesty for Confederate combatants in 1872 (Jenkins and Munger 2003), the 1956 School Aid Bill (Enelow 1981; Riker 1982; Denzau et al. 1985; Poole and Rosenthal 1997), the 1966 Civil Rights Bill (Enelow



1981), the Common Site Picketing Bill of 1977 (Enelow and Koehler 1980; Poole and Rosenthal 1997), the Panama Canal Treaties in 1977 and 1978 (Enelow and Koehler 1980; Poole and Rosenthal 1997), Senate minimum wage legislation in 1977 (Krehbiel and Rivers 1990), a series of votes on TV coverage of the U.S. Senate in the 1980s (Calvert and Fenno 1994), amendments to the Farm Bill in 1985 (Stratmann 1992), House minimum wage legislation in 1989 (Volden 1998), and examples in Scandinavian parliaments (Bjurulf and Niemi 1978).

As an illustration of the methodology, we analyze the politics surrounding the Powell amendment to the federal aid-to-education bill in the 84th House. The point of our discussion is not to suggest that previous analyses are incorrect. Rather, we use this history to highlight the approach and the ensuing advantages.

The history surrounding the Powell amendment is discussed in detail elsewhere (Enelow 1981; Riker 1982; Brady and Sinclair 1984; Poole and Rosenthal 1997). Brady and Sinclair (1984) summarize the politics surrounding the aid-to-education bills in the late 1950s:

A general aid-to-education bill reached the House floor for the first time in 1956. After the Powell amendment barring racial discrimination was adopted, the bill was defeated 224 to 194. In 1957, the 85th House, by the narrow margin of 208 to 203, voted to strike the enacting clause and thus kill an aid-to-education bill. The Powell amendment had not been added to this bill. During the 86th Congress both House and Senate passed a bill. The House adopted a Powell amendment yet the bill nevertheless passes, though on the close vote of 206 to 189.

The question of interest is whether the voting behavior of identifiable groups of legislators provides evidence that they anticipated that the adoption of the Powell amendment would result in the rejection of the amended aid-to-education bill in 1956. In other words, did some legislators perceive the Powell amendment to be a “killer amendment” when they were voting on it?

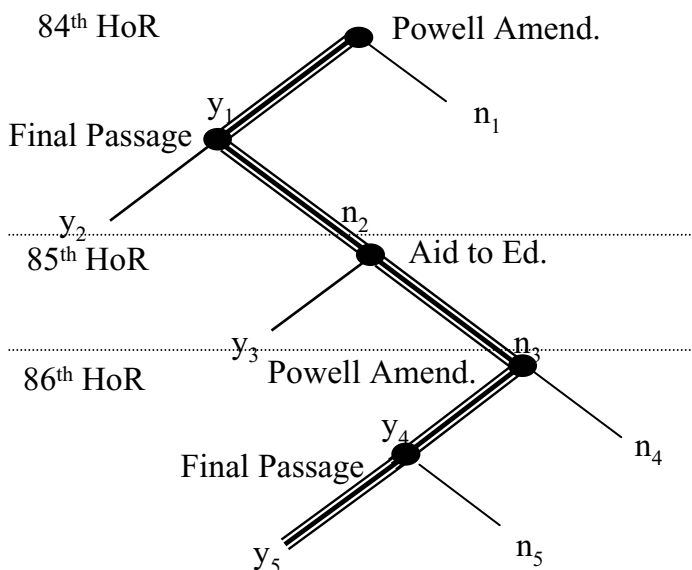
Using the description of Brady and Sinclair, it is possible to produce an agenda tree representation of the relevant voting behavior on the aid-to-education bills in the 84th, 85th, and 86th Houses. Figure 2 presents the agenda tree for the five pertinent roll call votes.

In terms of the framework articulated above, the statistical question is whether the location vector associated with the adoption of the Powell amendment in 1956 (i.e.,  $\mathbf{y}_1$ ) is identical to the location vector associated with the rejection of the amended bill (i.e.,  $\mathbf{n}_2$ ). Note that the location parameters represent vectors in  $\mathbb{R}^2$  to account for the two-dimensional nature of the problem (see Poole and Rosenthal 1991).

The agenda tree of Fig. 2 reveals that there are 10 proposal location parameter vectors that need to be estimated on the basis of five votes (i.e.,  $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4, \mathbf{y}_5, \mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \mathbf{n}_4, \mathbf{n}_5\}$ ). It is possible to impose a sufficient number of identifying constraints to make the estimation of these parameters possible without relying on the parametric form of the utility loss functions.<sup>6</sup>

Several *maintained hypotheses* can be imposed to gain leverage on the problem. First, since three outcomes result in no federal aid-to-education, we assume that the spatial locations associated with these outcomes are equivalent. In particular, rejecting the amended aid-to-education bill in 1956 ( $\mathbf{n}_2$ ), rejecting the aid-to-education bill in 1957 ( $\mathbf{n}_3$ ), and rejecting the amended aid-to-education bill in the 86th House ( $\mathbf{n}_5$ ) are all constrained to be

<sup>6</sup>Strictly speaking, the quadratic functional form in the Bayesian simulation estimator is sufficient to identify the yea and nay location parameters in a unidimensional model (see Clinton et al. 2003). As the policy space of interest is two-dimensional, this point is not relevant here.



**Fig. 2** Relevant agenda for aid-to-education votes (84th–86th Houses).

spatially equivalent (i.e.,  $\mathbf{n}_2 = \mathbf{n}_3 = \mathbf{n}_5$ ). Second, as passing the Powell amendment in the 86th House ( $\mathbf{y}_4$ ) resulted in the passage of the aid-to-education bill on the floor of the 86th House ( $\mathbf{y}_5$ ), the passage of the amended aid-to-education bill in the 86th House is the sophisticated equivalent of voting yea on the Powell amendment (i.e.,  $\mathbf{y}_4 = \mathbf{y}_5$ ). Finally, we assume that the spatial location resulting from the rejection of the Powell amendment in the 84th House is equivalent to the location that results from the rejection of the Powell amendment in the 86th House (i.e.,  $\mathbf{n}_1 = \mathbf{n}_4$ ).

With these constraints, six location parameter vectors need to be estimated using five roll call votes. Fixing the spatial location associated with passing the unamended aid-to-education bill in the 86th House (i.e.,  $\mathbf{y}_3$ ) to be  $(-.5, 0)$  both normalizes the space and reduces the number of unknown location parameters to 5:  $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_5, \mathbf{n}_1, \mathbf{n}_2\}$ . Table 1 presents the (relabelled) parameter estimates to be estimated for each vote.

To test whether legislators perceived the adoption of the Powell amendment in 1956 as resulting in the rejection of the aid-to-education bill when voting on the Powell amendment requires determining whether  $\theta_2 = \theta_3$  in Table 1. This is the *to be tested hypothesis*. To permit the possibility that only some legislators voted strategically (as Enelow 1981 suggests), we allow the perceived spatial location of some outcomes to differ across three groups of legislators: Republican, Southern Democrat, and non-Southern Democrat. The superscripts denote this allowance.

**Table 1** Parameters for Powell amendment

<i>Congress</i>	<i>Vote</i>	<i>Nay parameter</i>	<i>Yea parameter</i>
85th	Powell amendment	$\theta_1$	$\theta_2^{\text{ND}}, \theta_2^{\text{SD}}, \theta_2^{\text{R}}$
85th	Bill and Powell amendment	$\theta_3^{\text{ND}}, \theta_3^{\text{SD}}, \theta_3^{\text{R}}$	$\theta_4$
86th	Bill	$\theta_3^{\text{ND}}, \theta_3^{\text{SD}}, \theta_3^{\text{R}}$	$\theta_5 = (-.5, 0)$
87th	Powell amendment	$\theta_1$	$\theta_6$
87th	Bill and Powell amendment	$\theta_3^{\text{ND}}, \theta_3^{\text{SD}}, \theta_3^{\text{R}}$	$\theta_1$

The intuition of the exercise is straightforward. Assuming that legislators understood the legislative agenda (and formed reasonable beliefs about late-stage actions), was passage of the amendment viewed by legislators voting on the Powell amendment in 1956 equivalent to killing federal aid for education? In other words, given the maintained hypotheses, is it the case that  $\theta_2^{\text{ND}} = \theta_3^{\text{ND}}$ ,  $\theta_2^{\text{SD}} = \theta_3^{\text{SD}}$  and  $\theta_2^{\text{R}} = \theta_3^{\text{R}}$ ?

To test these hypotheses, we estimate two models. Model 1 assumes that the *to be tested hypotheses* are true (i.e.,  $\theta_2^{\text{ND}} = \theta_3^{\text{ND}}$ ,  $\theta_2^{\text{SD}} = \theta_3^{\text{SD}}$  and  $\theta_2^{\text{R}} = \theta_3^{\text{R}}$ ). Model 2 does not impose this constraint. Testing the hypotheses that some or all of the legislator groupings voted strategically is performed by comparing the performance of the two models and comparing the estimated parameter locations.

We estimate the two models using modified versions of the Bayesian simulation estimator described by Clinton et al. (2003).<sup>7</sup> We assume that legislator preferences are fixed in the 84th, 85th, and 86th Houses. Assuming that Rep. Fulton (R, PA) and Rep. Tabler (R, NY) are fixed at the locations from Poole's Optimal Classification (Poole 2000) estimate (and that  $\theta_5 = (-.5, 0)$ ) yields the six parameter constraints required to normalize the two-dimensional space.<sup>8</sup> Five hundred eighty-three ideal point estimates and parameters relating to the 452 nonunanimous votes are estimated. In all, 173,598 votes by legislators are analyzed.

The top graph in Fig. 3 plots the posterior mean of the resulting ideal point estimates for every legislator who voted in the 84th–85th Houses. As we orient the space using Poole's Optimal Classification scores, the distribution of resulting estimates are quite similar to existing estimates. The vertical dimension appears to capture issues on civil rights, and the horizontal dimension corresponds to issues pertaining to the traditional left–right ideological spectrum. The recovered estimates reveal the well-known Democratic fissure on issues of civil rights.

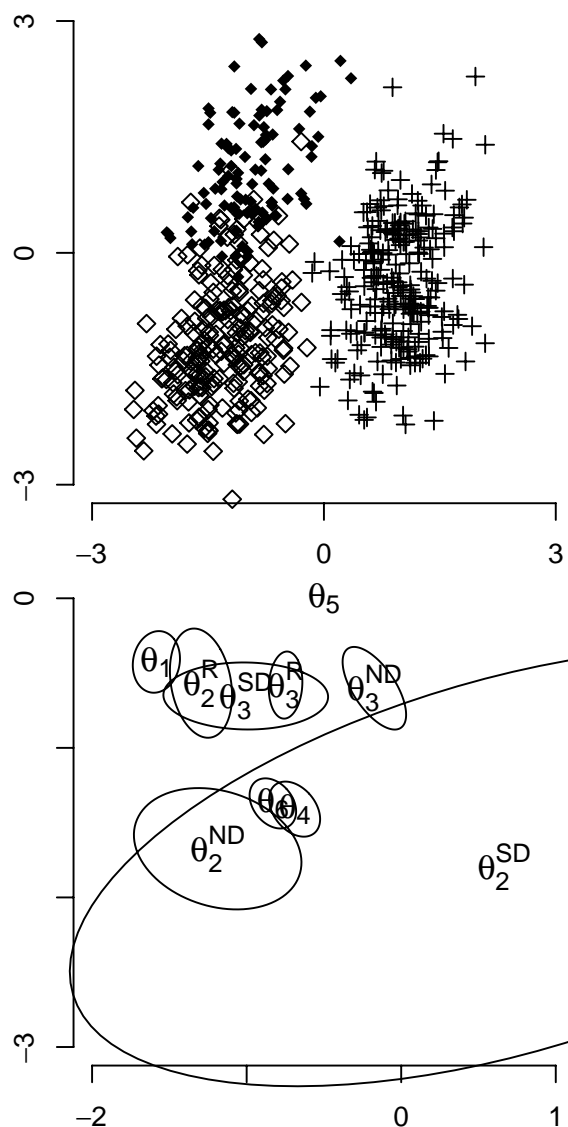
The bottom graph in Fig. 3 plots posterior means and draws from the posteriors of the bill parameters of interest from the model that does not impose the *to be tested hypotheses*. The approach is able to recover estimates of the perceived policy locations, as the *maintained hypotheses* are sufficient to identify the location parameters without having to rely on parametric assumptions. Previous work could only recover the former, and even those were of questionable neutrality with respect to the hypotheses of interest.

Several conclusions are notable in terms of the location parameter estimates. For both the Northern and Southern Democrats  $\theta_2$  and  $\theta_3$  are quite distinct—the spatial location associated with the passage of the Powell amendment is quite dissimilar to that resulting from the rejection of the amended aid-to-education bill in the 84th House. The posterior distributions of these two parameters for Republicans also have no noticeable overlap, although the civil rights coordinates of  $\theta_2^{\text{R}}$  and  $\theta_3^{\text{R}}$  seem to coincide.

Consequently, there is no evidence from these estimates that any of the groups viewed passage of the Powell amendment in the 84th House as equivalent to passage of the aid-to-education bill with the Powell amendment attached. Consistent with the argument advanced by Denzau et al. (1985), there is little indication that a combination of strategic voting

<sup>7</sup> A number of modifications to the basic model in Clinton et al. (2003) are employed. First, every vote except for the five described in the agenda tree are estimated using the linearized version of the estimator. Imposing the constraints of the *maintained hypotheses* prevents the use of the linearized version for the five votes of interest. Consequently, the nonreduced likelihood is used to estimate those votes. Second, for computational reasons, we assume that the error differences are logistic, not standard normal.

<sup>8</sup> Fixing Rep. Fulton at (.09, -1) and Rep. Tabler at (.68, -.3) ensures that the recovered space will be similar to those of existing estimates—positive (negative) estimates will be associated with “conservative” (“liberal”) positions. Note that fixing any six parameters is sufficient to normalize the space (Rivers 2003).



**Fig. 3** Parameter estimates for Powell amendment example. The top graph depicts the posterior means of legislator ideal point estimates. Open diamonds represent non-Southern Democrats, solid diamonds represent Southern Democrats, and crosses represent Republicans. The bottom graph denotes the posterior mean location for the parameters in Table 1. The estimated normal 95% “confidence regions” are indicated for each posterior. The  $x$ -axis represents the “liberal conservative” dimension and the  $y$ -axis represents the “civil rights” dimension.

and sincere voting combined to produce a “killer amendment” story.<sup>9</sup> However, of the three groups, and consistent with Enelow’s suggestion, the Republican estimates are most agreeable to the constraints required by sophisticated equivalence.

<sup>9</sup>This conclusion is tempered by the lack of precision in the estimates of  $\theta_2^D$  and  $\theta_2^{SD}$ . This may be due to heterogeneity in the perceptions of legislators within these groups.

Calculation of Bayes's factors for the two models (with and without the *to be tested hypotheses* serving as constraints) provides another means to test this version of the "killer amendment" story. With uniform priors  $\rho(\theta)$  and  $\mu(\theta)$ , integrating  $P_u(\theta)$  with respect to the prior is equivalent to  $\sum_t L(\theta_t; m)/T$ , where  $\theta_t$  represents the parameter values from draw  $t$  from the posterior and  $T$  is the number of sample draws from the estimated posterior.

Given the number of parameters being estimated, we are unable to calculate the required values (as  $\sum_t L(\theta_t; m)$  is vanishingly small). Instead, we observe that the maximum log-likelihood value for the unconstrained model is  $-99991.02$ , and the maximum log-likelihood value for the constrained model is  $-100113.9$  (over 50 draws from the posterior). Calculating the change in the Bayesian Information Criterion for these models yields a value of 207.57—strongly suggesting that the *to be tested hypotheses* are not satisfied because the model that does not impose the constraints fits the observed voting data better.<sup>10</sup> This result confirms the intuition from inspecting the proposal parameter estimates in Fig. 3. Consistent with the claims of Denzau et al. (1985), analyzing the observable roll call voting record provides no strong evidence that when legislators voted on the Powell amendment to the aid-to-education bill in the 84th House they perceived that the passage of the amendment would result in the eventual failure of the amended aid-to-education bill.

### 3.2 The Compromise of 1790

The approach can also be applied to test one of the earliest accounts of strategic voting in American history. The traditional story of the Compromise of 1790 involves a log roll over federal assumption of the states' Revolutionary War debts and the location of the temporary and permanent seats of government. The traditional account stems largely from three letters left by Jefferson detailing his involvement in a dinner party involving Hamilton and Madison in which a deal was struck.

It was observed, I forget by which of them, that as the pill [assumption of the state debts] would be a bitter one to the Southern states, something should be done to soothe them; and the removal of the seat of government to the Potomac was a just measure, and would probably be a popular one with them, and would be a proper one to follow the assumption.

(Thomas Jefferson in 1792 summarizing the outcome of the dinner party he held in mid-June 1790, quoted in Ellis 2000)

Although there is no dispute that a meeting took place between the principals at Jefferson's residence in mid-June, historians examining primary source material are divided over whether the Compromise was ever consummated. Clinton and Meirowitz (2003) reexamine roll call voting over these two issues and investigate how the legislators perceived the various voting options. In the study, the agenda (specifically the relationships between certain votes) and substantive knowledge of the issues being considered yield *maintained hypotheses* that dramatically reduce the number of bill parameters to be estimated. *To be tested hypotheses* involve how legislators perceived the critical votes on assumption and the capital location cast in the House during the summer of 1790.

The first vote of relevance to the compromise was on passage of S.12 on July 9, 1790, which located the temporary capital in Philadelphia and the permanent capital on the

<sup>10</sup>The change in the Bayesian Information Criterion (which denotes the change in moving from the constrained to unconstrained model) is given by  $-2 \log(\sup_c L(\theta; m)/\sup_{uc} L(\theta; m)) - (p_{uc} - p_c) \log(n)$ , where  $\sup_{uc} L(\theta; m)$  and  $\sup_c L(\theta; m)$  denote the supremum over the support of the prior of the likelihood for the unconstrained and constrained models respectively,  $p_{uc}$  and  $p_c$  denote the number of parameters in the unconstrained and constrained models respectively, and  $n$  denotes the number of observations (i.e., votes).

Potomac. The second piece of legislation involved with the log roll was passage of the Funding Bill on July 19 which did not provide for assumption. An amendment to the Funding Bill provided for assumption and was passed on July 29. The final piece of legislation involved with the log roll was an amendment that reduced the rate of interest paid on debt interest to state debt creditors passed on July 29.

If legislators believed that the Jefferson–Madison–Hamilton compromise was reached, then they (like Jefferson, Madison, and Hamilton) would be able to predict the consequences of each vote. Specifically, if the traditional log roll account is true then the sophisticated equivalent of voting for a permanent capital on the Potomac and a temporary capital in Philadelphia (S.12) is a world where the temporary capital is in Philadelphia, the permanent capital is on the Potomac, the Funding Bill passes, and is ultimately amended to provide for assumption with a reduced interest rate. While attributing an ability to predict outcomes might seem strong, the logic of the Compromise requires that legislators see the first vote as the first step in reaching the final agreed-upon outcome (i.e., the temporary capital in Philadelphia, the permanent capital on the Potomac, and assumption passed at a low interest rate).

Note that these relationships are an explicit statement of the traditional account of the Compromise, not an auxiliary assumption of the employed estimation procedure. More specifically, if the traditional account of the log roll is correct, Hamilton and Madison would have already brokered the appropriate deals by the time of the vote on S.12. Accordingly, it would have been known by the legislators that once they initiated the process of passing S.12, a chain of events would ensue causing the final outcome to be assumption at a low rate of payment. In other words, the *to be tested hypothesis* is that the yeas of each of these votes are identical:  $y_1 = y_2 = y_3 = y_4$ .

Posterior estimates from Bayesian simulation methods revealed that the *to be tested hypothesis* was unsupported. Integrating theory and measurement in the analysis of legislative voting behavior offered leverage on a historical debate that remains unresolved despite the presence of primary source material. Clinton and Meirowitz (2003) interpret the estimates as supportive of an alternative story.

Contrary to the conventional story when voting on the capital bill legislators did not anticipate that passage of this legislation would also entail the assumption of state Revolutionary War debts at the final agreed upon interest rate. When voting on the funding bill, legislators did not anticipate that subsequent amendments involving assumption would pass. The traditional account of the Compromise is not well supported by the roll call data and the theory of spatial voting. Instead, the questions of residence and assumption seem to have been resolved independently in the summer of 1790, with a compromise between assumption and reduced interest payments settling the contentious funding question.

## 4 Discussion

This article details a framework for integrating theory and estimation through the careful use of constraints in the context of roll call analysis. For a given theoretically derived hypothesis, the scholar constructs *to be tested hypotheses* and uses auxiliary assumptions of the model as well as other data and knowledge about the legislative history to form *maintained hypotheses*. The latter serve as constraints on the parameter space and the former serve to define a hypothesis test.

We contend that this approach can be fruitfully applied to real legislatures, and our discussion of the politics surrounding the Powell amendment and the politics of the First Congress illustrate how to impose and test constraints when scaling roll call voting data. So doing leverages off the structure of the posited estimation model to permit more nuanced

investigations into the politics (e.g., the spatial location of the proposals being voted upon and their relationship to legislator-induced preferences) and generates estimates that are neutral with respect to the *to be tested hypotheses*. However, there are costs and limits to the approach. Generating *maintained hypotheses* relating to the agenda may involve exhaustive reading of legislative histories to determine what the agenda tree actually looks like. Imposing structure based on the substantive import of the policy actions might involve subjective assumptions/interpretations that are open to debate. Deciding what *maintained hypotheses* follow from the auxiliary assumptions of strategic theories may require the construction of very explicit models. But in the end, increased knowledge and structured thought about legislative histories, efforts to incorporate legislative content into estimation, and careful construction of theories are all reasonable ways to enhance our understanding of legislative behavior. The construction of estimation procedures that are detailed to fit a specific question (or set of related questions) while more closely bridging theory and data seems like a promising and important direction for future methodological advancement in the field of roll call analysis.

## References

- Austen-Smith, David. 1987. "Sophisticated Sincerity: Voting over Endogenous Agendas." *American Political Science Review* 81:1321–1329.
- Banks, Jeffrey, and John Duggan. 2000. "A Bargaining Model of Collective Choice." *American Political Science Review* 94:73–88.
- Banks, Jeffrey, and F. Gasmi. 1987. "Endogenous Agenda Formation in Three-Person Committees." *Social Choice and Welfare* 4:133–152.
- Bjurulf, Bo, and Richard Niemi. 1978. "Strategic Voting in Scandinavian Parliaments." *Scandinavian Political Studies* 1:5–22.
- Brady, David, and Barbara Sinclair. 1984. "Building Majorities for Policy Changes in the House of Representatives." *Journal of Politics* 46:1033–1060.
- Calvert, Randall, and Richard Fenno. 1994. "Strategy and Sophisticated Voting in the Senate." *Journal of Politics* 56:349–376.
- Clinton, D. Joshua, Simon Jackman, and Douglas Rivers. 2003. "Statistical Analysis of Roll Call Data: A Unified Approach." Typescript. Stanford, CA: Stanford University.
- Clinton, D. Joshua, and Adam Meirowitz. 2001. "Agenda Constrained Legislator Ideal Points and the Spatial Voting Model." *Political Analysis* 9:242–259.
- Clinton, D. Joshua, and Adam Meirowitz. 2003. "Integrating Voting Theory and Roll Call Analysis: A Re-Examination of the Compromise of 1790." Typescript. Princeton, NJ: Princeton University.
- Denzau, Arthur, William Riker, and Kenneth Shepsle. 1985. "Farquharson and Fenno: Sophisticated Voting and Home Style." *American Political Science Review* 79:117–134.
- Ellis, J. 2000. *Founding Brothers: The Revolutionary Generation*. New York: Alfred A. Knopf.
- Enelow, James. 1981. "Saving Amendments, Killer Amendments, and an Expected Utility Theory of Sophisticated Voting." *Journal of Politics* 43:1062–1089.
- Enelow, James, and Melvin Hinich. 1984. *The Spatial Theory of Voting: An Introduction*. New York: Cambridge University Press.
- Enelow, James, and David Koehler. 1980. "The Amendment in Legislative Strategy: Sophisticated Voting in the U.S. Congress." *Journal of Politics* 42:396–413.
- Groseclose, Timothy, and James Snyder. 1996. "Buying Supermajorities." *American Political Science Review* 90:303–315.
- Guarnaschelli, S., Richard McKelvey, and Thomas Palfrey. 2000. "An Experimental Study of Jury Decision Rules." *American Political Science Review* 94:407–423.
- Herron, Michael. 1999. "Artificial Extremism in Interest Group Ratings and the Preference versus Party Debate." *Legislative Studies Quarterly* 24:525–542.
- Jenkins, Jeffrey, and Michael Munger. 2003. "Investigating the Incidence of Killer Amendments in Congress." *Journal of Politics* 65:498–517.
- Krehbiel, Keith, and Adam Meirowitz. 2002. "Minority Rights and Majority Power: Theoretical Consequences of the Motion to Recommit." *Legislative Studies Quarterly* 27:191–218.

- Londregan, John. 1999. "Estimating Legislators' Preferred Points." *Political Analysis* 8:35–57.
- McCarty, Nolan, Keith Poole, and Howard Rosenthal. 2001. "The Hunt for Party Discipline in Congress." *American Political Science Review* 95:673–688.
- McKelvey, Richard. 1976. "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control." *Journal of Economic Theory* 12:472–482.
- McKelvey, Richard, and Richard Niemi. 1978. "A Multistage Game Representative of Sophisticated Voting for Binary Agendas." *Journal of Economic Theory* 18:1–22.
- McKelvey, Richard, and Thomas Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior* 10:6–38.
- McKelvey, Richard, and Thomas Palfrey. 1996. "A Statistical Theory of Equilibrium in Games." *Japanese Economic Review* 47:186–209.
- McKelvey, Richard, and Thomas Palfrey. 1998. "Quantal Response Equilibria for Extensive Form Games." *Experimental Economics* 1:9–41.
- McKelvey, Richard, and Norman Schofield. 1987. "Generalized Symmetry Conditions at a Core Point." *Econometrica* 55:923–934.
- Mouw, Calvin, and Michael Mackuen. 1992. "The Strategic Agenda in Legislative Politics." *American Political Science Review* 86:87–105.
- Plott, Charles. 1967. "A Notion of Equilibrium and Its Possibility Under Majority Rules." *American Economic Review* 57:787–806.
- Poole, Keith T. 2000. "Nonparametric Unfolding of Binary Choice Data." *Political Analysis* 8:211–237.
- Poole, Keith, and Howard Rosenthal. 1991. "Patterns of Congressional Voting." *American Journal of Political Science* 35:228–278.
- Poole, Keith, and Howard Rosenthal. 1997. *Congress: A Political Economic History of Roll Call Voting*. New York: Oxford University Press.
- Riker, William. 1982. *Liberalism Against Populism*. Prospect Heights, IL: Waveland Press.
- Rivers, Douglas. 2003. "Identification of Multidimensional Item-Response Models." Typescript. Stanford: Stanford University.
- Signorino, Curt. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93:279–298.
- Snyder, James. 1991. "On Buying Legislatures." *Economics and Politics* 3:93–109.
- Snyder, James. 1992. "Artificial Extremism in Interest Group Ratings." *Legislative Studies Quarterly* 17:319–345.
- Snyder, James, and Timothy Groseclose. 2000. "Estimating Party Influence in Roll-Call Voting." *American Journal of Political Science* 44:193–211.
- Stratmann, Thomas. 1992. "The Effects of Logrolling on Congressional Voting." *American Economic Review* 85:1162–1176.
- Volden, Craig. 1998. "Sophisticated Voting in Supermajoritarian Settings." *Journal of Politics* 60:149–173.
- Wilkerson, J. 1999. "'Killer' Amendments in Congress." *American Political Science Review* 93:535–542.