Political Analysis, 9:3

Agenda Constrained Legislator Ideal Points and the Spatial Voting Model

Joshua D. Clinton
Department of Political Science,
Stanford University, Stanford, CA 94305
e-mail: jclinton@stanford.edu

Adam Meirowitz
Graduate School of Business,
Stanford University, Stanford, CA 94305
e-mail: ameirow@stanford.edu
and
Department of Politics,
Princeton University, Princeton, NJ 08544

Existing preference estimation procedures do not incorporate the full structure of the spatial model of voting, as they fail to use the sequential nature of the agenda. In the maximum likelihood framework, the consequences of this omission may be far-reaching. First, information useful for the identification of the model is neglected. Specifically, information that identifies the proposal locations is ignored. Second, the dimensionality of the policy space may be incorrectly estimated. Third, preference and proposal location estimates are incorrect and difficult to interpret in terms of the spatial model. We also show that the Bayesian simulation approach to ideal point estimation (Clinton et al. 2000; Jackman 2000) may be improved through the use of information about the legislative agenda. This point is illustrated by comparing several preference estimators of the first U.S. House (1789–1791).

1 Introduction

THE TASK OF estimating legislative preferences from a sequence of binary votes makes strong demands on coarse data (Londregan 2000a). Although several political scientists have successfully devised widely accepted methods of dealing with the problem (e.g., Poole and Rosenthal 1996; Heckman and Snyder 1997), most recent work in the area has involved refining the computation of these models (e.g., Poole 2000). Although this research is certainly important, in this paper we follow Londregan (2000b) in arguing that

Authors' note: The authors would like to thank John Ferejohn, Tim Groseclose, Simon Jackman, Keith Krehbiel, John Londregan, Nolan McCarty, Doug Rivers, Howard Rosenthal, and several anonymous referees for their helpful comments. Josh would like to thank the National Science Foundation Graduate Research Fellowship for support.

there is additional *theoretical* information about the process that generates roll call voting that can be usefully employed by both methodologists interested in estimation issues and Congressional scholars.

In the spatial model of voting (SMV) legislators' voting decisions are determined by their calculation of the utility differential between the state of the world where the proposal passes and the state where the proposal fails. Furthermore, the state that results from the proposal's failure is predetermined by the agenda. This induces a dependence between "nay" and "yea" locations in the issue space. However, existing preference estimation techniques assume that the nay locations are unrelated to the legislative agenda and therefore free parameters to be estimated.

That standard approaches to preference estimation treat all recovered parameters as unrelated to the legislative agenda is best observed by noting that the estimates are unaffected by a reordering of the sequence of votes. The estimates produced from estimating the actual Congressional agenda are identical to estimates produced by an arbitrary reordering of the voting sequence. In this paper, we contend that since the agenda reveals that the status quo point is related to the location of the last proposal to have passed, *the agenda should constrain nay location estimates*.

To demonstrate the intuition, consider an example from the first U.S. Congress (1789–1791). One of the most important issues of the second session of the first Congress was the determination of the funded rate of interest that the assumption of Revolutionary War debt would pay (Cooke 1970). Alternatives of 3, 4, and 6% were debated and voted upon during the Congress. This was a divisive issue, as some states had paid off most of their debts (e.g., Connecticut), while others had hardly paid off any (e.g., Massachusetts). Consequently, representatives from states without large debts (e.g,. Senator Oliver Ellsworth of Connecticut) generally preferred lower interest rates.

For the purposes of example, assume that the default or status quo interest rate was set by the Confederation Congress's pledge of 6%. Consider the agenda of first voting on a proposal to set the interest rate at 3% and then voting on a proposal to set the interest rate at 4%. If the 3% proposal passes, then a vote for the 4% proposal is a vote to *raise* the interest rate. In contrast, if the 3% proposal fails, then a vote for the 4% proposal is a vote to *lower* the interest rate. Although we would certainly expect Ellsworth to support the 3% proposal, our prediction about whether he would support the 4% proposal hinges on knowing whether the 3% proposal passes.[1] In other words, we need to know what the alternative to the 4% proposal is. This example illustrates that proposal "nay" locations are determined by the passage of previous proposals in the agenda. Ignoring this relationship neglects important and useful information.

As a first step toward the incorporation of information about the agenda in the estimation of legislator ideal points, in this paper we:

- Illustrate that current estimators do not incorporate all of the SMV's assumptions regarding the agenda.
- Prove analytically that this omission results in incorrect estimates of voter ideal points.
- Prove that the extent to which existing estimators are incorrect does not vanish as the number of observations tends to infinity. Thus, adding more observations (either proposals, legislators or both) does not solve the problem.

[1] Given that empirical estimation models must assume sincere voting to identify the problem, to be consistent with existing estimation techniques, we assume sincere voting in all the examples.

- Prove that if the analyst is sufficiently concerned with not *underestimating* the dimensionality of the policy space, maximum likelihood (ML) estimates of the dimensionality of the policy space from the unconstrained model will be no greater than those from the constrained ML model. Furthermore, with positive probability, estimates from the unconstrained ML model will be *strictly less* than those from the constrained ML model.
- Compare constrained and unconstrained Bayesian simulation estimators and DW-NOMINATE estimates for the first U.S. House using a simple procedure for incorporating constraints on the nay locations.

The results contained in this paper are important for three reasons. First, scholars can incorporate information about the agenda into their empirical work to impose more structure on their estimation problem. In the context we consider, incorporation of the constraint identifies proposal locations—providing researchers with potentially valuable information. Second, given the incongruence between the behavioral assumptions of the SMV and standard estimators, it is difficult to identify exactly what it is that standard estimators measure. As a consequence, tests of legislative theories imbedded in the SMV using these estimates may be problematic. Third, by failing to account for the structure imposed on the status quo by the SMV, estimates of the dimensionality of the policy space may be incorrect. This last point is important when one considers the fact that the choice of dimensionality in formal models of legislative behavior is not innocuous.[2]

As much can be done to advance preference estimation, we want to be clear about the current paper's scope. We do not:

- Address endogenous agenda formation or sophisticated voting.
- Address the consistency of preference estimators—although it is not possible that both the constrained and the unconstrained estimators are consistent.
- Exhaustively consider ways to constrain the nay locations for a given agenda structure. Instead, we present one example of a constraint.

The remainder of the paper is organized as follows. In Section 2 we consider simple examples illustrating how ignoring the relationship between proposals and status quos can result in erroneous preference estimates and misidentification of the policy space's dimensionality. Section 3 considers a large class of ML estimators and proves several general results about the effects of ignoring the constraint in ML estimation. Section 4 presents and implements a constrained Bayesian simulation estimator for the first U.S. House. Estimates are then compared to the unconstrained estimates and DW-NOMINATE. Section 5 concludes and discusses the next steps in developing this approach. Finally, the Appendix presents explicit definitions and proofs of the results in Section 3.

2 The Status Quo in the Spatial Model

The SMV, which is heavily relied upon in the derivation of theories about legislative behavior, consists of a policy space, legislator preferences defined over the policy space, and an agenda specifying the sequence of proposals to be voted on. In the SMV all payoff-relevant information about the world following the implementation of a proposal is incorporated into the proposal's location in the policy space.

[2]For example, Miller (1993) rationalizes the use of a unidimensional model of legislative and presidential interaction with the fact that NOMINATE recovers a unidimensional policy space.

In other words, a proposal's location in the policy space includes not only the values of dimensions changed by the proposal, but also the values of dimensions that are not changed. For example, a proposal that increases only the minimum wage must also have a defense spending component, a social security funding component, a congressional salary component, etc., *even though the actual proposal never mentions these issues*. To see why this is true, consider a three-dimensional space. Any object in that space is defined by its position in all three dimensions (i.e., a point in the space is of the form $x = \{a, b, c\}$). Hence, if there are three policy dimensions, the location of every proposal must be defined in each of the three policy dimensions, even if it explicitly deals with only one of the dimensions (in which case the proposal's value in the other policy dimensions corresponds to that of the status quo). While this observation is not novel, it is an important and subtle component of the SMV that needs to be taken seriously when interpreting ideal point estimates and testing theories of voting.

The standard legislator preference assumptions, which we also adopt, are that preferences are sufficiently well behaved so that each legislator has a most preferred point in the space (i.e., ideal point) and that utility is decreasing in Euclidean distance from that ideal point. We also make the usual assumption that we know the functional form of legislator utility functions up to an ideal point.

Restated within the framework of the SMV, roll call data are generated by rational voting defined with respect to three assumptions: (i) preference-relevant proposal information may be represented by considering policies as points in a subset of Euclidean space, (ii) legislators vote for a proposal if it is closer to their ideal point than the status quo is (i.e., voting is sincere), and (iii) the identity of a policy's status quo is that of the last successful proposal.

The implications of these three assumptions are illustrated in a second two-dimensional example. In this example, there are two dimensions: a Guns dimension and a Roses dimension. A location in the policy space is therefore described by the pair (G, R), where G represents the value in the Guns dimension, and R represents the value in the Roses dimension. Suppose that the initial status quo is (0, 0) and there are two proposals under consideration: one that increases Guns by only 1 unit and one that increases Roses by only 1 unit. Consider the following two scenarios.

First, if the Guns proposal is voted on first, and both proposals are successful, then the following path through the policy space results: (0, 0) to (1, 0) to (1, 1). In other words, in the first vote, the Guns proposal (1, 0) is considered against the status quo (0, 0). When the Roses proposal is voted upon, the choice is between (1, 1) and (1, 0) because the passage of the Guns proposal moved the status quo from (0, 0) to (1, 0).

Alternatively, if the Guns proposal is voted on first and only the Roses proposal is successful, the path is (0, 0) to (0, 1). In the first vote, the Guns proposal (1, 0) is considered against the initial status quo (0, 0). Since the Guns proposal fails, nothing changes. Thus, when the Roses proposal (0, 1) is considered, it too is considered against the initial status quo (0, 0).

This example illustrates three points. First, proposals have a policy component in every dimension, even in dimensions they do not explicitly affect. Second, there is a relationship between nay locations and previously passed proposals. Third, the agenda matters, as an arbitrary reordering of the agenda affects proposals' nay locations.

Failure to constrain the status quo also affects our ability to recover the correct dimensionality of the policy space. To see this, consider an example in which the legislators (and policies) are intentionally chosen to form a basis for $\mathbb{R}^2$. Figure 1 presents the "true" locations of legislators, proposals, and status quos.

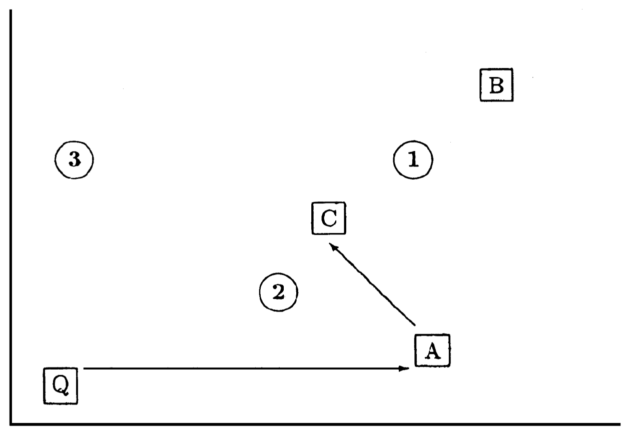

Fig. 1 Example true policy space. The true (by construction) locations of the ideal points of legislators {1, 2, 3} and proposal locations {A, B, C} are given. For the voting agenda {A, B, C}, the policy trajectory through the space is denoted by the arrows.

Au: (1) First & (2) last "sentences" in legend to Fig. 1 as meant?

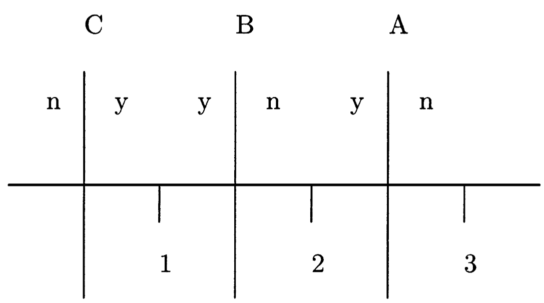

Fig. 2 Possible unidimensional representation of Fig. 1. This is one possible unidimensional representation of Fig. 1 consistent with the voting behavior. As before, {1, 2, 3} denotes the locations of legislator ideal points. {A, B, C} denotes possible cutpoints, with {y, n} indicating the direction of the vote imposed by the cutpoint.

The observed roll call votes of legislators {1, 2, 3} on proposals {A, B, C} consist of legislator 1 voting yes on all three proposals, legislator 2 voting no only on B, and legislator 3 voting yes only on C.

Figure 2 illustrates that with an unconstrained model, it is possible to recover cutpoints and ideal points that perfectly predict the votes using a single dimension despite the fact that the true dimensionality is two. This mistake is not possible under the constrained model.[3]

This section suggests that failing to constrain the status quo in a manner consistent with the SMV may produce ideal point and proposal location estimates that are ambiguously related to their true locations. Additionally, the analyst may be unable to recover the correct

[3]The unconstrained case recovers a unidimensional policy space as long as votes have "connected coalitions" (i.e., 1 and 2 vote together or 2 and 3 vote together). This guarantees that a cutpoint exists and allows for the selection of infinitely many appropriate yea and nay locations. The constrained model's inability to fit this data in one dimension results from the fact that although the votes on A and B imply that the yea location of proposal A is in the Pareto set (i.e., between the ideal points of either 1 and 2 or 2 and 3), the observed voting also indicates that all voters prefer to move to C from A, which is impossible if the yea location of A is in the Pareto set.

dimensionality of the space. As one may question the extent to which these examples are pathological, we now state and prove related results in some generality.

3 Agenda Constrained Maximum Likelihood Estimation

In this section, we consider the properties of two ML estimators—one that constrains nay locations in the manner discussed above and one that treats nay locations as free parameters to be estimated. The interpretation of these two estimators is that the constrained one is the correct ML estimator given the data generation process that results from the SMV with a particular type of noise. The unconstrained estimator is not the correct ML estimator. The results of this section demonstrate the consequence of using the wrong estimator.

It is important to note that the notion of constrained estimators in this paper differs from standard hypothesis testing settings. In this paper we argue that the constraint on nay locations must be imposed because the constraint is assumed in all of the theories that are to be tested or calibrated. Thus, the validity of the constraint is theoretically determined. Consequently, it is unclear what a test of the validity of the constraint (e.g., with likelihood ratio tests) implies because estimates that do not satisfy the constraint have no meaning in terms of the SMV.

Put differently, to test the constraint, the world where the constraint does not hold must make sense (i.e., be characterized by a probability model such that one can interpret its parameters). Interpretation of estimates of bill locations and ideal points requires a model of choice like the SMV, but the world where each bill has a free nay location is *not* one modeled by the SMV. Thus, if the constrained estimates perform worse than the unconstrained estimates (note that by construction they cannot perform better), then the conclusion is *not* that the SMV is true and it is not desirable to incorporate the constraint. Rather, the conclusion is simply that it is possible to write out a likelihood function with more parameters that attains a higher value at its optimum. Unfortunately, this better likelihood function is only ambiguously related to the theories that we are interested in testing and calibrating.

With this observation, we first prove that the constraint (usually) binds. As a consequence, the unconstrained estimator is not a feasible solution to the constrained problem. We then show that given an underlying data generating process satisfying weak conditions, if the constraint binds, the unconstrained ML ideal point estimates are *not* equivalent to ideal point estimates from the constrained problem. A corollary is that if a (strong or weak) law of large numbers holds for the constrained estimator, then one will generically *not* hold for the unconstrained estimator. These two results should be interpreted as arguments why it is not reasonable to pretend that the nay locations are unrelated to the yea locations.

We conclude the section by demonstrating that if one estimates the dimensionality of the policy space through a procedure of iteratively using likelihood ratio tests, then for sufficient levels of concern about concluding that the policy space is smaller than it is (i.e., the probability β of making a type II error in the Neyman–Pearson framework), the constrained model's results will be higher than those of the unconstrained model. Although this last result is weak in the sense of not holding for arbitrary levels of concern about type II errors, the intuition behind the proof clearly demonstrates the potential of unconstrained estimators to misestimate the dimensionality.

3.1 *Notation and Definition*

We begin by presenting a general model of binary voting which yields a general likelihood function. Let $X \subset \mathbb{R}^d$ denote the d-dimensional policy space. A generic element is expressed in both vector and scalar notation as either x or $(x^1, x^2, \ldots, x^j, \ldots, x^d)$. We assume that each legislator in the set of legislators $L = \{1, 2, \ldots, i, \ldots, L\}$ has preferences over proposals that are representable by a utility function with parameter $x_i \in X$. It is customary to think of these parameters as ideal points, and we denote the vector of legislator ideal points $\mathbf{x} = \{x_1, \ldots, x_L\}$. Thus, superscripts index dimensions, subscripts index legislators and proposals, and a boldfaced element denotes a collection of vectors.

A legislator's vote on a proposal is determined by comparing the utility that the legislator attains from voting for and against the proposal. Associated with each proposal is a location in the policy space representing the outcome induced by the proposal's passage. A list of T sequential proposals is an agenda $\mathbf{a} = (y_1, \ldots, y_t, \ldots, y_T)$, where $y_t \in X$ denotes the location in the policy space to which the status quo moves if the tth proposal passes. Thus, the vote on proposal y_1 occurs before y_2, which occurs before y_3, and so on, until y_T. A vote by legislator i for (against) proposal t is denoted $v_{it} = 1(0)$. A voting history of size T, denoted $\mathbf{h} = (v_{11}, \ldots, v_{L1}, v_{12}, \ldots, v_{L2}, \ldots, v_{LT}) \in \{0, 1\}^{LT}$, is a list of binary votes cast by the L legislators for each of the T proposals. The group decision on proposal t is denoted v_t, with $v_t = 1$ if the proposal passes and 0 otherwise.

We begin with the following **agenda assumption**: the status quo for the period t vote is the last proposal that passed. Thus, $q_t = y_{m(t)}$ for $t = 1, 2, \ldots, T$, where $m(t) := \max\{j : j < t \;\&\; v_j = 1\}$. Note that as long as the mapping $m(t)$ is a function $m : \{1, 2, \ldots, T\} \to \{1, 2, \ldots, T\}$, the results presented here attain.

The probability that legislator i votes yea or nay on vote t is given by

$$\begin{aligned} \Pr(v_{it} = 1) &= \rho(x_i, y_t, q_t) \\ \Pr(v_{it} = 0) &= 1 - \rho(x_i, y_t, q_t) \end{aligned} \tag{1}$$

with $\rho : X \times X \times X \to [0, 1]$. We assume that this mapping is smooth and not too flat (explicitly defined in the Appendix). Under this specification and the assumption that the voting lotteries are independent across both indices, the log-likelihood function is of the form

$$\log \pounds(\mathbf{a}, \mathbf{x}, \mathbf{q} \,|\, \mathbf{h}) = \sum_{i=1}^{L} \sum_{t=1}^{T} [v_{it} \log \rho(x_i, y_t, q_t) + (1 - v_{it}) \log(1 - \rho(x_i, y_t, q_t))]. \tag{2}$$

If the agenda assumption is ignored, the unconstrained ML estimator is

$$(\mathbf{a}^u, \mathbf{x}^u, \mathbf{q}^u) \in \arg \max_{\mathbf{a}, \mathbf{x}, \mathbf{q} \in X^{L+2T}} \log \pounds(\mathbf{a}, \mathbf{x}, \mathbf{q} \,|\, \mathbf{h}). \tag{3}$$

However, if the assumption is incorporated, the constrained problem becomes

$$(\mathbf{a}^c, \mathbf{x}^c, \mathbf{q}^c, \lambda) \in \arg \max_{\mathbf{a}, \mathbf{x}, \mathbf{q}, \lambda \in X^{L+2T} \times \mathbb{R}_+^{dT}} \log \pounds(\mathbf{a}, \mathbf{x}, \mathbf{q} \,|\, \mathbf{h}) + \sum_{t=1}^{T} \lambda_t'(q_t - y_{m(t)}) \tag{4}$$

where $\lambda_t' = (\lambda_t^1, \ldots, \lambda_t^j, \ldots, \lambda_t^d)$ in (4) is a d-dimensional row vector of Lagrange multipliers. By $\log \pounds^u(\mathbf{h})(d)$ and $\log \pounds^c(\mathbf{h})(d)$ we denote the values of the objective functions in (3) and (4) evaluated at the estimates $(\mathbf{a}^u, \mathbf{x}^u, \mathbf{q}^u)$ and $(\mathbf{a}^c, \mathbf{x}^c, \mathbf{q}^c, \lambda)$, respectively, for a fixed $d < \infty$. We call the estimator defined by (4) a constrained ML estimator.

3.2 *Results*

Our first result establishes that the estimation of (3) will generally not yield estimates that satisfy the constraint in Eq. (4). For example, inspection of NOMINATE proposal location estimates reveals that the constraint is not accidentally satisfied. More rigorous statements and proofs of all results appear in the Appendix.

Lemma 1 (Constraint Binds). Generically, ML estimates to the unconstrained problem (3) are not solutions to the constrained problem (4).

The intuition behind the result is straightforward. The constraint $q_t = y_{m(t)}$ is a linear restriction. In the space of possible estimates from unconstrained problems, X^{L+2T}, the set that also satisfies the constraint is a hyperplane of dimension $d(L+T)$. So, cases in which the constraint happens to be satisfied are knife-edged. This is not surprising, as one should not expect unconstrained estimators of a problem to turn out to also satisfy a constraint. In light of the lemma, we assume that the constraint binds.[4]

Proposition 1 (Ideal Point Difference). Under regularity conditions, if the constraint $q_t = y_{m(t)}$ binds, the ideal point estimates $\mathbf{x}^u$ from (3) are generically *not* equivalent to the ideal point estimates $\mathbf{x}^c$ from (4).

While proofs of the consistency of ML estimators of ideal points are plagued by the problems associated with the fact that as the sample size tends to infinity the number of estimated parameters and incidental parameters also tends to infinity, some results have been attained (Haberman 1977; Heckman and Snyder 1997; Kiefer and Walfowitz 1956; Wald 1948; Londregan 2000a). We sidestep this issue and make an observation regarding the asymptotic importance of imposing the constraint on nay locations.

Proposition 2 (Nonvanishing Differences). Under regularity conditions, if the constraint $q_t = y_{m(t)}$ binds in all but a finite number of sample sizes, and the estimators from (3) and (4) have limits (x^u_∞ and x^c_∞, respectively), then the ideal point estimates $\mathbf{x}^c$ from (4) and the estimates $\mathbf{x}^u$ from (3) generically do not have the same limit ($x^u_\infty \neq x^c_\infty$).

An immediate application of this result is that *if* the constrained ML estimator is consistent, then for a generic subset of true population ideal points, the unconstrained estimator of ideal points (3) is inconsistent.

For a fixed dimensionality d, the unconstrained and constrained estimators are well defined. Moreover, comparisons of the constrained (unconstrained) estimates with dimensionality of d and $d-1$ allows for a likelihood ratio test of the hypothesis that the policy space is $d-1$ dimensional (DeSarbo and Cho 1989; Londregan 2000a). The following can be shown.

Proposition 3 (Dimensionality Bias). If one is sufficiently concerned about avoiding Type II error, then the highest dimension which will not be rejected using likelihood ratio tests is (weakly) higher under the constrained model.

The interpretation of this proposition is that if we are sufficiently concerned about falsely concluding that the dimensionality is too small (i.e., β is close enough to 0), then the dimensionality estimate of the policy space computed via the comparison of likelihood ratios for the model that ignores the constraint [i.e., (3)] will lead us to conclude that the dimensionality is smaller than the algorithm that compares these ratios for the model that does not ignore the constraint [i.e., (4)].

The logic behind proposition 3 is similar to that behind the discussion of Fig. 2. Because the unconstrained model has more "free" parameters than the constrained model, there are configurations of true ideal points and bill locations which induce roll call voting that can be estimated in a lower-dimensional (unconstrained) model. The unconstrained estimator

[4] An empirical test of the validity of the constraint would require specifying the probability distribution of q_t conditional on $y_{m(t)}$. The outcome of testing the constraint of the form $\|q_t - y_{m(t)}\| \leq \delta$ for an arbitrary δ will hinge largely on the assumed conditional distributions of q_t given $y_{m(t)}$. Unfortunately, these distributions are likely to be unidentified.

achieves this by ignoring the relationship between yea and nay locations. The constrained estimator, however, does not have this freedom.

4 Agenda Constrained Bayesian Simulation Estimation

ML theory allows us to determine the significance of the constraint with a high degree of generality. To illustrate the effects of incorporating the constraint in a Bayesian simulation estimation procedure, we examine estimates of the first U.S. House (1789–1791). Specifically, this section compares constrained Bayesian simulation estimates to unconstrained and DW-NOMINATE estimates.

The construction of the agenda constrained Bayesian simulation estimator requires constraining parameter values of nay locations according to the legislative agenda. In the notation of Jackman (2000), a yea vote on vote j is a vote for the proposal location θ_j, and a nay vote is for the proposal location ψ_j. Given the discussion of the previous sections, we define the mapping between the status quo of vote j (i.e., ψ_j) and proposal parameters of vote $j-1$ by $\psi_j = \theta_{j-1}$ if θ_{j-1} passes and $\psi_j = \psi_{j-1}$ otherwise. The above mapping assumes that a vote for a proposal is a vote to move in the policy space, and a vote against a proposal is a vote to prevent movement. It is straightforward to show that incorporating the constraint identifies the proposal parameters (i.e., ψ_j and θ_j). Thus, in the one-dimensional setting examined, the constrained Bayesian simulation approach is able to recover proposal locations—not just the midpoint associated with vote j. As a consequence, it becomes possible to work with and test theories about the behavior of proposals.[5]

In the illustration that follows, we consider one-dimensional estimates for two reasons. First, this section is not intended as a source of empirical evidence in support of the analytic results in Section 3. Instead this section demonstrates the usefulness of imposing the constraint in Bayesian simulation estimation, just as the previous section demonstrates the consequences of failing to impose the constraint in ML estimation. Second, calculation of the unconstrained (and constrained) Bayesian simulation estimator in more than one dimension introduces several identification issues. Given the complications identified by Jackman (2001) in the simpler unconstrained Bayesian simulation estimator, addressing these concerns is a paper in its own right.

4.1 *Estimators Compared: The First U.S. House*

There are several aspects which make the first House an interesting illustration. First, the codification of the status quo that we employ is reasonable given the agenda structure that prevailed. Inspection of the voting agenda reveals that a large number of similar bills were considered sequentially (e.g., votes 13 through 23 and votes 63 through 75 all deal with determining the location of the new seat of government). Second, given the small number of issues debated, the gain from correctly capturing the interdependence of votes is potentially large. Louis-Guillaume Otto, the Chargé d’affaires of France, corresponding in 1790, described the business of the first Congress:

> The first session of [the first] Congress has had for its aim only the general organization of the government. The second will be more important and more delicate: it will decide about the money and the army . . . . A third object, much less interesting may give a more perceptible shock to the new confederation. It is the eternal discussion about the residence [i.e., location of the capitol]. (O’Dwyer 1964)

[5]Note that although the NOMINATE algorithm provides bill locations, their identification depends critically on the specific parametric assumptions made about the statistical and utility models.

Third, the small number of proposals enables us to summarize graphically the movement of the status quo through the policy space.

Figure 3 presents the legislator ideal point estimates produced by both the unconstrained and the constrained Bayesian simulation estimators.[6] The posterior distribution for each was generated via an initialization run of 400,000 iterations with uninformative (diffuse) priors. The subsequent 100,000 iterations were thinned by a factor of 100 to yield the 1000 realizations that were used to summarize the posterior. Thinning was utilized to reduce the autocorrelation between the draws and use iterations that explore more of the parameter space.

To highlight the resulting differences in the recovered estimates of the constrained and unconstrained estimator, both are presented in Fig. 3. The legislators are rank-ordered according to the posterior mean of the constrained estimator, with the open symbols representing the mean of the legislator posterior distributions, and the horizontal lines denoting the 95% posterior confidence intervals.

Consistent with previous results, Fig. 3 reveals that the uncertainty associated with the ideologically extreme legislators' estimates is greater than that of centrally located legislators. This differential uncertainty results from the lack of bills with extreme cutpoints, as it is the presence of cutpoints that enables us to distinguish between legislators. Second, because the Bayesian simulation approach provides for the (automatic) incorporation of missing roll call data by treating missing data as additional parameters to be estimated, estimates of legislators with a high number of abstentions (e.g., Bourn and F. Muhlengberg) are recovered with more uncertainty.

Using standard assessments of the similarity of estimators, the difference between the estimates seems negligible. The Pearson and Spearman (rank-order) correlations between the constrained and the unconstrained estimates are extremely high (.971 and .975 respectively). The Pearson and Spearman correlations of the constrained estimates and DW-NOMINATE are also fairly high, at .77 and .84, respectively, with the correlations between the unconstrained estimates and DW-NOMINATE being .84 and .90, respectively.[7]

Consideration of correlations alone may be misleading. Despite the appearance of similarity, it is possible to identify some clear differences between the estimators. First, since in Fig. 3 the legislators are ordered by the mean of the constrained posterior density (left), the fact that the unconstrained posterior means (right) are not monotonic in the ordered legislators demonstrates that the two estimators do not recover the same ordering, of legislators. If the estimators recovered the same ordering, then both the constrained and the unconstrained posterior means would increase as one moves from Matthews to Sturges.

Second, comparing the confidence intervals demonstrates that these differences in estimates are not trivial relative to the precision of the estimates. In the unconstrained estimates, the ideal point distributions of Sedgwick and Floyd contain hardly any overlap. However, the estimates of the constrained model reveal that Floyd and Sedgwick are almost indistinguishable. This phenomenon is not restricted to extreme legislators, as the ideal point distributions of the centrally located Seney and Sinnickson overlap in the unconstrained

[6]We maintain the distributional assumptions of Jackman (2000) where possible, although since the posterior distribution of the constrained estimator does not retain the conjugacy of the unconstrained estimator, the Metropolis–Hastings algorithm is employed.

[7]The DW-NOMINATE scores were graciously provided by Keith Poole. Note that there are three legislators in the ICPSR roll call data file without a DW-NOMINATE score: F. Muhlengberg, Bourn, and Giles. Note that some part of the difference in correlations is due to the fact that the constrained estimates are estimated with more variance.

P1: FIC/'Seema'
WV007B-Clinton May 14, 2001 21:7

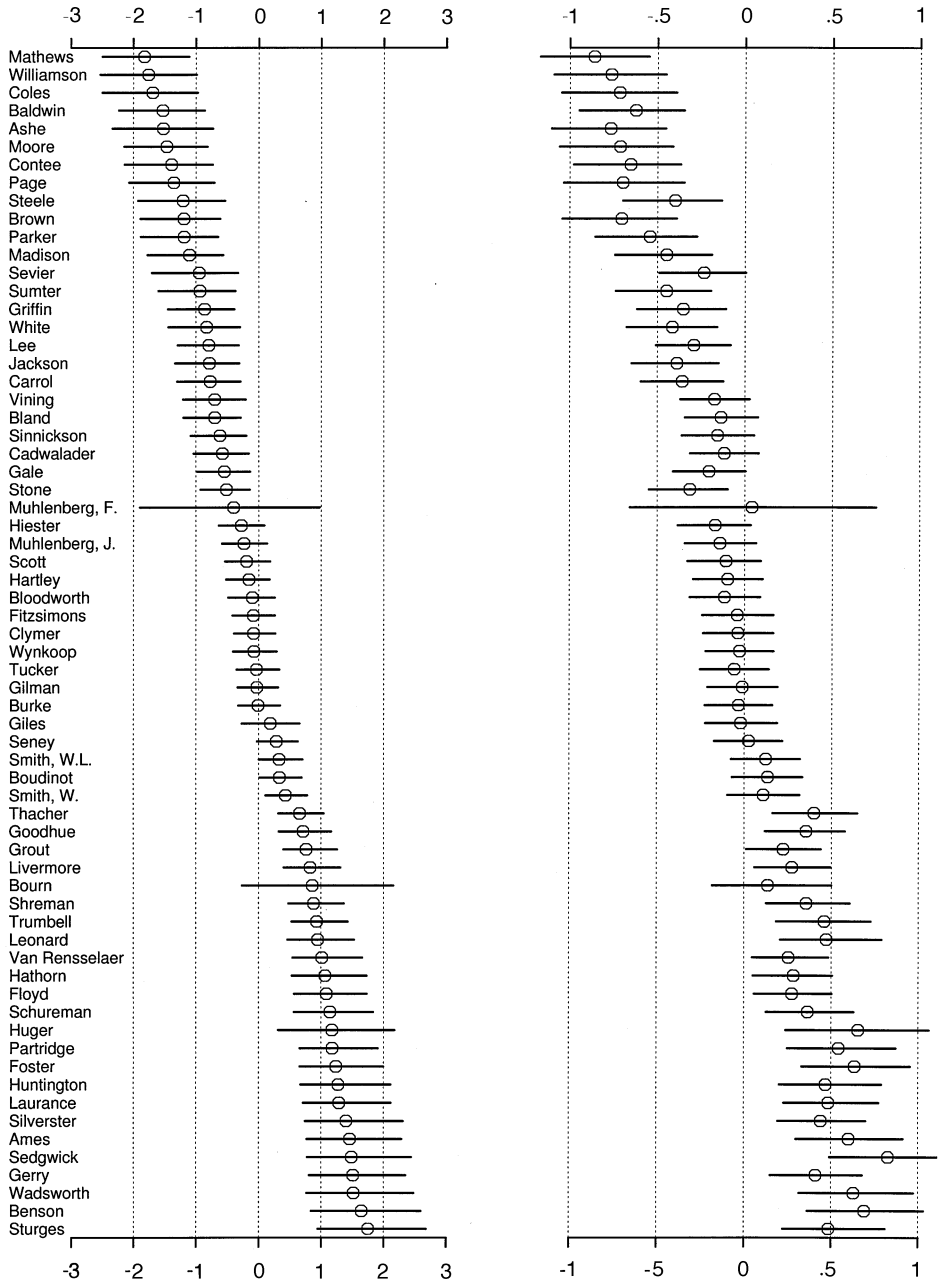

Fig. 3 Legislator ideal point estimates produced by the unconstrained (right) and the constrained (left) Bayesian simulation estimators.

estimator but not in the constrained estimator. Although certainly not definitive, such differences provide evidence that the constraint does indeed affect the ideal points that are recovered.

We quantify the relative uncertainty associated with the constrained and unconstrained ideal point estimates by comparing the standardized mean difference between the upper and the lower bounds of the 95% confidence intervals. Letting δ_i^c (δ_i^u) denote the length of the confidence interval of legislator $i \in L$ in the constrained (unconstrained) model and $\acute{\delta}^c$ ($\acute{\delta}^u$) denote the maximum of these distances over i for the respective estimators, the statistics are $\sigma^c = \sum_i \delta_i^c/(L\acute{\delta}^c)$ and $\sigma^u = \sum_i \delta_i^u/(L\acute{\delta}^u)$. We find that $\sigma^c = .40$ and $\sigma^u = .37$. The ordering of these two numbers is not surprising, as one would expect more error in the constrained model. Recalling the discussion in the beginning of Section 3, it is not possible to test the validity of the constraint without formulating a framework in which one can interpret the constrained and unconstrained estimates. Consequently the relative sizes of σ^c and σ^u should not be construed as evidence that the unconstrained model is preferred to the constrained model.

Although incorporation of the constraint yields different ideal point estimates, the greater innovation provided by the constraint is that it enables us to recover proposal locations—not just the midpoint (as is the case in the unconstrained estimator). As a result, it is possible to trace the trajectory of the policy agenda. Figure 4 presents this information.

The horizontal axis in Fig. 4 denotes the policy location of the proposal, and the vertical axis presents the agenda sequence—with the initial status quo represented by the open square (top), and subsequent proposals represented by open or filled diamonds depending on whether the proposal fails or passes. The locations of the legislators' posterior means are indicated by the short vertical lines located just above the horizontal axis, with the median legislator's ideal point denoted by the vertical line that extends through the graph.

There exists a substantial amount of uncertainty in the location of successful proposals. For the 108 proposals considered in the first U.S. Congress, 26 of the 51 (51%) successful proposals have 95% posterior confidence bounds that span the entire range of recovered (mean) ideal points. In contrast, only 21 of the 57 (37%) unsuccessful proposals span that range. This may seem counterintuitive given that the location of successful proposals appear in the likelihood function at least twice as many times as unsuccessful proposals (i.e., it is a yea location on the first vote and then a nay location until a proposal beats it). However, in a one-dimensional model, when a yea location enters into the likelihood function many times, unless the voting coalitions are connected on each vote, error is introduced into the yea location estimate.

This phenomenon provides weak evidence either that a one-dimensional solution is insufficient for the constrained estimation problem or that log rolls occurred on some votes—causing disconnected coalitions to form. The possibility of the latter is especially relevant in the first House, where a log roll between issues dealing with the assumption of Revolutionary war debt and the location of the capitol is commonly termed the Compromise of 1790 (Jefferson 1829). However, the existence of this log roll is highly contested (including Bowling 1968, 1971; Cooke 1970; Risjord 1976). Assessing whether a higher-dimensional model is better for this Congress is beyond the scope of this article.

Inspection of the proposal path through the issue space reveals several interesting findings. As expected, successful policies overlap the median voter's ideal point far more often (33 of 51) than do unsuccessful policies (21 of 57). However, it is still easy to find cases where a noncentrist proposal defeats a centrist status quo.

The seventh roll call passed H.R.19, a bill to allow for compensation to members of Congress, their clerks, and officers, by a vote of 30–16. Inspection of Fig. 4 demonstrates that the proposal's posterior mean is outside the Pareto set generated by the legislators' posterior means.

Although there is substantial movement between the seventh and the sixth roll call, it is important to recall the interpretation of such movement. Although the sixth proposal is a vote on receding from certain Senate amendments on a proposal establishing a duty on ships

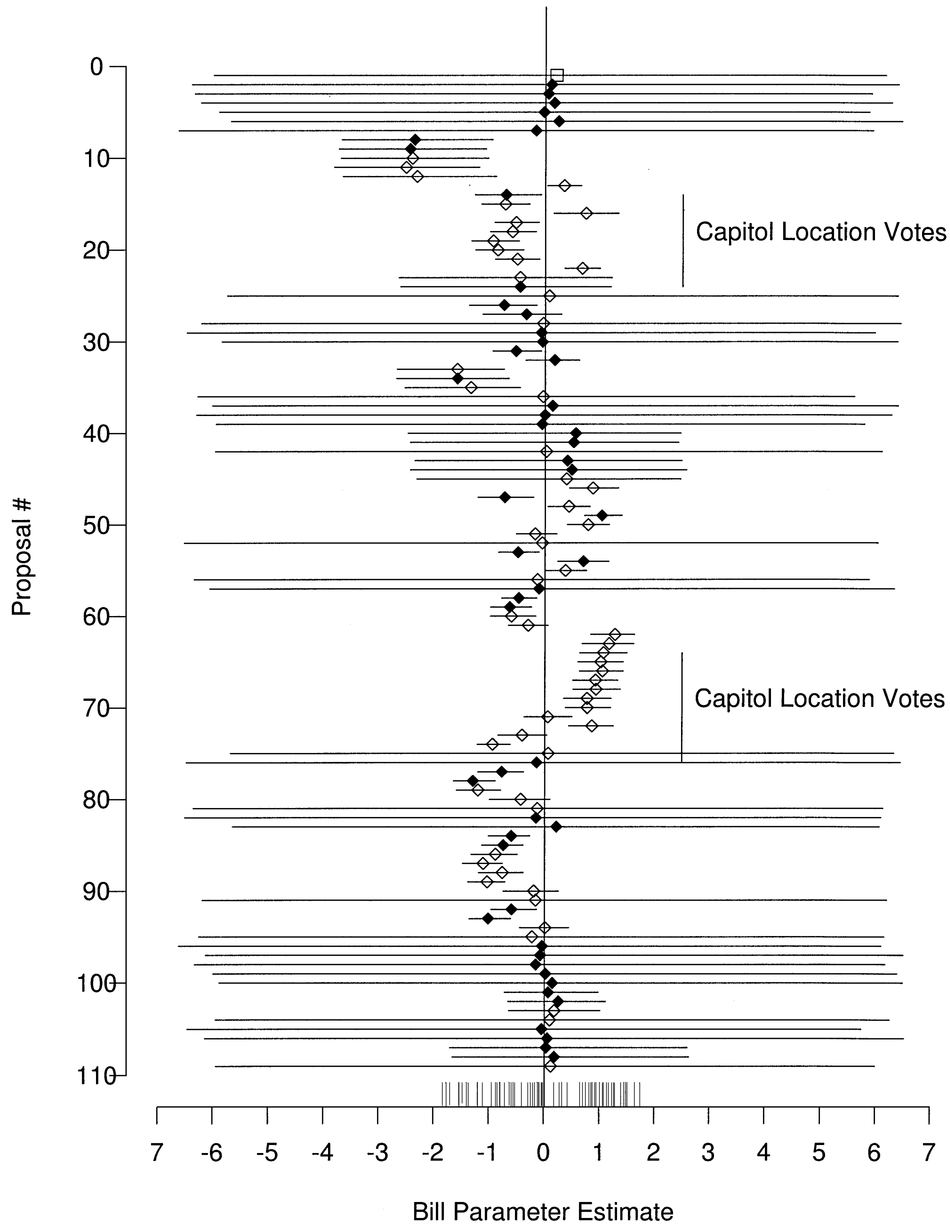

Fig. 4 Legislator ideal point estimates, proposal locations, and trajectory of the policy agenda using the constrained model.

built in the United States (H.R.5), the *location* of proposal 6 in Fig. 4 represents the state of affairs which results after the passage of the previous five bills. Similarly, the location of the seventh roll call in Fig. 4 summarizes the state of affairs with the modification that now Congressmen and their clerks and officers receive compensation. In other words, whereas the location of the sixth proposal represents the state of the world in which the first six proposals pass, the location of the seventh proposal represents the state of the world in which the first six proposals pass *and* Congressmen are compensated.

As a substantive observation, the 75th roll call passed S.12 and established Washington, D.C. (Philadelphia), as the permanent (temporary) capitol by a 32–29 margin. Despite Otto's prediction that this debate would be "uninteresting," 25 of the first 75 votes pertained to this issue, with Fig. 4 identifying two sequences of such votes. Examination of Fig. 4 also reveals that the posterior mean location of the 75th roll call passing S.12 is very close to the median legislator's ideal point—much closer than most of the previous (unsuccessful) proposals. Given the closeness of the vote, the size of the proposal's confidence interval is surprising—highly suggestive of the possibility of a log roll.[8]

Although we cannot compare proposal locations of the constrained and unconstrained estimators because the locations are identified only for the constrained estimator, it is possible to compare the recovered cutpoints of the roll call votes. The Pearson and Spearmen correlations between the constrained and the unconstrained cutpoints are .25 and .24, respectively. In contrast to the correlations of ideal point estimates, incorporation of the constraint greatly affects estimates of parameters associated with the legislative agenda.

This discussion is not intended to be a structured analysis of the first U.S. Congress. Such an endeavor would require far more investigation, consideration and knowledge than we have brought to the topic. However, the above observations demonstrate two main points. First, inclusion of the constraint does seem to affect the inferences that one draws. Second, identification of proposal locations by using information contained in the agenda opens up several avenues for future research.

5 Conclusion

To the extent that contemporary scholars of American politics utilize preference estimates to test and calibrate theories imbedded in the SMV, it is desirable for these estimates to share the basic assumptions of the SMV. We argue that the assumptions of existing procedures are incompatible with the SMV and illustrate the resulting pathologies. By constraining the nay locations, both ML and Bayesian simulation estimators of legislative roll call voting can be improved so as to make the resulting estimates more consistent with the theories being tested. The relevance of the constraint is analytically proven in the ML framework and illustrated with data in the context of the Bayesian simulation approach.

In addition, incorporation of the constraint provides researchers with an important new set of estimates—proposal locations. Given these estimates (or ones based on a more detailed description of the agenda process), it becomes possible to test questions such as, What is the behavior of the status quo over time? Are extreme bills more likely to get closed rules than open rules? and Do log rolls occur? Given the nature of these questions, it is clear that incorporation of the constraint will provide valuable information in the quest to characterize the behavior of Congress.

On a different scale, empirical investigation of proposal-making and roll call voting in specific policy areas is an endeavor in which the insight of this paper may be most rewarding (e.g., Krehbiel and Rivers 1988). In this setting, knowledge of procedural minutiae may be used to model the relationship between votes more accurately. Furthermore, with a small number of substantively related roll call votes and a large number of legislator ideal points to estimate, the identification granted by the use of the agenda may be especially valuable.

[8] In fact, we hypothesize that a statistical test of the log rolling hypothesis may be constructed using the uncertainty estimates on such a vote.

Appendix

In this appendix we make explicit the concepts discussed in Section 3 and rigorously restate the lemma and propositions with proofs.

We assume that the gradients $D_{x_i}\rho(x_i, y_t, q_t)$, $D_{q_t}\rho(x_i, y_t, q_t)$, $D_{y_t}\rho(x_i, y_t, q_t)$ and the corresponding Jacobian of second derivatives exist on X^3. We also assume that $\rho(\cdot, \cdot, \cdot)$ is not too flat. Specifically, for each i, we require that the subset of X^3 for which

$$\sum_{t=1}^{T} D_{x_i}\left(\frac{D_{x_i}(\rho(x_i^*, y_t, q_t))}{\rho(x_i^*, y_t, q_t)}\right) = \sum_{t=1}^{T} D_{x_i}\left(\frac{D_{x_i}(\rho(x_i^*, y_t, y_{m(t)}))}{\rho(x_i^*, y_t, y_{m(t)})}\right) \tag{A1}$$

holds has an empty interior.[9] If this condition fails, a very knife-edged distribution and agenda pair is being considered. Note that if $\rho(\cdot, \cdot, \cdot)$ is induced by normal or extreme value errors with legislators having quadratic or exponential preferences, then the condition holds.

We now consider Lemma 1. For fixed d, L, T, the space of possible optimization problems is infinite dimensional. However, for fixed d, L, T, the space of possible unconstrained estimates is X^{L+2T}, a subset of finite-dimensional Euclidean space. It is in this space that we work. Note that for any $A \subset X^{L+2T}$ with an empty interior, the Lebesgue measure of A [denoted $leb(A)$] is 0.

Lemma 1. In the space of unconstrained ML estimates to the unconstrained problem (3), the set which would also be solutions to the constrained problem (4) has Lebesgue measure 0.

Proof. If $(\mathbf{a}^u, \mathbf{x}^u, \mathbf{q}^u)$ is the solution to (3), a necessary condition for it to be the solution to (4) is that $q_t^u = y_{m(t)}^u$. In the space X^{L+2T} the set $A := \{(\mathbf{a}, \mathbf{x}, \mathbf{q}) : q_t = y_{m(t)}\}$ is the intersection of the hyperplane $H = \{(\mathbf{a}, \mathbf{x}, \mathbf{q}) : \mathbf{a} \in \mathbb{R}^{dL}, \mathbf{x} \in \mathbb{R}^{dT}, q_t = y_{m(t)}\}$ and X^{L+2T}. Since $leb(H) = 0$, $leb(A) = 0$. ■

A stronger result regarding the space of ML problems and not the space of ML estimates is available at the *Political Analysis* web site. We now prove propositions 1 and 2.

Proposition 1. If the constraint $q_t = y_{m(t)}$ binds, then ideal point estimates from (3) are generically not equivalent to ideal point estimates from (4).

Proof. Assume the hypothesis. The first-order conditions with respect to the ideal point estimates for the two problems for each i are, respectively,

$$\sum_{t=1}^{T} \frac{v_{it}}{\rho(x_i, y_t, q_t)} D_{x_i}(\rho(x_i, y_t, q_t)) - \sum_{t=1}^{T} \frac{1 - v_{it}}{\rho(x_i, y_t, q_t)} D_{x_i}(\rho(x_i, y_t, q_t)) = 0 \tag{A2}$$

$$\sum_{t=1}^{T} \frac{v_{it}}{\rho(x_i, y_t, y_{m(t)})} D_{x_i}(\rho(x_i, y_t, y_{m(t)})) - \sum_{t=1}^{T} \frac{1 - v_{it}}{\rho(x_i, y_t, y_{m(t)})} D_{x_i}(\rho(x_i, y_t, y_{m(t)})) = 0 \tag{A3}$$

[9] With slight abuse of matrix notation we use the fraction notation $D_{x_i}(\rho(\cdot, \cdot, \cdot))/\rho(\cdot, \cdot, \cdot)$ to denote the product of a scaler $1/\rho(\cdot, \cdot, \cdot)$ and a vector $D_{x_i}(\rho(\cdot, \cdot, \cdot))$. The term empty interior is relative to the Euclidean topology.

In (A3) the fact that $q_t = y_{m(t)}$ has been used. Moreover, by assumption the constraint is not satisfied in (3) so that $\mathbf{q}^u \neq \mathbf{q}^c$. It remains to verify that generically the same $\mathbf{x}$ cannot solve both (A2) and (A3). It is sufficient to show that given any $\mathbf{x}$ which solves both (A3) and (A2), for every neighborhood of that point in X^L there exists a point $\mathbf{x}'$ which solves (A3) but not (A2), as this implies that the set of constrained ML ideal point estimates $\mathbf{x}$ which are also unconstrained ML estimates has an empty interior and is thus the complement of a generic subset of X^L. To show this, assume that there is a vector $\mathbf{x}^*$ which solves both (A2) and (A3). From (A2) and (A3) this implies that, for each i,

$$\sum_{t=1}^{T} \frac{2v_{it} - 1}{\rho(x_i^*, y_t, q_t)} D_{x_i}(\rho(x_i^*, y_t, q_t)) = \sum_{t=1}^{T} \frac{2v_{it} - 1}{\rho\left(x_i^*, y_t, y_{m(t)}\right)} D_{x_i}\left(\rho(x_i^*, y_t, y_{m(t)})\right) \quad \text{(A4)}$$

We now consider an arbitrary ε−ball of $\mathbf{x}^*$ denoted $B(\mathbf{x}^*, \varepsilon)$. Since $v_{it} \neq \frac{1}{2}$, if (A4) holds at every $\mathbf{x} \in B(\mathbf{x}^*, \varepsilon)$, for each i we must have

$$\sum_{t=1}^{T} D_{x_i}\left(\frac{D_{x_i}(\rho(x_i^*, y_t, q_t))}{\rho(x_i^*, y_t, q_t)}\right) = \sum_{t=1}^{T} D_{x_i}\left(\frac{D_{x_i}\left(\rho\left(x_i^*, y_t, y_{m(t)}\right)\right)}{\rho\left(x_i^*, y_t, y_{m(t)}\right)}\right) \quad \text{(A5)}$$

at every $\mathbf{x} \in B(\mathbf{x}^*, \varepsilon)$, contradicting the assumption that $(\rho(\cdot,\cdot,\cdot), \mathbf{y})$ is not too flat. Thus there is a point arbitrarily close to $\mathbf{x}^*$ for which the equivalence in (A4) does not hold. Thus the result follows. ■

Proposition 2. If the constraint binds in all but a finite number of sample sizes, and the estimators from (3) and (4) almost surely converge to $\mathbf{x}^u_\infty$ and $\mathbf{x}^c_\infty$, respectively, then $\mathbf{x}^u_\infty \neq \mathbf{x}^c_\infty$.

Proof. Assume the hypothesis. By the fact that the Jacobian exists, the first-order conditions for problems (3) and (4) are continuous. Since $\mathbf{x}^c \to \mathbf{x}^c_\infty$ and $\mathbf{x}^u \to \mathbf{x}^u_\infty$, the continuous mapping theorem (Durrett 1995) implies that the sample averages of the left-hand side of the first-order conditions (A2) and (A3) converge to the constrained population first-order conditions evaluated at $\mathbf{x}^c_\infty$ and $\mathbf{x}^u_\infty$. But the arguments of the above proof applied to the limiting first-order conditions yield the result. ■

We now formalize the test of dimensionality. Assuming that (4) is the correct model, it is well known that the restriction from d to $d' = d - 1$ can be tested using $L^c(d, \mathbf{h}) := 2((\log \pounds^c(\mathbf{h})(d) - \log \pounds^c(\mathbf{h})(d - 1))$. We assume that this statistic has distribution $\mathcal{X}^2_{L+T}$ (α). We denote the critical value of the distribution $\mathcal{X}^2_{L+T}$ (α) by $c(L + T, \alpha)$.[10] Defining $L^u(d, \mathbf{h})$ in an analogous manner, we may consider the stochastic process of likelihood ratio values (LLRs) $\{L^c(d, \mathbf{h})\}_{d=1}^{\infty}$ and $\{L^u(d, \mathbf{h})\}_{d=1}^{\infty}$. The estimated dimensionalites are $d^c(\boldsymbol{h}; \alpha) := \min\{n \in \mathbb{N} : L^c(d, \boldsymbol{h}) \leq c(L+T, \alpha)\}$ and $d^u(\boldsymbol{h}; \alpha) := \min\{n \in \mathbb{N} : L^u(d, \boldsymbol{h}) \leq c(2T + L, \alpha)\}$, where $\mathbb{N}$ denotes the set of natural numbers. For a finite-dimensional data set and fixed α, both $d^c(\boldsymbol{h}, \alpha)$ and $d^u(\boldsymbol{h}, \alpha)$ are finite.

It is clear that if $L^c(d, \mathbf{h}) \geq L^u(d, \mathbf{h})$ for almost every $\mathbf{h}$, the constrained procedure will find a (weakly) higher dimensionality than the unconstrained procedure. Thus, it is sufficient to show that $L^c(d, \mathbf{h}) \geq L^u(d, \mathbf{h})$ and $L^c(d, \mathbf{h}) > L^u(d, \mathbf{h})$ for a nonnull set of

[10]The critical value $c(n, \alpha)$ is just the value found from a standard $\mathcal{X}^2$ table with n degrees of freedom and a P value of $1 - \alpha$.

h's to show that $E_{\mathbf{h}} d^c(\mathbf{h}; \alpha) > E_{\mathbf{h}} d^u(\mathbf{h}; \alpha)$. The operator $E_{\mathbf{h}}(\cdot)$ takes the expectation of an **h**-measurable function with respect to the underlying data (i.e., roll call votes).[11]

We assume that $\rho(\cdot, \cdot, \cdot)$ can be expressed as $\Pr(\omega < \|x_i - q_t\| - \|x_i - y_t\|)$, where ω is a random variable whose distribution has at most a countable number of discontinuities, and that for any **a**, **x**, **q**, any pattern of roll call voting is possible.

Proposition 3 (Dimensionality Bias). Fix L and T. There exists $\varepsilon > 0$ s.t. $E_{\mathbf{h}} d^c(\boldsymbol{h}; \alpha) > E_{\mathbf{h}} d^u(\mathbf{h}; \alpha)$ for all $\alpha > 1 - \varepsilon$.

Proof. Assume the hypothesis. First, we establish the existence of $E_{\mathbf{h}} d^c(\mathbf{h}; \alpha)$ and $E_{\mathbf{h}} d^u(\mathbf{h}; \alpha)$, then we establish the ordering. Since L, T are finite and any $\mathbf{h} \in \mathbf{H}$ (the set of $L \times T$ matrices with elements in $\{0, 1\}$) can be predicted via perfect voting in a sufficiently high-dimensional space, for any $\mathbf{h} \in \mathbf{H}$ there exists a $\overline{d^c}$ and $\overline{d^u}$ s.t. $L^c(\overline{d^u}, \mathbf{h}) = 0$ and $L^c(\overline{d^c}, \mathbf{h}) = 0$. This implies that $d^c(\mathbf{h}; \alpha)$ and $d^u(\mathbf{h}; \alpha)$ exist for each $(\mathbf{h}, \alpha)$. Since L, T finite implies that $\mathbf{H}$ is finite, the expectations $E_{\mathbf{h}} d^c(\boldsymbol{h}; \alpha)$, $E_{\boldsymbol{h}} d^u(\boldsymbol{h}; \alpha)$ exist for each α.

For some $\varepsilon > 0$ and for each $h \in H$, $\alpha \geq 1 - \varepsilon$ implies that $d^c(h; \alpha) = \overline{d^c}$ and $d^u(\mathbf{h}; \alpha) = \overline{d^u}$. It remains only to establish that for each $h \in H$, $\overline{d^c} \geq \overline{d^u}$ and for some $\mathbf{h} \in \mathbf{H}$, $\overline{d^c} > \overline{d^u}$. The former fact is true because any profile $(\mathbf{a}^c, \mathbf{x}^c, \mathbf{q}^c)$ from (4) for which perfect voting yields the data **h** is also admissible for (3), and under the constrained model this profile will yield perfect voting that induces the roll call **h**. To illustrate that for some $\mathbf{h} \in \mathbf{H}$, $\overline{d^c} > \overline{d^u}$, we exhibit an algorithm that for a given $L, T \geq 3$ will yield an **h** for which $\overline{d^c} > \overline{d^u}$: Taking $L_3 := \min\{n \leq L : L \text{ is divisible by } 3\}$ and $T_3 := \min\{n \leq T : L \text{ is divisible by } 3\}$, construct the $L \times T$, **h** matrix as

$$\mathbf{h}_{ij} = \begin{cases} 1 & \text{if} \quad \{i \leq L_3/3\} \quad \text{or} \quad \{i \leq 2L_3/3 \quad \text{and} \quad j \leq T_3/3\} \quad \text{or} \quad \{j \in (2T_3/3, T)\} \\ 0 & \text{otherwise} \end{cases}$$

The matrix h corresponds to the replication of the roll call in the third example of L_3 legislators and T_3 votes, with the leftover voters voting the same as the third $L_3/3$ voters. The leftover votes are all unanimous rejections. Since this roll call is perfectly fit in one dimension without the constraint and not with the constraint, we have $\overline{d^c} > \overline{d^u}$. Thus, the result is shown. ∎

References

Bowling, Kenneth R. 1968. *Politics in the First Congress, 1789–1791*, Ph.D. dissertation. Madison: University of Wisconsin.

Bowling, Kenneth R. 1971. "Dinner at Jefferson's: A Note on Jacob E. Cooke's 'The Compromise of 1790.'" *William and Mary Quarterly* 28:629–648.

Clinton, Joshua D., Simon Jackman, and Douglas Rivers. 2000. "The Statistical Analysis of Legislative Behavior: A Unified Approach," Stanford University typescript. Palo Alto, CA: Stanford University.

Cooke, Jacob E. 1970. "The Compromise of 1790." *William and Mary Quarterly* 27:523–545.

[11] We would like to establish that $L^c(d, \mathbf{h}) > L^u(d, \mathbf{h})$ and, thus, be confident that $E_{\mathbf{h}} d^c(\mathbf{h}; \alpha) > E_{\mathbf{h}} d^u(\mathbf{h}; \alpha)$ for all α levels. Proof of this claim requires analysis of the convexity properties of the likelihood function and offers little intuition. Instead, we worry only about sufficiently high α's so that the recovered dimensionalities $d^c(\mathbf{h}; \alpha)$ and $d^u(\mathbf{h}; \alpha)$ are those for which nearly perfect voting yields the recovered data set. Thus, the result should be interpreted as stating that the unconstrained estimate yields an *underestimate* of the dimensionality of the policy space if one is sufficiently concerned about not finding the dimensionality of the policy space smaller than it really is. Since α and β are inversely related, we formulate the analysis for sufficiently high α so as to pertain to the case of sufficiently low β (type II error).

DeSarbo, Wayne, and Jaewun Cho. 1989. "A Stochastic Multidimensional Scaling Vector Threshold Model for the Spatial Representation of *Pick Any/N* Data." *Psychometrika* 54:105–129.

Durrett, Richard. 1995. *Probability: Theory and Examples*, 2nd ed. Belmont, CA: Duxbury Press.

Haberman, S. J. 1977. "Maximum Likelihood Estimates in Exponential Response Models." *Annals of Statistics* 5:815–841.

Heckman, James, and James Snyder. 1997. "Linear Probability Models of the Demand for Attributes with and Empirical Application to Estimating the Preferences of Legislators." *Rand Journal of Economics* 28:142–189.

Jackman, Simon. 2000. "Estimation and Inference Are Missing Data Problems: Unifying Social Science Statistics via Bayesian Simulation." *Political Analysis* 8:307–322.

Jackman, Simon. 2001. "Estimation and Inference for Legislative Ideal Points by Bayesian Simulation: Extensions and Elaborations," Stanford University typescript. Palo Alto, CA: Stanford University.

Jefferson, Thomas. 1829. Memoir, Correspondence and Miscellanies, from the Papers of Thomas Jefferson, ed. Thomas Jefferson Ranoolph. Charlottesville, VA: F. Case and Co.

Kiefer, J., and J. Wolfowitz. 1956. "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters." *Annals of Mathematical Statistics* 27:887–896.

Krehbiel, Keith, and Doug Rivers. 1988. "An Analysis of Committee Power: An Application to Senate Voting on the Minimum Wage." *American Journal of Political Science* 32:1151–1174.

Londregan, John. 2000a. "Estimating Legislator's Preferred Points." *Political Analysis* 8:35–56.

Londregan, John. 2000b. *Legislative Institutions and Ideology in Chile*. Cambridge: Cambridge University Press.

Miller, Gary J. 1993. "Formal Theory and the Presidency." In *Researching the Presidency*, eds. George C. Edwards III et al. Pittsburgh: University of Pittsburgh Press.

O'Dwyer, Margaret M. 1964. "A French Diplomat's View of Congress 1790." *William and Mary Quarterly* 21:408–444.

Poole, Keith T. 2000. "Nonparametric Unfolding of Binary Choice Data." *Political Analysis* 8:211–237.

Poole, Keith, and Howard Rosenthal. 1996. *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford Press.

Risjord, Norman K. 1976. "The Compromise of 1790: New Evidence on the Dinner Table Bargain." *William and Mary Quarterly* 33:309–314.

Wald, A. 1948. "Estimation of a Parameter When the Number of Unknown Parameters Increases Indefinitely with the Number of Observations." *Annals of Mathematic Statistics* 19:220–227.