*Patterns of Congressional Voting**

Keith T. Poole, *Carnegie Mellon University*
Howard Rosenthal, *Carnegie Mellon University*

Congressional roll call voting has been highly structured for most of U.S. history. The structure is revealed by a dynamic, spatial analysis of the entire roll call voting record from 1789 to 1985. The space is characterized by a predominant major dimension with, at times, a significant, but less important second dimension. In the modern era, spatial positions are very stable. This stability is such that, under certain conditions, short run forecasting of roll call votes is possible. Since the end of World War II, changes in congressional voting patterns have occurred almost entirely through the process of replacement of retiring or defeated legislators with new members. Politically, selection is far more important than adaptation.

1. Introduction

The Congress of the United States is a complex legislative institution subject to a myriad of formal and informal rules. Legislative action typically requires the assent of numerous committees and subcommittees, as well as the support of party leaders. Furthermore, legislation is shaped not only by the 535 members of Congress and attendant thousands of staff but also by influences arising in the executive, organized lobbies, the media, and from private individuals. One important outcome of these various processes are the recorded roll call votes taken on the floors of the two houses of Congress.

Beneath the apparent complexity of Congress, we find that these roll call decisions can largely be accounted for by a very simple dynamic voting model.[1] In a spatial model (see Enelow and Hinich 1984; Ordeshook 1986), each legislator is represented by a point in s-dimensional Euclidean space. Each roll call, whether it be a key vote on a civil rights bill or a mundane motion to restore Amtrak service in Montana, is represented by two points that correspond to the policy consequences of the yea and nay outcomes. The spatial model holds that a legislator prefers the closer of the two alternatives. The extent of preference is expressed by a utility function. The closer an alternative is to the legislator's

*We thank John Londregan and Tom Romer for many helpful comments. Our research was supported in part by National Science Foundation grant SES-8310390. The estimation was performed on the Cyber 205s of the John Von Neumann Center at Princeton University. Other portions of the analysis were carried out at the Pittsburgh Supercomputer Center.

[1]The fact that roll call voting can be accounted for by a simple model does not imply that all strategic complexities in Congress fit into this mold. Van Doren (1990) has stressed a number of ways that focusing solely on roll calls induces "sample selection bias" in arriving at substantive conclusions. In particular, Krehbiel (1986) and Smith and Flathman (1989) have emphasized that a great deal of important business is handled by unanimous consent agreements or voice vote.

American Journal of Political Science, Vol. 35, No. 1, February 1991, Pp. 228–78

ideal point, the greater the preference for the alternative, and the higher the utility.

Although our work shows that a low dimensional Euclidean model largely captures the structure of congressional voting, we should stress that the work says nothing about how specific issues get defined in terms of the structure. We cannot, for example, explain why Robert Bork was rejected by the Senate while a perhaps equally conservative Supreme Court nominee, Antonin Scalia, was confirmed by a 99–0 vote. Later in the paper, we do show that it would have been possible, using our model, to have accurately predicted the Bork vote on the basis of announced positions by members of the Judiciary Committee. In other words, once the positions of the alternatives have been defined, a spatial model can predict the outcome. But we have to leave to other research the all-important task of predicting how substantive issues get mapped into alternatives in the space.

What the spatial model does assert is that voting alignments must largely remain consistent with spatial positions. Thus, the lobbying process—involving interest groups and the White House—can be seen as a set of efforts to alter the location of the cutting line on an issue.

The development of supercomputing has enabled us to estimate the spatial model for the period from 1789 through 1985. The spatial structure uncovered is very stable, with two exceptions that occurred when the two-party system had major breakdowns. The first was from 1815 to 1825, after the collapse of the Federalist party; the second was in the early 1850s during the collapse of the Whigs and the division over slavery. Since the Civil War, the structure has been sufficiently stable that the major evolutions of the political system can be traced in terms of repositioning within the structure. The Great Depression, for example, witnessed a massive influx of "liberals," but there was no sharp break with pre-Depression voting patterns.

The paper proceeds, in part 2, with a description of the behavioral model that represents this simple structure of roll call voting and a brief explanation of the estimation method, dubbed D-NOMINATE for *D*ynamic *Nomina*l *T*hree-Step *E*stimation. (The Appendix provides details concerning the estimation technique, statistical issues in the estimation, and Monte Carlo tests of the method.) In part 3 we present evidence that, on the whole, the space is of low dimensionality in which legislators occupy temporally stable *relative* positions. The argument in part 4 is directed at establishing that issues of slavery and civil rights for Afro-Americans are the major source of exception to a unidimensional, stable space. In part 5 we briefly illustrate the predictive capacities of the model with an analysis of the confirmation vote on the Bork nomination to the Supreme Court. In the conclusion we point to two key findings. On the one hand, in contrast to earlier historical periods, political change must now be accomplished by the selection of new legislators through the electoral process rather than by the adaptation of incumbent legislators to changes in public demands. On the other hand, the possibility of major political change has been sharply reduced because the average distance between the two major parties has fallen dramatically in this century.

2. Estimation of a Probabilistic, Spatial Model of Voting

An Overview of the Model

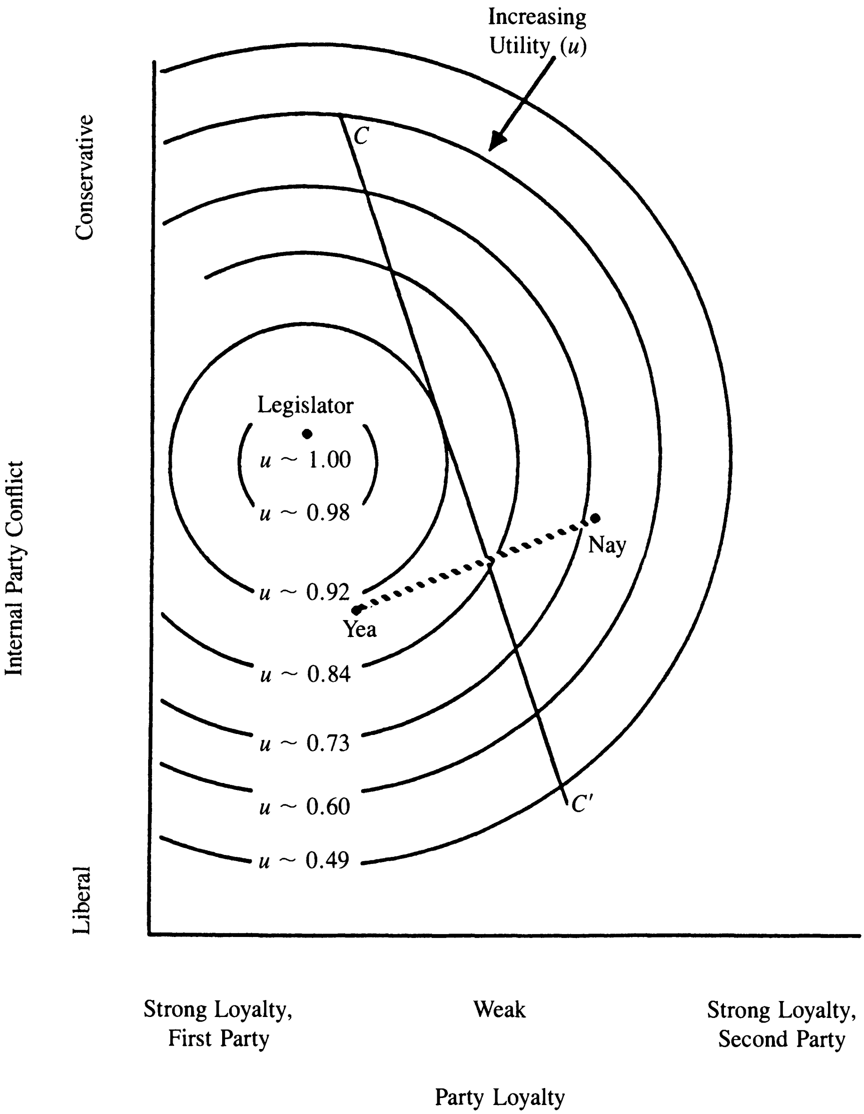
Expressions such as "liberal," "moderate," and "conservative" are part of the common language used to denote the political orientation of a member of Congress. Such labels are useful because they quickly furnish a rough guide to the positions a politician is likely to take on a wide variety of issues. A contemporary liberal, for example, is likely to support increasing the minimum wage, oppose aid to the Contras, oppose construction of MX missiles, support mandatory affirmative action programs, and support federal funding of health care programs. Indeed, this consistency is such that just knowing that a politician favors increasing the minimum wage is enough information to predict, with a fair degree of reliability, the politician's views on many seemingly unrelated issues.

This consistency or constraint (Converse 1964) of political opinions suggests that a politician's positions on a wide variety of issues can be summarized by a simple formal structure, where, as mentioned above, legislators are points and roll calls pairs of points in an Euclidean space. Insofar as a spatial model can capture congressional roll call voting, it is unnecessary to use a large number of dimensions. We find that one dimension captures most of the spatial information while a second dimension makes a marginal but important addition to the model. Adding more dimensions does not help us to understand congressional voting.

Although the dimensions are mathematical abstractions, the reader can think of one dimension as differentiating strong political party identifiers from weak ones. Except for very brief periods, the United States has always had a two-party political system. It is not surprising, therefore, that one dimension ranges from strong loyalty to one party (Democrat-Republican or Democrat) to weak loyalty to either party to strong loyalty to a second, competing, party (Federalist, Whig, or Republican). Another dimension differentiates "liberals" from "conservatives" within the two competing parties. The distinction between the two dimensions is a fine one. Loyalty to a political party and loyalty to an ideology have a similar behavioral implication of consistent, stable voting patterns. This is the reason that—especially during periods of stability—a one-dimensional model accounts for most voting in Congress.

In Figure 1 we show a two-dimensional example of a legislator's ideal point along with the points representing the yea and nay alternatives on a roll call vote. The circles centered on the legislator represent contours of the utility function employed in our study. If spatial proximity were the only consideration, the

Figure 1


Increasing
Utility (u)
C
Conservative
Internal Party Conflict
Liberal
Legislator
u ~ 1.00
u ~ 0.98
u ~ 0.92
Yea
Nay
u ~ 0.84
u ~ 0.73
u ~ 0.60
u ~ 0.49
C′
Strong Loyalty,
First Party
Weak
Strong Loyalty,
Second Party
Party Loyalty

legislator would clearly vote yea. Furthermore, consider the perpendicular bisector CC' of the line joining the yea and nay outcomes. This bisector, termed the cutting line, should pick out legislators who vote yea from those who vote nay. Those legislators whose ideal points are on the nay side of the bisector should vote nay.

We say "should" because the model will obviously not be successful in accounting for every individual decision. To allow for error, we employ the logit model. In this model it is only more likely that the legislator votes for the closer alternative. For a legislator whose ideal point falls on the cutting line, the probability of voting yea will be 0.5. Legislators with ideal points far from the cutting line will have a probability close to zero or one. Consequently, most "errors" (i.e., legislators who voted nay when on the yea side of the line and vice versa) in classifying actual data should fall close to the cutting line. We shall later use Figure 2 to return to the topic of error. The top panels of that figure use tokens to show the estimated ideal points of senators and the estimated cutting lines for two actual votes.

The dimensions of the space are related to policy areas considered by the legislature. We shall show that "1.5" dimensions can account for essentially all the behavior that can be accounted for with a simple spatial model that allows for probabilistic voting. We say "1.5" because, while a second dimension adds significantly in some Congresses, the second dimension is clearly less important than the first. That is, projecting all the diverse issues treated by Congress onto one dimension accounts for about 80% of the individual decisions and adding a second dimension adds only another 3%.

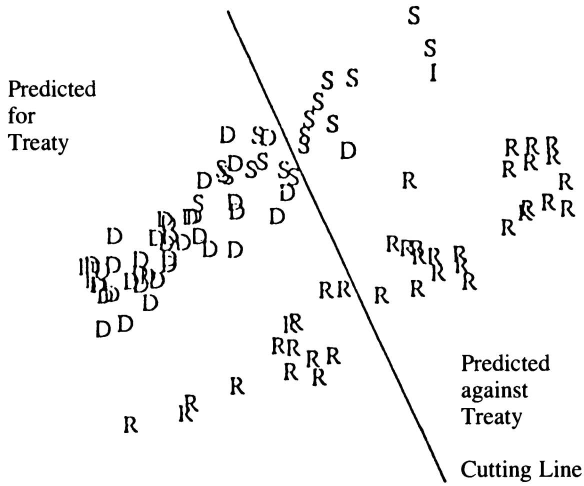
Of course, how specific issues map onto the dimensions may change over time. In the postwar period, an interpretation is fairly clear. The lineup on overriding Truman's 1947 veto of the Taft Hartley Act was almost a pure division along the first, left-to-right, dimension. Similarly, minimum wage and most other "economic" votes tend to line up on this dimension. In contrast, final passage of the 1964 civil rights act in the Senate, which was close to a straight South-North vote, was almost a pure division along the second, top-to-bottom, dimension. However, most of the roll calls designated by *Congressional Quarterly* as key votes in the postwar period, such as the Panama Canal treaty (see Figure 2), the Jackson amendment on SALT I, and a 15 May 1974 vote on school busing, tend to be "Conservative Coalition" votes, dividing the two parties internally at an angle of $-45°$ to the left-right dimension.[2]

Indeed, the results of our scaling algorithm readily admit to more than one interpretation. Defining the first dimension to be roughly along a 45° line in Figure 2, we can differentiate liberals (southwest quadrant) from conservatives

[2]For a set of figures like those in Figure 2 covering all *Congressional Quarterly* key Senate roll calls from 1945 to 1985, see Poole and Rosenthal 1989b. In addition, the Jackson amendment to the 1972 SALT I treaty is analyzed in Poole and Rosenthal (1988) and the Taft-Hartley, 1964 Civil Rights Act final passage, and busing votes are analyzed in Poole and Rosenthal (1989a).

Figure 2.A

Senate 95, 1977–78, Roll Call 755
Panama Canal Treaty



Senate 95, 1977–78, Roll Call 755
Panama Canal Treaty

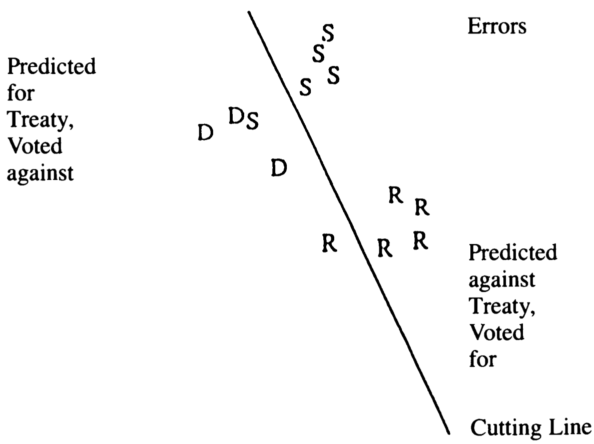
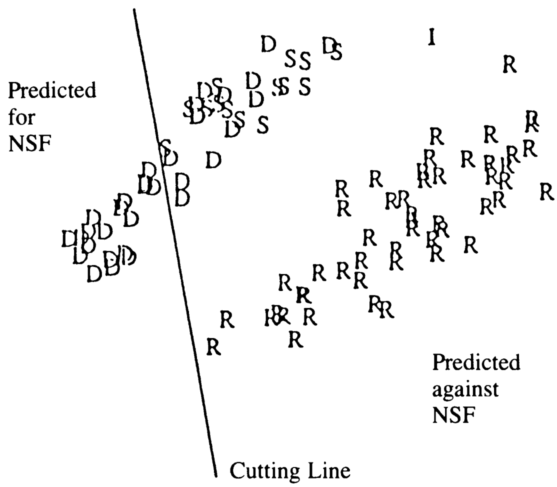
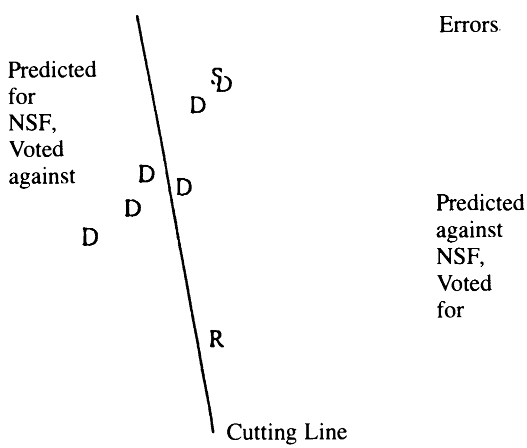


Figure 2.B

Senate 97, 1981–82, Roll Call 70
Restore NSF Funding



Senate 97, 1981–82, Roll Call 70
Restore NSF Funding



within each party. The orthogonal dimension is a party loyalty dimension. (The scaling algorithm does *not* use information about party.) This interpretation follows Poole and Daniels's (1985) two-dimensional analysis of interest group ratings.

A party dimension is present throughout nearly all of U.S. history, while an orthogonal dimension captures internal party divisions. Thus, the division between southern and northern Democrats after World War II finds a parallel in the division between southerners and northerners in both the Democratic and Whig parties in the 1840s and in the division between eastern and western Republicans from the 1870s through the 1930s.[3]

The fact that there is more than one substantive summary of our results is not troubling. Indeed, the "economic" versus "regional or social" and the "liberal-conservative" versus "party" interpretations both provide insight into the results. Moreover, the major finding of this study is that an *abstract* and parsimonious model can account for the vast bulk of roll call voting on a very wide variety of substantive issues. Our "1.5" departs from much of the previous literature which has either exogenously imposed a larger number of dimensions (e.g., Clausen 1973) or used methodologies that were inappropriate for the recovery of spatial voting (Morrison 1972).

Estimation Methodology

The data. Our estimation includes every recorded roll call between 1789 and 1985 except those with fewer than 2.5% of those voting supporting the minority side. For a given Congress (two-year period), we included every legislator who cast at least 25 votes. Pairs and announced votes were treated as actual votes. Observations with other forms of nonvoting (absent, excused) were not included in the analysis.

The spatial parameters. After thus excluding near unanimous roll calls and legislators with very few votes, the estimation requires Euclidean locations for 9,759 members of the House of Representatives, 1,714 senators, and 70,234 roll calls. In a two-dimensional setup, this requires estimating 303,882 parameters when spatial positions are invariant in time. This number is nevertheless small relative to the 10,428,617 observed choices.

Additional parameters are required to allow for spatial mobility. Some legislators clearly do not occupy stable positions in the space. For instance, there is the remarkable conversion of Senator Richard Schweiker (R-PA) from a weak liberal to a strong conservative after being tapped as Ronald Reagan's vice-

[3]The post–World War II split is aptly illustrated by the 88th Senate panel in Figure 7. The antebellum and postbellum divisions are shown graphically in Poole and Rosenthal (1989a). The postbellum Republican split continued into the 73d Senate as shown in Figure 7. Western Republicans, such as Borah (ID), Nye (ND), and Frazier (ND), tend to be at the top of the figure along with the two Progressives, La Follette (WI) and Norbeck (SD).

presidential running mate in 1976. To allow for spatial movement, we permitted all legislator coordinates to be polynomial functions of time. Nonetheless, we found great stability in legislator positions. A slight improvement in fit results from allowing linear trend;[4] higher order polynomials make virtually no additional contribution.

In the estimated model, the locations of legislators and roll calls are identified only up to a translation and rigid rotation. When we speak later of dynamics or realignments, the movement is always relative to any global translations or rotations. In contrast, the relative scale of the space is identified intertemporally. One cannot arbitrarily shrink or stretch the space over time. As a result, we can discuss changes in the degree of polarization of the political system: when legislators are spread further apart in the space, the system is more polarized.

Functional representation of the model. We use a specific functional model of choice to represent our hypothesis that roll call voting is sincere Euclidean voting subject to "error" induced by omitted factors. To eliminate notational baggage, we develop an s-dimensional model where legislator coordinates are quadratic functions of time. The extension to higher order polynomials is direct.

Legislators are indexed by i. At time t a legislator's Euclidean position is given by $(x_{i1t}, \ldots, x_{ikt}, \ldots, x_{ist})$ where

$$x_{ikt} = x_{ik}^0 + x_{ik}^1 t + x_{ik}^2 t^2, \qquad k = 1, 2, \ldots, s$$

Time is measured in terms of Congresses. Within a Congress, time is held constant. For each legislator serving in four or more Congresses, all three coefficients are estimated; for each legislator in three Congresses, the constant and linear coefficients are estimated; only x_{ik}^0 is estimated for legislators who served in only one or two Congresses.

Each roll call, indexed by j, is represented by two points in the space, one corresponding to an outcome identified with a yea (y) vote and the other to the nay (n) vote. The coordinates are written as z_{jyk} and z_{jnk}. (We omit time subscripts on the roll calls.)

If there was pure Euclidean voting, each legislator would vote yea if and only if his or her location were closer to the yea location than to the nay location.[5] In two dimensions, for example, this would be:

$$d_{ijy}^2 \equiv (x_{i1t} - z_{jy1})^2 + (x_{i2t} - z_{jy2})^2 < (x_{i1t} - z_{jn1})^2 + (x_{i2t} - z_{jn2})^2 \equiv d_{ijn}^2$$

[4] If we were to allow every legislator to have trend parameters, 23,146 additional parameters would be required for the two-dimensional model. A smaller number is used in practice, since no trend term is estimated for legislators serving in only one or two Congresses.

[5] One might be tempted to generalize our model to allow individuals to have salience weights for each dimension. However, the weights and the Euclidean coordinates cannot be identified simultaneously. That is, a large weight and a small coordinate would be equivalent to a small weight and a large coordinate.

Pure spatial voting ignores the "errors," or omitted variables, that influence voting. To allow for error, we assume that each legislator has a utility function given by:

$$U_i(z_{jl}) = u_{ijl} + \varepsilon_{ijl}, \qquad l = y, n$$

$$u_{ijl} = \beta \exp[-d_{ijl}^2/8]$$

where β is an additional parameter estimated in the analysis, ε is a "logit model" error which is independently distributed as the log of the inverse exponential, and 8 is an arbitrary scale factor.

The parameter β is essentially a signal-to-noise ratio. As β is increased, perfect spatial voting occurs—all probabilities approach zero or one. We have imposed a common β for all of U.S. history. The estimation was not substantially improved by allowing a distinct β for each Congress.

As a result of our choice of error distribution, we are able to write the probability of voting yea as:

$$\Pr(Yea)_{ij} = \frac{\exp[u_{ijy}]}{\exp[u_{ijy}] + \exp[u_{ijn}]} \qquad (1)$$

We chose the above specification for a number of reasons. First, the spatial utility function (u) is bell-shaped.[6] This allows for the possibility that individuals do not attribute great differences to distant alternatives. For example, Ted Kennedy might see little to choose from in a proposal that was at John Warner's ideal point rather than at Jesse Helms's. But Helms might see a very large difference between two such proposals.

Second, using a stochastic specification permits developing a likelihood function that is a function of the coordinates to be estimated. As this function is differentiable, it can be maximized by standard numerical methods. If we were

[6] We did not make utility linear in distance in order to preserve differentiability. We did not use quadratic utility because the roll call locations are not identified (although legislator locations and cutting lines are). While identification that occurs via choice of functional form can be tenuous, when data is generated, in a Monte Carlo experiment, by simulated behavior that corresponds to our posited model, we are able to recover the outcome coordinates. In practice, though, roll calls are likely to vary substantially as to level of error (β). This variation in error level and the fact that, empirically, most distances are in the concave region of our estimated utility function, make for very noisy estimation of the outcome coordinates. When the Monte Carlo work generates the data with variable β and D-NOMINATE is used assuming a common β, we still obtain accurate estimates of legislator coordinates and cutting lines. Moreover, estimation of legislator coordinates and of cutting lines may be very robust to the functional form used in the utility function. Ladha (1987) used a one-dimensional quadratic utility specification and obtained legislator coordinates and cutting lines very similar to our one-dimensional estimates for recent Senates. Ladha also shows that using a "probit" rather than a "logit" model for the errors makes little difference to the results. To sum up, we believe we have a very robust procedure for recovery of legislature coordinates and midpoints. More details are available in the Appendix.

concerned solely with *ordinal* scaling, we could eschew this approach in one-dimensional problems. That is, we could start with a configuration of legislators and then find a midpoint for each roll call that minimized classification errors. Next, the midpoints could be held constant, and classification errors could be further reduced by reordering the legislators. This process can then be iterated to convergence. Such a procedure indeed makes fewer classification errors than ours, which, in maximizing a likelihood, heavily weights errors that correspond to low-probability choices. Ordinal scaling of this form, however, is wholly impractical in more than one dimension.

Finally, by using the logit form of error, we can calculate the probabilities in the closed form (equation 1). The standard alternative to our nonlinear logit model would be probit. As this involves numerical integration, more time is needed to estimate the model. Further details about the estimation procedure appear in the Appendix.

3. Spatial Structure of Congressional Voting

Let us begin by showing typical but recent "snapshots" of the voting model. Figure 2 shows all voting members of the Senate in their estimated positions and the cutting lines for two specific votes, the Panama Canal treaty vote on 18 April 1978 and a proposal to restore funding for the National Science Foundation on 2 April 1981. A *D* token represents northern Democrats; *S*, southern Democrats;[7] and *R*, Republicans (the one *I* is Harry Byrd [VA]). Similar positions of senators produced overstriking. However, an *R* token is always overstruck by another *R*, and *S*'s and *D*'s are always overstruck by other *S*'s and *D*'s. The bottom part of each panel shows only those senators who were, given their location relative to the cutting line, "errors." As explained above, the probability of an error should be greatest for senators closest to the cutting line. The data conform to the expected pattern; errors in voting are far more likely for senators close to the cutting line than for those who are distant. When senators' Euclidean positions provide a clear indication of which side they should join, forces not captured by our simple structure are rarely strong enough to produce a vote that is inconsistent with the spatial model.

These two roll calls are quite representative of post–World War II voting in Congress. Typically, roll calls divide at least one of the two parties and have estimated cutting lines roughly parallel to the two shown in Figure 2.[8] The ten-

[7] Southern Democrats are those from the 11 states of the Confederacy, Kentucky, and Oklahoma.

[8] We performed significance tests on our estimated roll call coordinates. We tested both the null hypothesis that all roll call coordinates were zero and the null hypothesis that the vote "split the parties." To carry out the tests, we began by estimating the legislator coordinates and β using sets of roll calls that did *not* include the vote in question. This procedure eliminates all the statistical problems discussed in the Appendix. The Panama Canal Treaty vote was the 755th in our estimation for the 95th Senate. For the test we used the first 754 roll calls in that Senate to estimate the legislators

dency for cutting lines to be parallel explains why a one-dimensional model provides a useful approximation that accounts for most voting decisions.

Returning to the question of error: How general is the pattern shown by the snapshots? A straightforward method of fit is the percentage of correct classifications across all roll calls. The classification results for the two-century history of both Houses of Congress are shown in Table 1. The table reports classification both for all roll calls in the estimation and for close roll calls where the minority got over 40% of the vote cast. With a two-dimensional model, classification is better than 80% for close votes as well as all votes.

It can be seen that a reasonable fit is obtained from a one-dimensional model where each legislator's position is constant throughout his or her career. On the other hand, there is considerable improvement—about three percentage points—from adding a second dimension. Allowing for a linear trend in legislator positions adds another percentage point. (That we get less of a boost in the percentages from the time trend than the dimensions is expected. When we add a time trend, we add only one parameter per legislator per dimension. In contrast, adding a dimension adds two parameters per roll call as well as additional legislator parameters. Since roll calls outnumber legislators by over five to one, it is not surprising that classification shows more improvement when we increase the dimensionality of the space than when we increase the order of the time polynomial.)

Introducing more parameters in a dynamic spatial model—through extra dimensions or higher order polynomials—does not appreciably add to our un-

---

and β. The NSF vote was the 70th in the 97th Senate. We used the last 896 votes in that Senate. (To avoid using two hours of supercomputer time per significance test, we did not rerun the full dynamic estimation excluding only the roll call in question. The estimated legislature configurations from these subsets of roll calls are virtually identical to those obtained from the dynamic estimation.) Treating β and the x's estimated in this first step as fixed parameters, we then used the roll call of interest to estimate the two-dimensional roll call coordinates.

The null hypothesis that all coordinates are zero implies that the yea and nay outcomes occupy identical locations and thus that legislators flip fair coins on the vote. In other words, the hypothesis that all coordinates are zero is equivalent to the more general hypothesis that $z_{jy1} = z_{jn1}$ and $z_{jy2} = z_{jn2}$. Under the null hypothesis the log-likelihood is simply $L(H_0) = N_j \ln(1/2)$, where N_j is the actual number voting or paired on roll call j. The log-likelihood of the alternative, denoted, $L(H_a)$, is computed by D-NOMINATE. For the Panama Canal Treaty, we find $L(H_0) = -69.31$, $L(H_a) = -14.06$. Using the standard likelihood-ratio test, we obtain $\chi^2 = 110.32$ with 4 d.f. (since there are four roll call parameters in two dimensions), and $p < 10^{-22}$. Similarly, for the NSF vote, we have $\chi^2 = 110.24$.

To test the null hypothesis that the vote split the parties, we first created a pseudo–roll call in which all Democrats voted yea and all Republicans and Harry Byrd voted nay. We then estimated the outcome coordinates that maximized the log-likelihood for this roll call. The null hypothesis was that the coordinates for the Panama Canal (NSF) vote were those of the party line pseudo–roll call. The chi-square statistic calculated from the log-likelihoods was 122.48 (170.74), again extremely significant for 4 d.f.

Table 1. Classification Percentages and Geometric Mean Probabilities

	House			Senate		
Degree of	Number of Dimensions			Number of Dimensions		
Polynomial	1	2	3	1	2	3
Classification percentage, all scaled votes:						
Constant	82.7[a]	84.4	84.9	80.0	83.6	84.1
Linear	83.0	85.2	—[c]	81.3	84.5	85.5
Quadratic	83.1	85.3	—	81.5	84.8	85.1
Cubic	83.2	85.4	—	81.6	85.0	86.1
Classification percentages, votes with at least 40% minority:						
Constant	80.5[b]	82.9	83.7	78.9	82.7	83.4
Linear	80.9	83.8	—	79.4	83.6	84.8
Quadratic	81.0	83.9	—	79.7	83.8	85.1
Cubic	81.1	84.1	—	79.8	84.0	85.3
Geometric mean probability, all scaled votes:						
Constant	.678	.696	.707	.660	.692	.700
Linear	.682	.709	—	.666	.704	.716
Quadratic	.684	.712	—	.668	.708	.721
Cubic	.684	.714	—	.670	.708	.725

[a]The percentage of correct classifications on all roll calls that were included in the scalings (i.e., those with at least 2.5% or better on the minority side).

[b]The percentage of correct classifications on all roll calls with at least 40% or better on the minority side.

[c]Higher polynomial models for three dimensions were not estimated because of computer time considerations.

derstanding of the political process. Adding extra spatial parameters results in only a very marginal increase in our ability to account for voting decisions. For example, consider adding to the two-dimensional linear model in the Senate. Allowing for a quadratic term in the time polynomial improves classification only by 0.3% at a cost of 1,456 additional parameters (two dimensions × 728 senators serving in four or more Congresses). Allowing for a third dimension improves classification by only 1.0% at a cost of 77,479 more parameters (two more per roll call and one or two additional parameters per legislator). Allowing for both, generates an improvement of only 1.1%. Thus, the important regularity we have found is that somewhat over 80% of all individual decisions can be accounted for by a two-dimensional model where individual positions are temporally stable. This regularity is an important pattern, but the pattern does not arise from a well-specified theoretical model that would fix the dimensionality of the space. The decisions the spatial model cannot account for are likely to reflect either very specific sets of constituency and other interests or logrolling and other forms of strategic voting[9] that lie outside the paradigm that forms the basis for our statistical estimation.

The Dimensionality of Congressional Voting

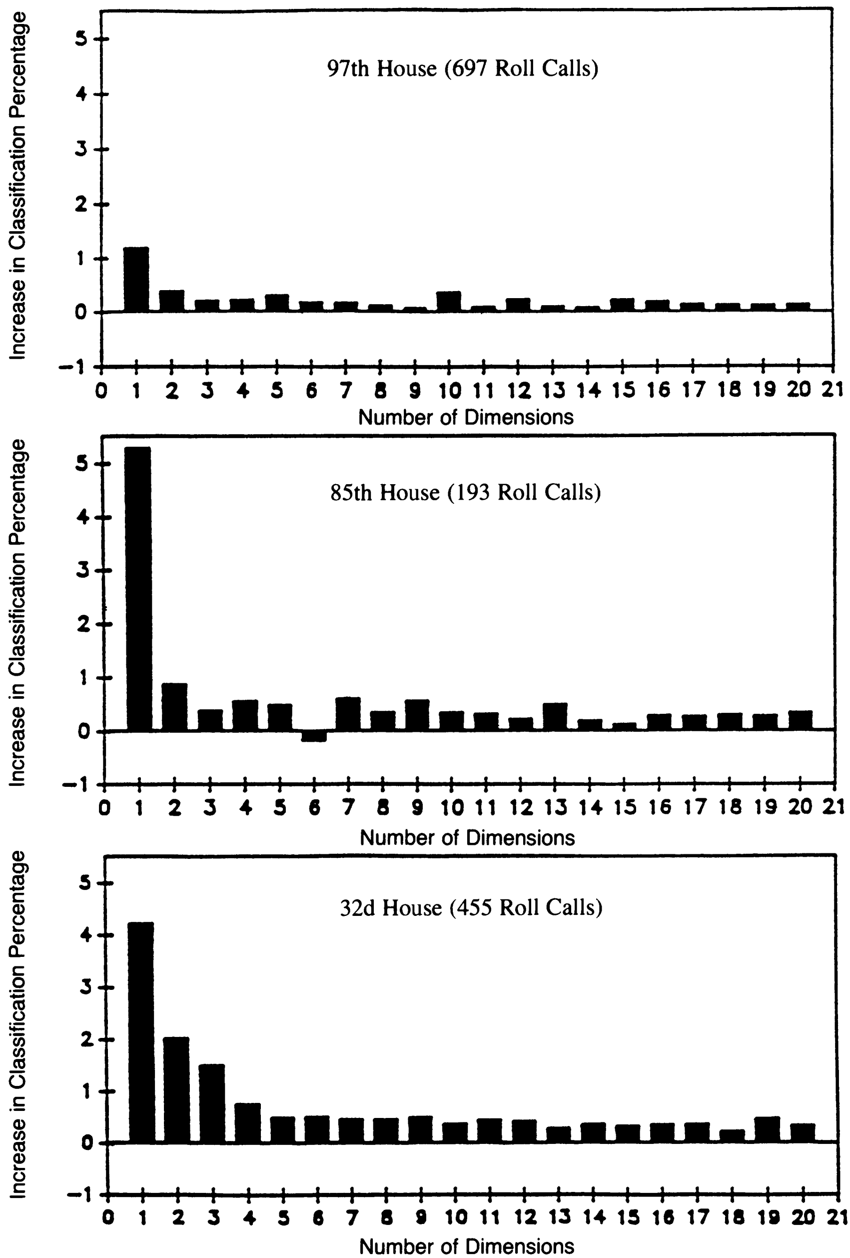
Since low dimensionality is an important and, to many, unexpected empirical result, we shall discuss several different sets of supporting evidence for it. First, for three Houses, we show the increments to the percentage classified correctly when D-NOMINATE is estimated with as many as 21 dimensions. Second, we evaluate the classification ability of the second dimension from the two dimensions with linear trend estimation and compare this to the first dimension. Third, we compare our ability to classify with a one-dimensional model with what might be expected if legislators and roll calls were distributed within an s-dimensional sphere. Fourth, we show that the results of D-NOMINATE are reasonably stable when the algorithm is applied to subsets of roll calls that have been defined in terms of substantive content. Fifth, we show that an alternative measure of fit, the geometric mean probability of the observed choices, gives similar results to those based on classification percentages. Sixth, since dimensionality may depend on the agenda, we compare the model's performance with measures of the diversity of the agenda. Seventh, we ask which issues in U.S. history led to an important role for a second dimension.

1. Models of high dimensionality. As our first check on dimensionality of our dynamic models, we selected three Houses and estimated the static model up to 50 dimensions. The 32d House (1851–52) was chosen because it was one of the worst-fitting Houses in two dimensions and thus was a good candidate to exhibit high dimensionality. The 85th House (1957–58) was chosen because it was analyzed with other methods by Weisberg (1968). In addition, the 85th House is part of a period when the two-dimensional linear model clearly dominates the one-dimensional linear model. Finally, the 97th House (1981–82) is included because it appears that roll call voting became nearly unidimensional at the end of the time series.

Figure 3 displays the classification gains for the 2d through the 21st dimensions for each of the three Houses. The classification percentage for the first dimension was 70.2 for the 32d House, 78.0 for the 85th, and 84.1 for the 97th. The bars in the figure indicate how much the corresponding dimension adds to the total of correctly classified. (The horizontal axis is labeled such that 3 corresponds to the fourth dimension, etc.) Note that the bars do not drop off

[9] We can show, however, that estimates of the legislator locations will not be biased by strategic voting over binary amendment agendas (Ordeshook 1986, 281–84) in a complete information setting.

Figure 3. Classification Gain by Dimensionality



smoothly—in fact, on one occasion the bar is negative—because the algorithm is maximizing likelihood not classification.

The 97th House is at most two dimensional with the second dimension being very weak. After two dimensions the added classifications are minuscule. There is a clear pattern of noise fitting beyond two dimensions. In contrast to the 97th, the 85th House is strongly two dimensional, but again there is little evidence for additional dimensions. While the 32d does show evidence for up to four dimensions, even four dimensions account for only 78% of the decisions. These results argue that either voting is accounted for by a low-dimensional spatial model, or it is, in effect, spatially chaotic. There appears to be no middle ground.

2. *The relative importance of the second dimension.* Although the evidence presented in Table 1 and Figure 3 suggests a marginal role for at most a second dimension, and a weak one at that, it is important to evaluate the second dimension by other than its marginal impact. Specifically, Koford (1989) argues that a one-dimensional model will provide a good fit even when spaces have higher dimensionality. For example, in a truly two-dimensional space, one dimension will have some success at classifying any vote that is not strictly orthogonal to the dimension. As a result, the marginal increases in fit on the order of 3% may understate the importance of the second dimension.

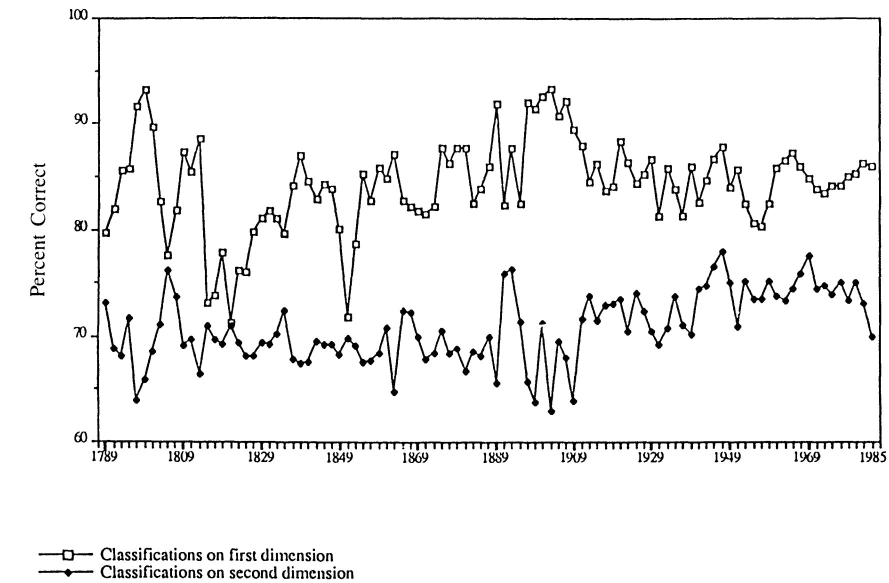
The natural question, then, is how well does the second dimension do in classifying by itself. To study this, we took the second-dimension legislator coordinates from our preferred model, two dimensions with linear trend, and, for each roll call, found a cutpoint that minimized classification errors. We used the minimum errors to compute overall classification percentages. We made the same computation for the first dimension.

The results of these computations for the House[10] are shown in Figure 4. The averages of the 99 biennial figures show the first dimension correctly classifies 84.3% of the votes, but the second dimension accounts for only 70.8%. The 70.8% is particularly unimpressive given that predicting by the marginals would lead to 66.7%. If the two dimensions were indeed of equal importance, then in some Congresses dimension two might do better than dimension one. But in all 99 Houses, one did better. The Senate results are slightly weaker—83.8% for one dimension versus 73.6% for two. The marginals here were 66.1. In addition, two does better in Senates 2, 17, and 18. But clearly the second dimension is a second fiddle.

3. *How well should one dimension classify?* In addition to this empirical comparison between our two dimensions, following Koford (1989), we consider

[10]To save space we do not present a full set of results for both Houses of Congress. Except where noted, results are similar for the two Houses.

Figure 4



the issue of unidimensional fit theoretically. Specifically, we assume an n-dimensional uniform spherical distribution of ideal points and consider the projection of these ideal points onto one dimension under the conditions of errorless spatial voting in the n-dimensional space. As for the distribution of roll call cutting lines or separating hyperplanes, note that each roll call hyperplane can be represented as tangent to a sphere of radius r that has a common center with the ideal point sphere. For fixed r, we assume that the distribution of tangency points is uniform on the sphere. As for the distribution of r, we make use of the fact that, with errorless spatial voting, there is a one-to-one relationship between r and the expected split y (y% in the majority, $100 - y$ in the minority) on the roll call. We use the empirical distribution of y (i.e., the historical distribution of the marginals) to define the distribution of r. Given the empirically generated distribution of r and the assumed uniform distributions of tangency points and ideal points, one can calculate the percentage of correct classifications that would be made by a one-dimensional projection.

For $s = 2$, we can calculate the exact percentage correct for errorless spatial voting. For the empirical distribution of splits over all roll calls included in our analysis, 78.9 would be classified correctly in both the House and Senate if legislators and roll calls were spherically uniform for $s = 2$. Thus, it might be the case that the 80% correct classification that we obtain in the one-dimensional constant model might arise from *perfect* two-dimensional spatial voting. However, with two dimensions, we correctly classify 83.5% (Table 1). Thus, a better benchmark model would be two-dimensional voting with an error rate of 16.5%. The 83.5% voting correctly in two dimensions would be projected correctly with probability of 0.789. The 16.5% voting incorrectly would have their error "corrected" by an incorrect projection with probability 0.211. Therefore, a one-dimensional projection would correctly classify only 83.5(0.789) + 16.5(0.211) = 69.4% of the individual votes.

For $s \geqslant 3$, we conducted simulations. We had 5,000 voters randomly drawn within the unit sphere vote perfectly on 900 randomly drawn roll calls. Using the empirical distribution of splits, Table 2 shows how the percentage correctly classified declines with s. As the dimensionality increases, the percentage correct for a one-dimensional projection approaches the average value of y or the percentage correct for the "Majority" model. The table indicates that for a one-dimensional model to classify at the 80% level, the underlying distributions of ideal points and cutting lines must be nearly one-dimensional rather than spread uniformly about some space of even modestly higher dimensionality.

4. Do different issues give different scales? In contrast to our emphasis on low dimensionality, Clausen (1973) has argued that there are five "dimensions" to congressional voting represented by the issue areas of government management, social welfare, agriculture, civil liberties, and foreign and defense policy. We have coded every House roll call from 1789 to 1985 in terms of these five categories and, for completeness, a sixth category termed miscellaneous. If the

Table 2. Percentage Correct Classifications, One-Dimensional Fit to *S*-Dimensional Space

House	Senate	"True" Dimensionality
100.0	100.0	1[a]
78.9	78.9	2[a]
74.8	74.7	3
73.2	73.0	4
71.2	71.0	5
70.9	70.8	6
69.9	69.8	7
69.9	69.7	8
69.2	69.1	9
65.9	65.9	majority model

Note: Legislators uniformly distributed through an s-dimensional sphere. Roll call lines distributed to reproduce marginals found in congressional data.

[a]Calculation based on closed form expression.

issues are really distinct dimensions, we ought to get sharp differences in legislator coordinates when the issues are scaled separately.

To conduct this experiment of separate scalings, we chose the 95th House because it had the largest number of roll call votes (1,540). There were 714 government management votes, 286 social welfare votes, 311 foreign and defense votes, and, to have enough votes for scaling, 229 in a residual set that combined agriculture, civil liberties, and miscellaneous. We then ran one- and two-dimensional (static) D-NOMINATE on each of these four clusters of votes. Because it is difficult to compare coordinates from two-dimensional scalings directly, we based our comparisons on correlations between all unique pairwise distances among legislators.[11]

Correlations between the management, welfare, and residual categories for one-dimensional scalings are, as shown in Table 3, all high, around 0.9. Correlations between the foreign and defense policy category and the other three were somewhat lower, in the 0.7 to 0.8 range.[12] As a whole the results hardly suggest that each of these clusterings of substantive issues generates a separate spatial dimension.

When the same subsets of votes are scaled separately in two dimensions, the correlations are somewhat *lower* than they are in one dimension (again see Table 3). This result is not surprising. The 95th House had nearly unidimensional voting. From the D-NOMINATE unidimensional scaling with linear trend that was applied to the whole data set, we find that one dimension correctly classifies 83% of the votes in each of the four categories. With two dimensions the percentages increase only to 84% for social welfare and foreign and defense and to 85% for the other two categories.[13] Moving from one to two dimensions doubles the number of estimated parameters with only slight increases in classification ability. In breaking down the roll calls into four categories and estimating separately, the number of legislator parameters is effectively quadrupled. With a further doubling of all parameters, by moving from one to two dimensions, one is likely to be fitting idiosyncratic "noise" in the data. The fit to the noise weakens the underlying strong correlations between legislator positions. We also note that the spirit of Clausen's (1973) work suggests that each category should be scaled in one dimension only. In summary, our breakdown of the 95th House in terms

[11] See the Appendix for discussion of why distances rather than coordinates are analyzed when the analysis is not limited to unidimensional spaces.

[12] It is difficult to pin these lower correlations on a specific item. In the 95th Congress, foreign and defense policy votes included 14 on CIA, spying, or intelligence; 11 on South Africa or Rhodesia; eight on military pensions or veterans benefits; seven on the Panama Canal; seven on the B-1 bomber; five on arms control; and five on the United Nations.

[13] Fits were not as good for the 49 votes in the agriculture category, with 75% correct classifications in one dimension and 80% in two. Because of the small number of votes, agriculture had to be placed in the residual category.

Table 3. Interpoint Distance Correlations, Clausen Category Scalings, 95th House

Category	1	2	3	4
1. Government management	1.0	.914[a]	.796	.908
2. Social welfare	.883[b]	1.0	.765	.881
3. Foreign and defense policy	.770	.654	1.0	.724
4. Miscellaneous policy, civil liberties, and agriculture	.832	.746	.613	1.0

[a] Numbers above diagonal are correlations from one-dimensional scalings.

[b] Numbers below diagonal are correlations from two-dimensional scalings.

of Clausen categories indicates that the categories represent highly related, not distinct, dimensions.

5. *Evaluation by geometric mean probability.* In addition to computing classification percentages, the model may be evaluated by an alternative method that gives more weight to errors that are far from the cutting line than to errors close to the cutting line—a vote by Edward Kennedy (D-MA) to confirm Judge Robert Bork to the Supreme Court would be a more serious error than a similar decision by Sam Nunn (D-GA). Such a measure is the geometric mean probability of the actual choices, given by:

$$gmp = \exp(\text{log-likelihood of observed choices}/N)$$

where N is the total number of choices.

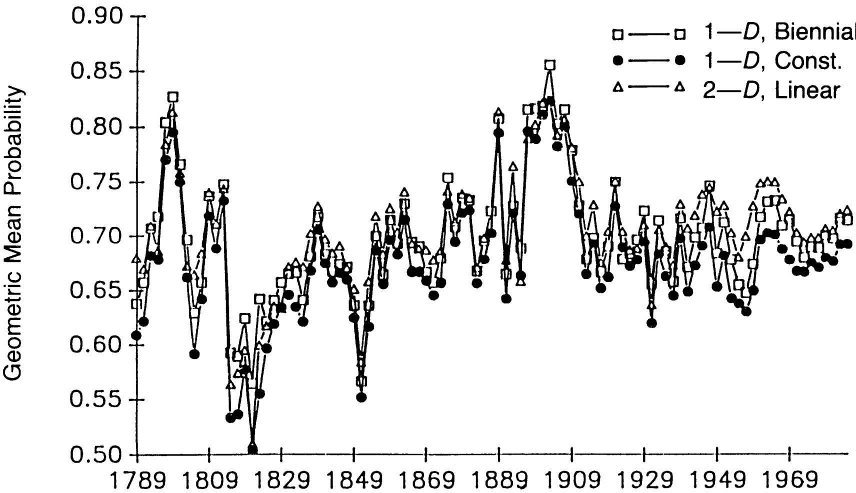
Summary *gmp*s for the various estimations are presented in Table 1. The pattern matches that found for the classification percentages—little is gained by going beyond two dimensions or a linear trend.

In Figure 5, we plot the *gmp* for each House for the following models: (1) a two-dimensional model with legislator positions constrained to a constant plus linear trend; (2) a one-dimensional model with legislator positions constrained to a constant; (3) a one-dimensional model that is estimated separately for each of the first 99 Congresses. Note that in this model there is no constraint on how legislator positions vary from Congress to Congress.

Motivation for the third model came from a recent argument by Macdonald and Rabinowitz (1987) that American political conflict is basically one dimensional within the time span of any one Congress but that the dimension of conflict evolves slowly over time.[14]

[14] Macdonald and Rabinowitz (1987) support their hypothesis on the basis of an analysis that combines our model 3 results with time series of state returns in congressional and presidential elections.

Figure 5. Geometric Mean Probability, House



The Macdonald-Rabinowitz hypothesis, unidimensionality within any Congress, is sustained for the entire period following the Civil War. As shown in Figure 5, the *gmp* for model 3 has not fallen below 0.64 since 1853–54, oscillating in the 0.64 to 0.74 range with the exception of the very high *gmp*s that occurred in the period of strong party leadership around the turn of the century. The hypothesis of slow evolution is supported by our result that voting patterns can be largely captured by a one-dimensional model where individual positions are constant in time. In Figure 5 the curve for model 2 closely tracks the curve for model 3. Model 1 provides a slightly better tracking.[15] Because political change is slow, roll call voting reflects changes in the substance of American politics either as a trend for an individual legislator in a two-dimensional space

[15] Because of the vast amount of data involved, significance tests for statements made concerning Figure 5 would be of little value. For example, the lowest *gmp* for our biennial scalings is 0.564 for the 17th Congress. Assume the null hypothesis were all z's equal to 0, that is all probabilities 0.5. Performing a likelihood-ratio test for that scaling using the approximation that $\sqrt{2\chi^2} - \sqrt{2d.f. - 1}$ is normal for large $d.f.$ yields a Z-statistic of 63.78. Even if the *gmp* was only 0.51, the Z-statistic would still be a hefty 9.65. Of course, for the 17th Congress, the one dimensional constant model is not very much better than 0.5 (*gmp* 0.504). To carry out the appropriate likelihood-ratio test, one would have to constrain the z's for the 17th Congress only to be zero and reestimate the full dynamic model. This would be a waste of computer resources. After all, the difference between 0.504 and 0.500 is of no *substantive* importance. In general, given our very large number of observations, we focus on substance rather than statistical significance.

or as replacement of some legislators by others with different positions in the space.[16]

6. The agenda and dimensionality. One basis for the Macdonald-Rabinowitz argument would be that short-term coalition arrangements enforce a logroll across issues that generates voting patterns consistent with a unidimensional spatial model. Another potential consideration is that short-run unidimensionality may reflect the fact that, in any two-year period, Congress must place some restriction on the issues that can be given time for consideration. If this is so, Congresses that consider a diversity of issues should be less unidimensional.

To test this diversity hypothesis, in at least a crude way, we computed, for each of the 99 Congresses, the Herfindahl concentration index[17] for the six Clausen (1973) categories. We also coded all House roll calls using a finer-grained set of 13 categories developed by Peltzman (1984). The Herfindahl index was also computed for the Peltzman categories. The indices validate, but very weakly ($R = .362$).[18] Just over half the variation in the index for Peltzman categories is explained by trend ($R = -.709$), as government has expanded over time. The index for the Clausen categories is more weakly related to trend ($R = -.405$). Both indices are significantly correlated with the geometric means from the two-dimensional, linear trend model but in a counterhypothesis direction. As the roll call set becomes more diverse, the model fits better ($R = -.302$ for Clausen, $-.369$ for Peltzman). The result is undoubtedly spurious. The worst-fitting years occur early in the time series while the agenda has become more diverse over time. Indeed, diversity of the agenda, at least as measured by these indices, is not significantly related to the ability (difference in geometric means) of the two-dimensional model to *improve* over the one dimensional linear model ($R = -.089$ for Peltzman, $-.158$ for Clausen).

One reason for these basically negative results for the diversity hypothesis is that the indices have exhibited little variation. For Congresses 40–99, the index for Clausen averaged 0.355 with a standard deviation of 0.062; for Peltzman, the average is 0.090, and the standard deviation 0.020. In the last 100 years, Congress has had a full and wide-ranging agenda. Low-dimensional voting has not occurred simply because votes are restricted to a narrow topical area.

[16] Indeed, the close fit between the biennial scalings in Figure 5 and the dynamic scaling shows that our results are not strongly influenced by the time period chosen. For example, if we had scaled only the twentieth century, the results for that period would be almost identical to those obtained by scaling all 99 Congresses together.

[17] If ρ_a is the proportion of roll calls in category a, the index H is given by $H = \Sigma_a \rho_a^2$. Index H equals 1.0 if all the votes are in one category. For the Clausen categories, H would reach a minimum of 1/6 if the roll calls split evenly among the six categories.

[18] For descriptive purposes, note that every R reported in this paragraph above 0.2 in magnitude is "significant" at 0.002 or better while those below 0.2 are not significant, even at 0.1.

Table 4. Clausen Categories with Second Dimension PRE Increases at Least 0.1

PRE > 0.5			PRE × 100		PRE < 0.5			PRE × 100	
House	Categ.	# Votes	1-D	2-D	House	Categ.	# Votes	1-D	2-D
1	Mgt	92	41	56	9	Mgt	73	16	35
	Civil	11	26	65		Civil	15	28	40
2	Misc	26	41	58	10	Civil	13	29	42
9	F&D	54	39	58	11	Civil	16	30	41
	Misc	12	35	52	12	Misc	34	33	46
10	F&D	150	45	56	15	Mgt	49	17	29
18	Mgt	68	43	54	15	F&D	25	11	26
23	Mgt	180	45	59	17	Mgt	49	11	24
24	Mgt	190	42	54	19	Civil	12	19	30
	Civil	65	37	54	22	F&D	30	24	45
25	Civil	46	46	64	24	Welfr	11	07	45
26	Civil	33	46	63	30	Civil	30	36	47
28	Civil	23	40	64	32	Mgt	290	22	36
31	Civil	17	44	55		Civil	17	26	40
33	Mgt	440	42	55	33	Agr	10	24	38
53	Mgt	140	46	61	40	Welfr	17	26	39
64	Welfr	21	20	50	41	Mgt	370	36	47
76	Welfr	22	32	52	42	Mgt	280	38	49
	Civil	11	23	56	43	Mgt	220	26	37
77	Civil	14	21	61	51	Agr	13	34	44
	F&D	40	41	53	53	Agr	12	20	47
78	Civil	12	32	61	62	Welfr	16	03	27
	F&D	31	45	56	63	Welfr	25	30	46
79	Welfr	18	40	59	65	Agr	20	27	43
	Misc	20	31	66	66	Welfr	27	09	38
81	Welfr	45	59	70	67	Welfr	14	18	37
	Civil	15	94	64	69	Agr	11	22	34
	Misc	28	26	75	75	Welfr	16	20	45

7. *Issues and the second dimension.* There have been relatively few issues that have consistently sparked a second dimension in spatial terms. We demonstrate this by considering the proportional reduction in error (PRE) over the marginals [PRE = 1 − (D-NOMINATE errors)/(number voting on minority side)] within each of the Clausen categories. We use PRE to control for differences in marginals across categories. We computed the PRE for the linear models in one and two dimensions. We obtained PREs for six categories × 99 Houses. We then filtered these into a subset that contains only those category-Congress pairs that were (1) based on at least 10 roll calls, (2) had a two-dimensional PRE of at least 0.5, and (3) had an increase in the PRE of at least 0.1 between one and two

Table 4 (*continued*)

PRE > 0.5			PRE × 100		PRE < 0.5			PRE × 100	
House	Categ.	# Votes	1-D	2-D	House	Categ.	# Votes	1-D	2-D
82	Welfr	19	52	67	77	Agr	16	24	38
	Agr	14	45	59	80	Welfr	19	37	47
	F&D	34	50	61	84	F&D	16	30	49
	Misc	10	58	73	86	F&D	24	12	48
83	Welfr	18	29	63	89	Agr	17	37	48
	F&D	18	33	52	90	Agr	29	27	38
84	Welfr	10	45	65	91	Agr	12	17	39
	Agr	12	60	70	92	Agr	24	31	50
85	Mgt	110	41	53	93	Agr	49	22	45
	Agr	12	31	51	94	Agr	40	22	35
	F&D	29	24	53	95	Agr	49	21	37
86	Agr	14	40	57	97	Agr	23	26	41
	Civil	10	21	71					
87	Welfr	32	57	67					
	Agr	20	47	58					
	F&D	36	36	57					
88	Welfr	36	59	69					
	Agr	14	53	64					
	Civil	11	35	65					
89	Civil	23	49	70					
	Misc	24	58	69					
90	Civil	30	37	60					
91	Welfr	65	53	64					
	F&D	63	44	55					
	Misc	26	40	52					
94	Misc	48	44	55					

dimensions.[19] In other words, we found sets of roll calls that were highly spatial and where the second dimension made an important difference. These appear in Table 4.

It can be seen that the second dimension was important in only six of the first 23 Houses. It appeared sporadically in different areas. In Houses 24–31, the second dimension emerges in five Houses, and civil liberties is the key. Note for further reference, however, that, after the Compromise of 1850, civil liberties

[19]The substantive results are not sensitive to the values used in the filter. In particular, the 10 roll call requirement is sufficiently low that the analysis is not affected by the lower number of total roll calls in earlier years.

(essentially slavery) vanishes as an issue that is accommodated by introducing a second dimension. In fact, from the 32d through the 75th Congresses, there are only three occasions (1853–54, 1893–94, and 1915–16) when the second dimension made a key difference, each time only in one issue area. The realignments of the 1890s and 1930s were largely accommodated not by a shift in the space but by the infusion of new blood (Republicans in the 1890s and Democrats in the 1930s) in the existing space.

In contrast to the first 75 Houses, the second dimension was systematically important in Houses 76 (1939–40) through 91 (1969–70). Except for the 80th Congress, every House in this period appears in the table. Civil liberties is the most frequent category. (The 80th is eliminated because there were only eight civil liberties votes; however, their PRE was 0.62 in two dimensions, an increase of 0.26). Moreover, in only one case (the 90th Congress) did the second dimension matter in fewer than two issue areas. In other words, when the space became strongly two dimensional, it became consistently so across a wide variety of topics. As stated earlier, the dimensions are not so much defined by topics as they are abstractions capable of capturing voting across a wide set of topics. Finally, from Congress 92 out, the second dimension appears only once when the PRE is over 0.5, reaffirming our earlier conclusion that the House is currently virtually unidimensional.

Table 4 also shows groups of roll calls where the second dimension increases PRE by 0.1, but the total PRE is below 0.5. In other words, here we have roll calls where fit is improved but where a spatial model does not explain most of the variance. The first 50 Houses account for a higher fraction of the entries here, reflecting the poor fits in some years. For example, in our two worst-fitting Houses, the 17th and 32d, a second dimension helps the category with the most votes, government management, but the PRE remains low.

Agriculture, particularly since the 89th House, appears to be the one category that is not captured by a spatial model. (In the first 60 Houses, there were only three with at least 10 votes in the category.) Agriculture seems to have gradually fallen out of the spatial framework. Immediately after World War II, agriculture had one-dimensional PREs over 0.6. In the 1950s and early 1960s, the one-dimensional PREs fell but agriculture still fit spatially via the second dimension. In the 1970s and 1980s, voting on agriculture has been largely nonspatial. Indeed, from the 88th House forward, agriculture has consistently the lowest PRE of the six categories (see Poole and Daniels 1985 for similar conclusions based on interest group scalings).

We repeated the above analysis, using the same filters, for the Peltzman coding. The results were quite similar. Results for Peltzman's domestic policy code were like those for the Clausen civil liberties code. As with the Clausen codes, voting in a variety of other areas also scaled on the second dimension in Houses 76–91. In summary, our analysis of PREs reinforces our findings of low dimensionality. Particularly in recent times, when a second dimension has an impact on fit, the impact is on the one area, agriculture, that is essentially nonspatial.[20]

Spatial Stability

We have seen that, to whatever extent roll call voting can be captured by a spatial model, a low-dimensional model (say, "1.5" dimensional) suffices. But what of the temporal stability of the model. We address three issues here: (1) Does the model consistently fit the data in time? (2) Is the major, first dimension stable in time? (3) Are individual positions stable in time?

1. Stability of fit. Inspection of Figure 5 shows that there are only two occasions when spatial models fit poorly. Poor fit occurs between 1815 and 1825, when the Federalist party collapsed and gave way to the Era of Good Feelings, and in the early 1850s, when the destabilization induced by the conflict over slavery was marked by the collapse of the Whigs. Thus, in periods of political stability, roll call voting can be described by a low-dimensional spatial model. In contrast, voting is largely chaotic in unstable periods when a political party expires and a new one is formed. In the twentieth century, the spatial model has consistently provided a good fit to the data, even if the agenda was buffeted by a fast pace of external events, including four prolonged armed conflicts overseas and the Great Depression of the domestic economy.

2. The stability of the major dimension. Given the pace of events, it would be possible for the major dimension to shift rapidly in time. In our dynamic model, rapid shifts are to some extent foreclosed by our imposition of the restriction that individual movement can be only linear in time. While the small gains in fit from higher order polynomial models (see Table 1) constitutes evidence that legislators do not shift back and forth in the space, we thought it important to evaluate stability in a manner that allows for the maximum possible adjustment.

To perform this evaluation, we return to model 3 where one-dimensional, static D-NOMINATE was run 99 times, once for each Congress. This gets the best one-dimensional fit for each Congress and allows for the maximum adjustment of individual positions. Since there is no constraint tying together the estimates, we cannot compare individual coordinates directly, but we can compute the correlations between the coordinates for members common to two Houses or two Senates. Rather than deal with a sparse[21] 99×99 correlation matrix, we focus on the correlations of the first 95 Congresses with each of the succeeding four Congresses. This allows us to look at stability as far out as one decade.

[20] More generally, issues that involve redistribution that is geographically concentrated are likely not to be captured by a "spatial" model.

[21] There are no members common to the 1st and 99th Congress and many other pairs.

In the upper portion of Table 5, we average these correlations across the first 95 Congresses and for four periods of history. For both Houses of Congress, the table shows that the separate scalings are remarkably similar, especially since the end of the Civil War. After 1861 a senator could count on a stable alignment, relative to his colleagues, over an entire six-year term ($t + 1$ and $t + 2$).

In the lower portion of Table 5, to save space, we display the individual pairwise correlations only for situations where either the correlation of $t + 1$ was less than 0.8 or a later correlation was less than 0.5; that is, we show periods of instability. Consistent with the preceding discussion, the low correlations are overwhelmingly concentrated in pairs where at least one Congress preceded the end of the Civil War.

It is further noteworthy that a preponderance of the low correlations fall, for both Houses of Congress, in the Era of Good Feelings (at least one pair in the correlation in Congresses 14–18), and, for the Senate, in the period around 1850. These cases are not spatial flip-flops, where two solid major dimensions bear little relation to one another, but simply cases of bad fit where there is not a strong first dimension in some Congress.[22] (The only geometric means below 0.6 for the House occur in Congresses 14, 15, 17, and 32; Congress 17 has the lowest geometric mean for the 99 Senates.)

Subsequent to the Civil War, there is only a $t + 1$ correlation below 0.8 for Senate 44 (1875–76), which preceded the end of Reconstruction, Senate 61 (1909–10), Senate 77 (1941–42), and Senate 69 (1925–26). We note that none of the years in question involves either the realignments of the 1890s or the Great Depression. Thus, the realignments were not shifts in the space but mainly changes in the center of gravity along an existing dimension. The first dimension is remarkably stable; the stability persists through the period in the 1940s, 1950s, and 1960s when a second dimension was also important.

3. Individual stability and reputations. It is possible to obtain high correlations when individuals are moving in the space. If members serving at time t all had nearly equal trend coefficients, their coordinates would remain highly correlated even if they were moving relative to individuals elected later than t.

To assess the stability of individual positions in the space, for each legislator we computed the annual movement implied by the estimated trend coefficients in our two-dimensional linear estimation. Given that the space we estimate is

[22] More precisely, one can hypothesize that two factors will affect the correlation. One is that bad fits lower the correlation. To capture fit, we created a variable that was the average of the two *gmps* for the Congresses in the correlation. The other is that the correlation increases in time, as the political system stabilizes, but at a decreasing rate. This we measured as the logarithm of the Congress number. Corresponding to the columns of Table 5, the eight multiple regressions of the correlations on these variables all showed coefficients with the expected signs. T-statistics had p-levels below 0.005 except for the coefficient on the fit variable in the two $t + 4$ regressions. The R^2 values were 0.33, 0.29, 0.20, and 0.24 for the Senate and 0.25, 0.22, 0.24, and 0.27 for the House.

Table 5. Correlations of Legislator Coordinates from Static, Biennial Scalings

Averages for Years	Senate				House				Number of Congresses
	$t+1$[a]	$t+2$	$t+3$	$t+4$	$t+1$	$t+2$	$t+3$	$t+4$	
1789–1860	.78	.77	.74	.63	.85	.81	.75	.68	36
1861–1900	.90	.87	.86	.86	.93	.91	.89	.91	20
1901–44	.89	.85	.83	.81	.90	.88	.87	.86	22
1945–78	.95	.93	.91	.89	.92	.91	.89	.87	17
1789–1978	.86	.84	.82	.77	.89	.86	.83	.80	95
Individual Congress:[b]									
1	.45	.43	.62	.18	.69	.85	.75	.85	
2	.47	.46	.05	−.65	.86	.80	.62	.44	
3	.49	.73	.50	.13					
4	.79	.76	.95	1.0					
6					.87	.89	.73	.46	
8					.68	.79	.84	.77	
11	.79	.75	.93	.98					
12					.92	.78	.62	.49	
13					.88	.43	.31	.17	
14	.83	.63	.19	−.24	.44	.32	−.11	−.30	
15	.55	.31	−.22	.56	.76	.55	.39	.37	
16	.59	.32	.56	.42	.72	.56	.63	.72	
17	.41	.50	.63	.48	.67	.71	.74	.85	
18	.52	.75	.66	.68	.58	.66	.77	.66	
19					.78	.89	.82	.54	
20					.91	.83	.59	.44	
22					.75	.64	.77	.82	
29	.89	.68	.96	.47					
30	.70	.85	.51	.50					
31	.33	.88	.87	.78					
32	.23	.24	.25	.22					
35	.86	.95	.94	.23					
36					.89	.90	.65	.47	
37					.97	.93	.37	.80	
39	.69	.72	.76	.67					
44	.77	.95	.93	.94					
61	.75	.89	.89	.91					
69					.77	.79	.81	.76	
77	.78	.68	.81	.81					

[a]The notation $t + k$ refers to the correlation of legislator coordinates for legislators serving in Congress t, given in the left-hand column, with their coordinates in Congress $t + k$. Correlation computed only for those legislators serving in both Congresses.

[b]Results shown only for those cases where either the $t + 1$ correlation was less than 0.8 or where any $t + k$ correlation, $k = 2, 3, 4$, was less than 0.5.

identified only up to translations and rotations, one has to interpret the movements in relative terms. The trend coefficient tells us whether an individual is moving relative to legislators whose careers have overlapped the individual's.

Average trends for each Congress are shown in Figure 6.[23] The figure reports results only for legislators serving in at least five Congresses—roughly a decade or more. Similar curves for legislators with shorter careers would appear systematically above those plotted in the figure. Thus, annual movement decreases with the total length of service. Two hypotheses are consistent with this observation. On the one hand, legislators with abbreviated periods of service tend to be unsuccessful legislators; their movement may reflect attempts to match up better with the interests of constituents. On the other, short-run changes in the key issues before Congress (such as the Vietnam War or the current trade gap) may result in the spatial position resembling an autoregressive random walk. In this case, the estimate of the magnitude of true trend will be biased upward, with greater bias for shorter service periods.

Another result, one we see as more important than the finding that spatial movement is limited for legislators with long careers, is shown in Figure 6. Prior to the Civil War, there is a choppy pattern in the figure, most likely in part a consequence of the smaller number of legislators in this period, both in terms of the size of Congress and of the fraction of Congress that served long terms. The key result occurs after the Civil War. It can be seen that spatial movement, which was never very large relative to the span of the space, has been in secular decline, except for upturns in the 1890s realignment and the realignment following the Depression.[24] Since World War II, individual movement has been virtually nonexistent.[25] (The visibility of party defections by Wayne Morse and Strom Thur-

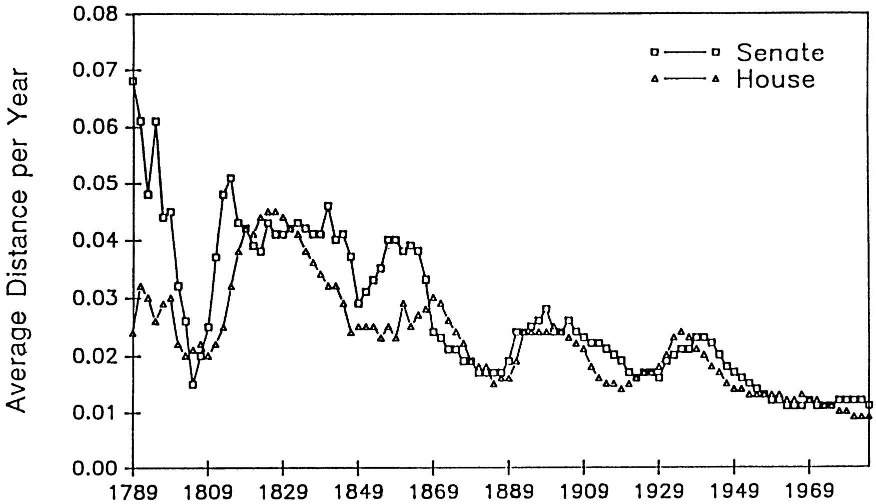
[23] To obtain some idea of the confidence intervals that would bound the curves in the figure, we used the standard errors of the x_{ik}^{1} produced by D-NOMINATE and standard first-order Taylor series methods to estimate standard errors for the numbers plotted. As explained in the Appendix, these errors differ from those produced in a standard MLE setup because the standard errors for the x's are calculated without the full covariance matrix and because D-NOMINATE uses heuristic constraints. For this reason we rely on nonparametric tests in the ensuing two footnotes. Nonetheless, our Monte Carlo work suggests that the standard errors are reasonably accurate. In the House, the estimated standard errors are always below 0.0005 beginning in the 5th House and below 0.0001 beginning in the 64th. In the (smaller) Senate, 0.0005 begins with the 15th Senate and 0.0001 with the 84th. In contrast, the average distance per year is always greater than 0.01. Thus, the numbers reported in Figure 6 would be precisely estimated even if the standard errors on the x's were downwardly biased by a factor of 10!

[24] We tested the proposition that mobility has decreased by carrying out a nonparametric runs test (Mendenhall, Scheaffer, and Wackersley 1986) that compared the distances in Figure 6 for the 50 Congresses in the nineteenth century with the 43 Congresses in the twentieth century. The null hypothesis was no difference, and the alternative was less movement in the twentieth century. The null hypothesis was rejected both for the Senate ($Z = 4.873$, $p = 5 \times 10^{-7}$) and for the House ($Z = 5.293$, $p = 6 \times 10^{-8}$). (In this and later runs tests, we use the large sample approximation.)

[25] Runs tests results comparing the first 22 Congresses in the twentieth century with the 21 Congresses starting in 1945 reject the null hypothesis of no difference under the alternative hypothe-

Figure 6. Annual Movement of Senators and Representatives

Members in at Least Five Congresses



mond proves the rule.) An immediate implication of this result is that changes in congressional voting patterns occur almost entirely through the process of replacement of retiring or defeated legislators with new blood.[26] Politically, selection is far more important than adaptation. Of course, Congress as a whole may adapt by, for example, moving to protectionism when jobs are lost to foreign competition. But as such new items move onto the agenda, their cutting lines will typically be consistent with the preexisting, stable voting alignments.

The current lack of spatial mobility is likely to reflect the role of reputation in U.S. politics. While politicians might choose to adapt to changes in the issues that are salient to their constituency and to changes in the demographic composition of the constituency, the process of adaptation may result in voters who believe the politician is less predictable. In turn, risk-averse voters will value predictability (Bernhardt and Ingberman 1985). Therefore, a politician faces a trade-off between maintaining an established reputation and taking a position that is closer to the current demands of the constituency. Politicians also may find a

sis of less movement since 1945. For the Senate we have $Z = 4.475$, $p = 3 \times 10^{-6}$; for the House, $Z = 5.093$, $p = 2 \times 10^{-7}$.

[26]Lott (1987) shows, using a variety of interest group ratings, that how members of Congress vote is unrelated to whether they face reelection or are planning to retire. In addition, Poole and Daniels (1985) show that members of the House who later are elected to the Senate also tend not to change how they vote.

reputation useful in cultivating campaign contributors. Some mixture of reduced change in constituency demands, increased incentives to maintain a reputation, and perhaps, other factors are manifested by the reduced spatial mobility of legislators.

4. What Is Not Stable and Unidimensional: North and South

Since the Civil War, U.S. politics, in spatial terms, has been remarkably stable. Issues have largely been dealt with in terms of being *mapped* onto a generalized liberal-conservative dimension. Even such major political events as the realignments of the 1890s and 1930s have been accommodated in this manner (Figure 5). During the realignments, legislators changed their positions to a somewhat greater extent than usual (Figure 6), but the changes were largely ones of movement within the existing space. And over time, movement has become more restricted, with, for example, lesser movement during the 1930s realignment than during the 1890s.

North versus South, or perhaps race, may be used to label the major issue that has not fit into the liberal-conservative mapping. While at times the model fits well prior to the Civil War, the fit is less pervasive. The conflict over the extension of slavery to the territories produced the chaos in voting in the 1850s. In the twentieth century, while voting is spatial throughout, a second dimension becomes important in the 1940s, 1950s, and 1960s, when the race issue reappeared as a conflict over civil rights, particularly with respect to racial desegregation and voting rights in the South. The civil rights issue, rather than fully destabilizing the system, was accommodated by making the system two dimensional.

To this point, however, our analysis has not examined the race issue directly. We have only noted anomalies in the fits in periods when race is thought to have been a key issue. Results based on the Clausen and Peltzman categories helped somewhat, but Clausen's civil liberties and Peltzman's domestic social policy codes cover many nonrace issues such as freedom of speech and the Hatch Act. However, we have our own more detailed coding of all roll calls in terms of substantive issues. There is a specific code for slavery and another for civil rights votes that mainly concern Afro-Americans rather than other groups of individuals. The analysis for these issues is contained in Table 6, which provides results for all Houses that had at least five roll calls coded for the topic.

The first appearance of slavery is in 1809–10. The issue—we suspect like many that do not have a sustained appearance on the agenda—does not scale well, even in two dimensions, and disappears for over a decade, until the 15th and 16th Houses. Although these Houses are quite poor in overall fit, the slavery votes fit quite nicely along the first dimension.

From the 23d through the 38th House, there are substantial numbers of slavery votes in every House. From the late 1830s through 1846, slavery is ac-

Table 6. PRE Analysis for Slavery and Civil Rights Roll Calls

				Percent Correct Classification		PRE Over Marginals	
House	Issue and Years	Total Roll Calls	Issue Roll Calls	One Dim.	Two Dim.	One Dim.	Two Dim.
Slavery:							
9	1805–06	158	13	65	73	−.01	.22
15	1817–18	106	13	92	92	.84	.84
16	1819–20	147	16	88	87	.65	.63
20	1827–28	233	9	82	81	.52	.50
23	1833–34	327	5	76	79	.36	.44
24	1835–36	459	70	81	86	.39	.54
25	1837–38	475	39	84	91	.50	.73
26	1839–40	751	27	80	88	.49	.68
27	1841–42	974	84	84	90	.62	.76
28	1843–44	597	44	80	88	.43	.65
29	1845–46	642	8	73	89	.35	.72
30	1847–48	478	29	71	78	.32	.48
31	1849–50	572	26	75	82	.36	.54
32	1851–52	455	14	70	76	.28	.42
33	1853–54	607	159	92	94	.79	.83
34	1855–56	729	115	95	95	.88	.89
35	1857–58	548	65	92	92	.79	.79
36	1859–60	433	24	88	90	.69	.73
37	1861–62	638	92	92	93	.79	.81
38	1863–64	600	13	95	96	.87	.89
Civil Rights:							
37	1861–62	638	43	89	90	.65	.67
38	1863–64	600	30	96	97	.90	.91
39	1865–66	613	32	90	91	.66	.67
40	1867–68	717	7	95	95	.83	.82
42	1871–72	517	23	94	94	.82	.84
43	1873–74	475	87	97	97	.91	.91
48	1883–84	334	9	90	91	.78	.79
49	1885–86	306	5	78	80	.25	.33
56	1899–1900	149	9	97	97	.94	.94
67	1921–22	362	14	97	97	.90	.91
77	1941–42	152	8	80	95	−.01	.73
78	1943–44	156	7	78	92	.29	.76
79	1945–46	231	6	69	87	.00	.57
81	1949–50	275	9	67	89	.04	.69
85	1957–58	193	6	70	92	.01	.73

Table 6 (*continued*)

House	Issue and Years	Total Roll Calls	Issue Roll Calls	Percent Correct Classification One Dim.	Percent Correct Classification Two Dim.	PRE Over Marginals One Dim.	PRE Over Marginals Two Dim.
Civil Rights:							
86	1959–60	180	8	72	92	.08	.74
87	1961–62	240	5	80	92	.01	.59
88	1963–64	232	7	77	94	.33	.82
89	1965–66	394	22	83	91	.53	.75
90	1967–68	478	8	78	91	.35	.73
91	1969–70	443	8	85	90	.60	.73
92	1971–72	649	22	80	84	.45	.55
93	1973–74	1,078	17	80	83	.48	.56
95	1977–78	1,540	6	81	81	.56	.58
96	1979–80	1,276	17	85	86	.58	.61
97	1981–82	812	9	86	88	.50	.59

Note: One Dim. and Two Dim. refer to the one-dimensional and two-dimensional dynamic scalings with linear trends in legislator positions.

commodated within the spatial structure by a second dimension. Classifications and PREs are high during this period. The destabilization of the issue begins in 1847 and continues through 1852, a period centered on the Compromise of 1850. There is a substantial reduction in the ability of the spatial model to capture slavery votes. A temporary reduction in the actual number of slavery roll calls perhaps testifies to the difficulty of dealing with the issue. The issue, in fact, could not be accommodated by the existing party system. Free Soilers and States Rights adherents appear in Congress in increasing numbers in the early 1850s, and the elections of 1856 mark the virtual completion of the process of the replacement of Whigs by Republicans.

As the party system changed, true spatial realignment occurred. From 1853–54 onward, when the Kansas-Nebraska Act was passed, slavery votes fit the model exceptionally well *on the first dimension*. Particularly, in 1853–54, when slavery represented a quarter of all roll calls, slavery most likely defined this dimension. First-dimension classifications and PREs are remarkably high.

The North-South conflict persists on the dimension via civil rights votes from the Civil War through the 43d House, 1873–74. Classifications and PREs remain high. (The alternative label "race" is suggested by the fact that roll calls in the category nullification/secession/reconstruction also line up on the first dimension, but with somewhat lower PREs than civil rights.) From 1875 through 1940, civil rights votes occur only sporadically. In line with our contention that race is the major determinant of spatial realignment, this period is stable and largely unidimensional.

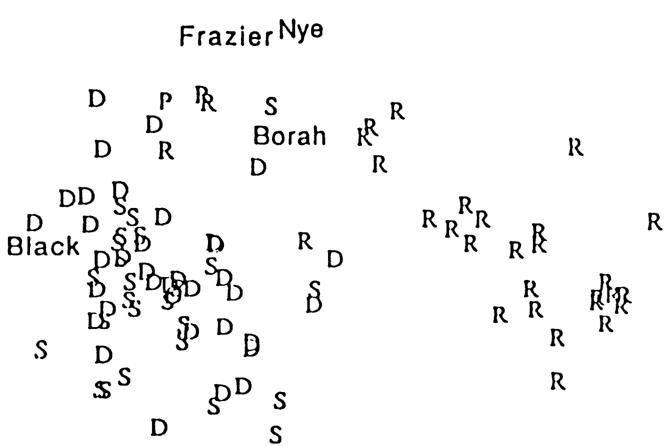
Civil rights reappears on the agenda in 1941–42. For the next 30 years, it remains as an issue where the second dimension is always important. Through 1971–72, the two-dimensional scaling always adds at least 0.10 to the PRE. Indeed, the first dimension is often orthogonal to civil rights; several one-dimensional PREs in Table 6 are close to zero. In five instances, the second dimension increased PRE by over 0.5. Despite the fact that our detailed coding of issues produced 978 occurrences of issue codes with at least five votes in a Congress, a PRE improvement over 0.5 occurred only one other time in the 99 Houses—public works in 1841–42.

The pattern of PRE improvements is echoed by Figure 5; the period from roughly 1941–42 to 1969–70 is the only period where the two-dimensional model consistently and rather substantially out-performs the other two models. In fact, the second-dimension curve in model 1 is even further above that of model 3 for the Senate during this period, reflecting in part the many cloture votes the Senate took on civil rights filibusters. In the House, unlike slavery votes in the 1850s, civil rights votes were never a substantial fraction of the votes before Congress. The debate was contained. By the 1970s and 1980s, the debate had shifted from one of changing the status of southern blacks to measures that would have a national impact on civil rights. Correspondingly, although civil rights votes occur in the 93d, 95th, and 97th Houses, the second dimension no longer makes a major contribution.

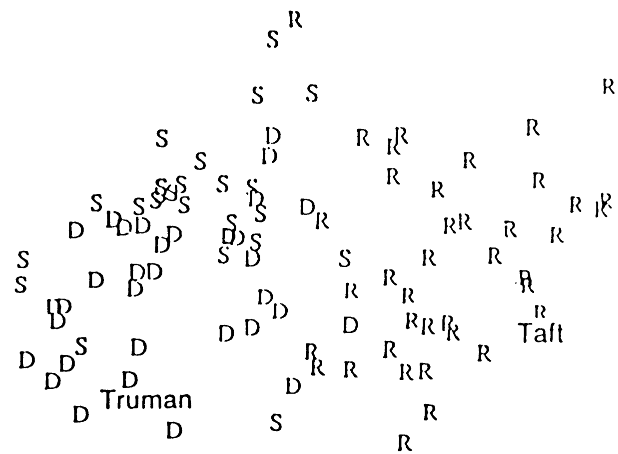
The accommodation of the civil rights issue to the political system is traced, for the Senate, in Figure 7. At the inception of Franklin Roosevelt's administration, race was not an important political issue, and the primary concern of southern Democrats remained the South's economic dependence on northern capital. As a result, southern Democrats appeared largely as a random sample of all Democrats, or to the extent they were differentiated, they represented the liberal wing of the party (Figure 7.A). As the importance of the race issue intensified in the 1940s, southern Democrats began to be differentiated from northern members of the party (Figure 7.B). By the mid-1960s, when civil rights was the dominant item on the congressional agenda, southern Democrats had separated nearly completely (Figure 7.C), and there was a virtual three-party system. In these years roll call votes on civil rights issues tended to have cutting lines parallel to the horizontal axis, opposing southerners, at the top of the picture, and a few highly conservative Republicans to the rest of the Senate. In later years, confronted with increasingly large numbers of registered black voters, southern Democrats gradually took on the national party's role of representing such groups

Figure 7

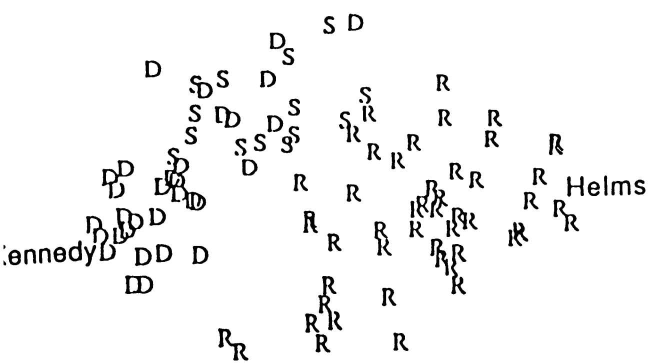
Senate 73, 1933–34



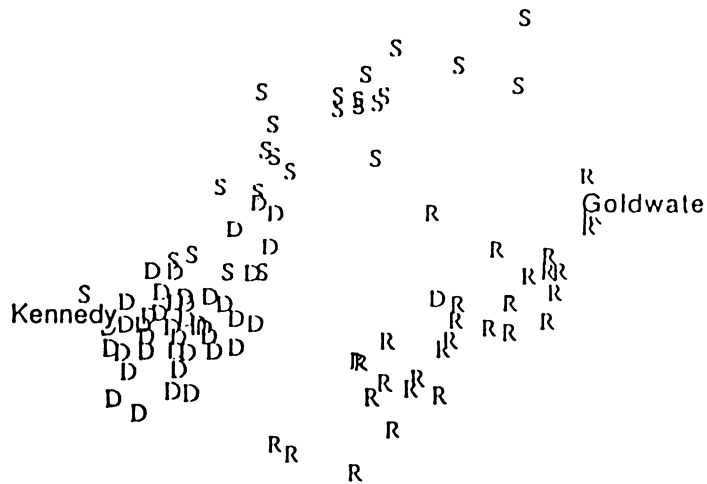
Senate 78, 1943–44



Senate 99, 1985–86



Senate 88, 1963–64



D = Democrat   R = Republican
S = Southern Democrat

as minorities and public employees.[27] Consequently, the differentiation of southern from northern Democrats decreased (Figure 7.D).[28]

On the whole, the process we have traced occurred largely through changes in the membership of the southern Democrat delegation in Congress. Those who entered before the passage of key legislation in the 1960s tended to locate in more conservative positions than those who entered after (on this point see Bullock 1981). The changes by southern Democrats have resulted in the 1980s being not only a period in which spatial mobility is low but also one which is nearly spatially unidimensional.

5. Predictions from the Spatial Model

The great stability of individual positions implies that the spatial model can be used for short-term forecasting. To illustrate, consider the vote on Judge Bork in 1987. The spatial positions, shown in Figure 8 result from model 3 estimation with only 1985 data used. Quite early in the confirmation process, the five most-liberal members of the Judiciary Committee came out in opposition to Bork, and the five most-conservative members supported him. The last four members of the committee to take a public stance were between these two groups, a finding parallel to our earlier result that "errors" tend to occur close to cutting lines.

As soon as Arlen Specter (R-PA) made known his opposition, it was possible to predict, accurately, that the final committee vote would be 9–5 against, since the remaining three undecided members were all more liberal than Specter. At this time, using the fact that Grassley (R-IA) had announced support for Bork, one could predict a final Senate vote of 59–41 against. (The four senators between Specter and Grassley on the scale were predicted to split 2–2, senators elected since 1985 were predicted to vote on party lines.) In Figure 8 we present

[27] Cox and McCubbins (1989) note that southern Democrats who deviated too far from the northern wing in the 1970s were punished in the House by having their seniority violated. Thus, the movement of southern Democrats may also reflect the internal dynamics of the majority party as well as constituency changes.

[28] To provide statistical backup for the statements concerning differentiation of northern and southern Democrats, we carried out a runs test. To carry out the test, we calculated the interpoint distance between each pair of Democratic senators. Pairs were then tagged as to whether they were the same (North-N or South-S) or opposite N-S. They were then rank-ordered by distance, and the runs statistic was calculated. The null hypothesis was that there was no difference between same and opposite. The alternative was that opposite distances were greater than same. The results are: $Z = -0.52$, $p = 0.302$, 73d Senate; $Z = -0.99$, $p = 0.161$, 78th; $Z = -17.65$, $p < 8 \times 10^{-10}$, 88th; $Z = -2.93$, $p = 0.002$, 99th. These results show that there was a very slight, insignificant increase in differentiation from the 73d to the 78th Senate, a sharp increase from the 78th to the 88th and a substantial decrease between the 88th and 99th. Although the runs for the 99th Senate are "significantly" less than expected by chance, it is also true that we reject the null hypothesis $D \equiv (R_{99} - R_{88}) - (E(R_{99}) - E(R_{88})) = 0$ using the one-tailed alternative hypothesis $D < 0$ with $p < 8 \times 10^{-10}$, where R_t is the number of runs for Senate t.

some information on the temporal ordering of announcements as well as the final vote. Note that, echoing the committee, moderates tended to announce relatively late.[29] The actual final vote was 58–42. At the individual level, the spatial model correctly forecast the vote of 93 of 100 senators. As was the case with Figure 2, the errors tend to be close to the cutting line.

6. Conclusion: Consensus and Impasse

Although the spatial model has an applied use in short-term prediction, its greater relevance is in what it indicates about long-term changes in our political system. The average distance between legislators within each of our two major parties has remained remarkably constant for more than a century (Figure 9). The maintenance of party coalitions apparently puts considerable constraint on the extent of internal party dissent.

In contrast, the average distance between the parties—and by inference the average distance between all legislators—has shrunk considerably in the past 100 years.[30] The conflict between Edward Kennedy and Jesse Helms is undoubtedly narrower than that which existed between William Jennings Bryan's allies and the Robber Barons. Symptomatic of the reduced range of conflict is the willing-

[29] Statistical support for this statement is furnished by a McKelvey-Zavoina (1975) ordinal probit analysis where the dependent variable is coded 1 = Announced before 7 Oct., 2 = Announced on 7 Oct., 3 = Announced after 7 Oct., the regressors were a constant and the absolute value of the distance of the senator from the midpoint of Specter and Grassley, and the sample was the 72 noncommittee members serving in 1985. The null hypothesis that each senator chose an announcement date according to the marginal frequencies (60/72, 7/72, 5/72) was rejected with $p = 5 \times 10^{-4}$. As is standard procedure, the variance of the probit was set to unity, and the cutpoint between the first two categories was set to zero. The null hypothesis of a zero slope on distance was rejected with $p = 0.0019$ and of a zero cutpoint between the second and third categories with $p = 0.010$. The *New York Times* and *Washington Post* were used as sources for the announcement dates.

[30] As with previous figures, the numbers displayed in Figure 9 are very precisely estimated. For example, using the variance-covariance matrices of the estimated x coefficients for each legislator and Taylor series methods, we computed standard errors for the within-party average distances shown in the figure. (Previous caveats apply.) The Z-statistic is the ratio of the estimate to the estimated standard error; the *minimum* (over 64 Congresses) Z was 6.06 for House Democrats and 5.23 for House Republicans. Because of the precision of the estimates, small differences in the graph will often be "statistically significant." For example, we can directly compute a Z for the difference between the two within-party distances. When the Democrats are more heterogeneous, the Z is greater than 2.0 in magnitude in Congresses 36, 40, 48, 52, 53, 64, 66, and 76–99. Although the Republicans are estimated to be more heterogeneous than Democrats for some Congresses, the Z never exceeds 2.0 in these cases. The greater heterogeneity of the Democrats in Congresses 76–99 (see also Figure 7) reflects the important civil rights conflict. Statistical significance is aided by the fact that our x's are more precise for these years as a result of longer periods of service and more stable individual voting patterns (Figure 6). While statistically significant and often substantively important, the changes in heterogeneity are dwarfed in importance by the changes in the distance between the parties.

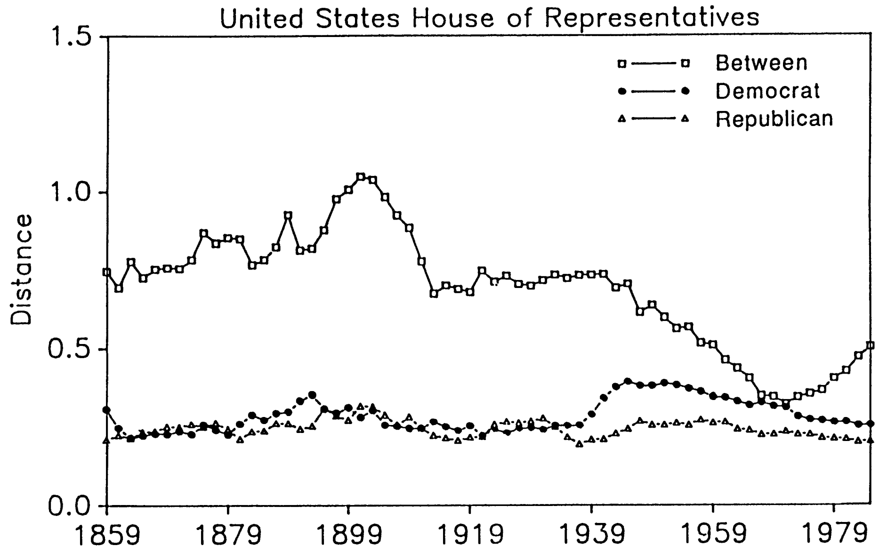
Figure 8. The Bork Vote

		Judiciary Committee: Committed	Judiciary Committee: Undecided	Other Senators: Announced before Oct. 7	Other Senators: Announced on Oct. 7	Other Senators: Announced after Oct. 7	
	−1						
				Sarbanes			
		Simon		Inouye			
	−0.8	Metzenbaum		Melcher, Harkin			
				Burdick, Levin, Riegle			
		Kennedy		Kerry, Cranston, Pell			
				Lautenberg, Moynihan			Predicted
				Mitchell, Dodd			Anti-
				Baucus, Rockefeller	Sasser		Bork
Liberal	−0.6	Leahy		Gore, Pryor			
		Biden	Byrd	Bumpers, Bradley,	Chiles		
				Glenn, Ford			
				Bingaman, Johnson			
					Dixon		
					Exon		
	−0.4		Deconcini				
				Bentsen		Proxmire	
				Weicker		Stennis	
			Heflin	**Boren***		Nunn	
						Hollings*	
			Specter		**Hatfield***		Cutting Line
	−0.2						
				Durenberger, Cohen		Heinz	
		Grassley		**Stafford***			
				Chaffee*, **Packwood***		D'Amato	

	0			Kasten, Danforth Pressler			
				Stevens, Evans Kassebaum			
	0.2						
				Cochran, Rudman Murkowski, Nickles Trible, McConnell Boschwitz, Roth		**Warner***	Predicted Pro- Bork
Conservative	0.4			Lugar, Wilson, Dole Domenici			
		Quayle Simpson					
		Hatch		Thurmond			
	0.6						
				Garn, Goldwater			
				Armstrong			
	0.8	Humphrey		Gramm, Hecht			
				McClure			
	1			Helms, Wallop, Symms			

*Prediction errors in bold.

Figure 9. Average within and between Party Distances



ness of most corporate political action committees to spread their campaign contributions across the entire space, the most liberal Democrats in the southwest quadrant excepted (see Poole and Rosenthal 1989b). Although a well-defined two-party system persists and although liberals and conservatives maintain stable alignments within each party (Figure 7.D), the range of potential policy change has been sharply reduced. While our earlier contention (Poole and Rosenthal 1984) that polarization increased in the 1970s is supported by Figure 9, the long-term, more-relevant pattern has been toward a national consensus. The cost of consensus can perhaps best be seen in the alleged wasteful concessions to special interests in such programs as agriculture, space, and defense and in the alleged failure of the nation to address a variety of inequities that befall various groups of citizens. The benefits are perhaps made apparent by recalling the Civil War and the period of intense conflict surrounding the labor movement in the late nineteenth and early twentieth centuries, and by observing the fragility of democracy in some other nations. We do not pass judgment but simply point to a regularity in a system that, as manifest by its ability to absorb the supposed Reagan Revolution, is likely to be with us indefinitely (for a similar substantive conclusion, see Fiorina 1989, 141).

Manuscript submitted 21 April 1989
Final manuscript received 27 April 1990

APPENDIX

The Estimation Algorithm

The algorithm employed to estimate themodel simply extends the procedure developed in detail in Poole and Rosenthal (1985). The earlier model was restricted to unidimensional, constant coordinates. Here we briefly sketch the new procedure, D-NOMINATE, with emphasis on modifications needed to handle the more general model. As before, the procedure begins by using a starting value for β and a set of starting values for the senator parameters x_{ik}^{0}. (The other polynomial coefficients are initially set to zero.) These are obtained by metric similarities scaling (Torgerson 1958; Poole 1990).

For the metric scaling, an agreement score is formed for each pair of legislators with a common period of service. The score is simply the percentage of times they voted on the same side. Scores vary from 0 to 100 and thus are analogous to the interest group ratings that were subject to metric unfolding scaling in Poole and Daniels (1985). The matrix of these scores is converted into squared distances by subtracting each score from 100, dividing by 50, and squaring. The matrix of squared distances is double centered (the row and column means are subtracted, and the matrix mean is added, to each element) to produce a cross-product matrix. Eigenvectors are extracted from the cross-product matrix and used to start the metric scaling procedure. This procedure produces the starts for D-NOMINATE.

D-NOMINATE proceeds from the starts by using an alternating algorithm (a common procedure in psychometrics). In a first stage, the roll call coordinates are estimated on the first dimension, holding the legislator coordinates and β constant. Since roll call coordinates are independent across roll calls (for fixed β and legislator coordinates), each roll call can be estimated separately. In a second stage, the legislator coordinates are estimated on the first dimension, again holding everything else constant. Because each legislator's choice depends only upon his own distances to the roll call outcomes, if β and the roll call alternatives are held fixed, each legislator's choice is independent of those of all other legislators. Independence allows us to estimate each legislator's coordinate separately. These two stages are then repeated for each higher dimension. In a third stage, we estimate the utility function parameter β, holding constant all x and z values. Within each stage, we use the method of Berndt et al. (1974) to obtain (conditional) maximum likelihood estimates of the parameters. These three stages define a *global iteration* of D-NOMINATE. Global iterations are repeated until both the x's and the z's correlate above 0.99 with their values at the end of the previous global iteration.

Constraints

In Poole and Rosenthal (1985), we explain in detail why, even when the underlying model presented in part 1 accurately represents behavior, estimated x and z values can run amok, taking on values with exceptionally large magnitudes. The problem is basically an identification problem. Legislators who are highly liberal, for example, will tend almost always to vote on the liberal side of an issue. As there is thus not enough information to pin down a precise location for these legislators, their estimates are constrained. The problem should not be exaggerated, however. We still know, reliably, that the individual is an extremist in the direction of the constrained estimate.

Similarly, it is hard to pin down the location of a "Hurrah" vote, one where almost everyone votes on the same side of the issue. We included such lopsided votes because, noisy as they are, they provide some information that helps us to differentiate legislators at the extremes of the space. We only excluded roll calls that failed to have at least 2.5% of the vote cast for the minority side. This cutoff rule reflects experimentation (Poole and Rosenthal 1985) with one-dimensional estimation of the 1979–80 Senate. A better multidimensional, dynamic algorithm might result from exploring alternative cutoff rules, but we elected not to allocate scarce computer resources to such a study.

Identification problems are accentuated by the obvious specification error implicit in the assumption of a homoscedastic logit error. In fact, some roll calls are highly noisy. D-NOMINATE will try to arrange their cutting lines so that all legislators are predicted to vote with the majority; this will cause the cutting line (and at least one z) to drift outside the space spanned by the legislators and invoke a constraint. Conversely, some roll calls are noiseless, representing perfect spatial voting. Although the cutting line is precisely identified, maximum likelihood will try to put both z's at a very large distance from the legislators.

To deal with these identification problems, constraints are imposed. After unconstrained x estimates have been obtained on a dimension, the maximum and minimum coordinates for each Congress are used to define constraint coordinates:

$$\tilde{x}_{tk} = \left(\frac{[\max_i x_{itk}]^2 + [\min_i x_{itk}]^2}{2} \right)^{1/2}$$

For each legislator, we then define a constraining ellipse by computing, for $k = 1, \ldots, s$

$$\bar{x}_{ik} \equiv \frac{1}{n_i} \sum_{t \in T_i} \tilde{x}_{tk}$$

The $\bar{x}_{ik}$ and the origin define an ellipse. We then compute for $k = 1, \ldots, s$

$$\overset{*}{x}_{ik} \equiv \frac{1}{n_i} \sum_{t \in T_i} x_{itk}$$

where $T_i = \{t \mid i \text{ is in the estimation in Congress } t\}$ and $n_i = |T_i|$, the number of Congresses served in by i.

Thus, averages are being taken over all Congresses for which the legislator was in the data set. If the point defined by the $\overset{*}{x}$ is in the interior of the legislator's ellipse, the legislator is unconstrained. Otherwise, the parameters x^1_{ik} and x^2_{ik} on the dimension currently being estimated are set to zero and x^0_{ik} is constrained to keep the legislator inside the ellipse. In other words, a legislator is not allowed to drift too far from the origin relative to others who overlap his or her service period. Note that, unless T_i is identical for all i, the entire configuration is not constrained to the same ellipse.

A similar constraint is imposed in the roll call phase. For each Congress, we use the $\tilde{x}_{tk}$ and the origin to define an ellipse. If the roll call midpoint, defined by coordinates $(z_{jyk} + z_{jnk})/2$, is inside the ellipse, the roll call is unconstrained. Otherwise, the coordinate currently being estimated is adjusted to keep the midpoint within the ellipse.

In the actual estimation, constraints were invoked for 4.2% of the legislators in the two-dimensional, linear House estimation. A check of the estimates in the postwar period shows that the constrained legislators are overwhelmingly known "extremists." More constraints are needed for roll calls, where 14.1% are constrained. However, breaking down the roll calls by margin and by time period in Table A.1 shows that the problem is less serious than indicated by this aggregate percentage. First, it is clear that, as argued above on theoretical grounds, constraints are most often invoked for lopsided roll calls. Constraints are invoked in under 1% of the roll calls that are closer than 55–45. Almost all roll calls of interest to scholars will be unconstrained. In contrast, constraints are needed on more than half of the most lopsided votes. Second, controlling for margin, the fraction constrained is less beginning with the 76th Congress than before. The larger overall proportion of constrained roll calls in the modern period simply reflects an increase in the relative number of lopsided "Hurrah" votes.

Standard Errors

The use of constraints implies that conventional procedures for computing standard errors do not apply. Another problem with the standard errors produced by the D-NOMINATE program is that

Table A.1. Constrained Roll Calls by Margin

Margin	Congresses 1–75 Percent Constrained	Congresses 1–75 Total Roll Calls	Congresses 76–99 Percent Constrained	Congresses 76–99 Total Roll Calls
50–55	1.0	5,933	0.5	1,756
55–60	4.2	5,012	2.6	1,571
60–65	8.0	3,510	4.4	1,241
65–70	11.6	2,788	8.0	1,022
70–75	18.0	1,949	13.3	791
75–80	24.5	1,308	15.2	677
80–85	38.7	874	20.3	661
85–90	57.7	655	28.2	611
90–95	69.9	625	53.5	905
95–97.5	76.4	399	64.7	665
Total	13.0	23,053	16.3	9,900

they are based only upon a portion of the covariance matrix. A standard error for a legislator coefficient in the linear dynamic model comes, for example, solely from the inversion of the 2 × 2 outer product matrix computed when the legislator's coordinates are estimated for a given dimension in the legislator phase of the alternating algorithm. The appropriate procedure would be to compute, at convergence, the estimated information matrix for all parameters and invert. This computation is impractical, even on supercomputers. With the dynamic models covering 1789–1985, the matrices would be larger than 100,000 by 100,000.

For the reasons outlined above, the standard errors reported in the text must be viewed as heuristic descriptive statistics. To get a handle on the reliability of the reported errors, we applied Efron's (1979) bootstrap method to estimate the standard errors of the legislator coordinates for a model 3 estimation of the 94th Senate. That is, we took the 1,311 actual roll calls and drew 50 samples of 1,311 roll calls. Each roll call was sampled with replacement, so in any particular sample, some actual roll calls will not appear while others will appear more than once. We ran D-NOMINATE for each of the 50 samples and then computed the standard deviation of the 50 estimates for each senator. The results appear in Table A.2. The largest bootstrap standard error is 0.051, and 73 of 100 senators have bootstrap standard errors under 0.03. Since the space has a range of two units, the senator locations are precisely estimated. We did not apply the bootstrap method to multidimensional or dynamic estimation as a matter of economizing computer time. However, it is clear, at least in one dimension, that the dynamic model estimates will be even more precise than those reported in Table A.2. This is because, typically, three or more Senates, rather than one, have been used to estimate the location of a senator.

Consistency

In addition to the statistical problem posed by our imposition of constraints, we have an additional problem that reflects the fact that every legislator and every roll call has a specific set of parameters. Therefore, we always have additional parameters to estimate as we add observations. This generates what is known as an "incidental parameters" problem in the econometrics literature. In fact, every parameter we estimate, except for β, is "incidental." As a result, the standard proof of the consistency of maximum likelihood does not apply. We are not guaranteed that, even with "infinitely" many observations, maximum likelihood estimates will converge to the true values of

Table A.2. Distribution of Bootstrap Standard Errors for Senators, 94th Senate

Range of Bootstrap Standard Error	Number of Senators
0.00–0.01	0
0.01–0.02	11
0.02–0.03	62
0.03–0.04	21
0.04–0.05	5
0.051	1
Total	100

Note: One-dimensional scaling.

the parameters. At a practical level, this caveat is not important. The key point is that data is being added at a far faster rate than parameters. Consider the two-dimensional linear model. Assume our time series were augmented by a new Senate with 15 freshman senators and 500 roll calls. We would eventually add 60 parameters for the senators (assuming they all acquired trend terms) and 2,000 parameters for the roll calls. To estimate these 2,060 new parameters, we would have 50,000 (100 × 500) new observations. The ratio of observations to parameters is 25 to one. For the House of Representatives, a similar ratio would be over 100 to one. Many empirical papers published in professional social science journals have had far lower ratios.

Haberman (1977) obtained analytical results on consistency for a problem closely related to ours. He treated the Rasch model from the educational testing literature. In place of p legislators voting on q roll calls, p subjects take q tests. Each subject has an "ability" parameter, and each test has a "difficulty" parameter. The role of yea and nay votes is played by correct and incorrect answers. A version of the Rasch model analyzed by Lord (1975) is in fact isomorphic with a one-dimensional Euclidean model of roll call voting developed by Ladha (1987).

Haberman considered increasing sequences of integers $\{q_n\}$, $\{p_n\}$ for which $q_n \geqslant p_n$. In other words, the number of roll calls always exceeds the number of legislators. In addition, make the (innocuous) technical assumption that $\log(q_n)/p_n \rightarrow 0$ as $n \rightarrow \infty$. Under these conditions, Haberman establishes consistency for the Rasch model.

Few actual processes, including Congress, can be thought of as satisfying Haberman's stringent requirement that p and q both grow as n grows. But Haberman cites Monte Carlo studies by Wright and Douglas (1976) that show excellent recovery for conventional maximum likelihood estimators for p between 20 and 80 and q of 500. In D-NOMINATE, the effective p is 100 in the modern Senate and 435 in the modern House. The effective q is about 900. Similar Monte Carlo evidence is contained in Lord (1975).

Another source of comfort, in addition to the work on the Rasch model, is provided by Poole and Spear (1990). They proved, for a wide class of error distributions, that Poole's (1990) method of scaling interest group ratings gives a consistent estimate of the ordering of the legislators. Since, for the modern period, interest group scaling and D-NOMINATE give similar results, it is likely that D-NOMINATE also gives a consistent ordering.

Nonetheless, the D-NOMINATE model uses individual choices directly whereas the Poole (1990) method scales distances. In the D-NOMINATE model, utility is, unlike in the Rasch model, a nonlinear function of the parameters. Thus, while the theoretical work of Haberman (1977) and Poole and Spear (1990) and the previous Monte Carlo results are suggestive, it is appropriate to conduct direct Monte Carlo tests of our algorithm.

Monte Carlo Results

Summary of previous experiments. Previously, in Poole and Rosenthal (1987), we reported on extensive Monte Carlo studies of the one-dimensional NOMINATE algorithm. As "true" locations, we used the estimated senator coordinates from a scaling of 297 roll calls that took place in 1979. We used a wide variety of alternative sets of "true" roll call coordinates and used alternative "true" values of β between 7.5 and 22.5. Over all runs, the squared correlations between the recovered legislator coordinates and the true ones exceeded 0.98. The standard error of recovery of individual coordinates (the variability across Monte Carlo runs) was on the order of 0.05 relative to a space of width 2.0. Recovered β's are slightly higher than the true. Recovery came closer to the true value as the number of observations was increased. Thus, our one-dimensional estimates are highly accurate under the maintained hypothesis of the model. Moreover, as most of the discussion in the paper is essentially based on summaries of parameter estimates (e.g., Figure 6 uses average distances), statistical significance would not be an issue.

We also did the counterfactual experiment of setting β to zero by generating data with a random coin toss. The recovered distribution of legislators was far more tightly unimodal than any recovered from actual data. More important, the geometric mean probability was only 0.507, far lower than almost all values reported in Figure 5.

An extensive test of the static model in one and two dimensions. The work we have just summarized was done under the assumption that the data was generated by a true one-dimensional world with a constant signal-to-noise ratio, β. We subsequently undertook additional work that examines the performance of D-NOMINATE with a true two-dimensional world. In addition, we studied how robust D-NOMINATE was to violation of the constant β assumption. All the work was done using the static model on one Senate of a hypothetical single Congress. We did not pursue simulations of the dynamic model because such simulations are too costly in computer time.

To simulate a Senate, we had 101 senators vote on 420 roll calls. Finding good recovery with only 420 roll calls should bode well for our actual estimations, since, in the past two centuries, legislators have averaged 909 votes in their careers. Thus, to generate the "observed" choices, in each simulation we drew $84{,}840 = 2$ (choices) $\times$ 420 (roll calls) $\times$ 101 (senators) random numbers from the log of the inverse exponential distribution.

In each simulated Senate, we drew the senators' coordinates uniformly from $[-1, +1]$. Thus, in one dimension the senators were distributed on a line of length two; in two dimensions, on a square of width two. Midpoints of roll calls followed the same distribution.

In our simulation design, one comparison was a true one-dimensional world versus a true two-dimensional world. A second was comparing a world with a fixed β, set equal to 15.75 to match typical estimates from actual data, and a variable β. When the variable β model was used, for each roll call we drew $1/\beta$ uniformly from [.043, .123]. The β's were thus in the interval [8.13, 23.26]. The median of the distribution of the random β's was 15.75. The third design factor was low versus high error rates. The error rate is the percentage of choices that are for the further alternative, that is, the percentage of choices for which the stochastic error dominates the spatial portion of the utility function. After the legislator positions, the βs, and the midpoints have been assigned, the error rate can be controlled by scaling the distance between the yea and nay outcomes. The smaller the average distance, the greater the error rate. Scale factors were chosen to keep the average rate in the low condition close to 14%, about the level of classification error attained by D-NOMINATE with the congressional data. In the high condition, the error rate was set to 30%, above that for the worst Congresses in our actual scalings.

Our design yielded $8 = 2 \times 2 \times 2$ conditions. In each condition, we ran 25 simulations, for a total of 200. We emphasize that each simulation had the following sources of randomization: (1) spatial locations of legislators and roll calls; (2) utility of each choice; (3) β for each roll call (in variable β condition only).

We computed two sets of statistics to assess the recovery by D-NOMINATE. First, we computed the 5,050 distances that represented all distinct pairs of the 101 legislators. For every one of the 200 simulations, we did this both for the "true" distances and the "recovered" distances. Since substantive work using the scaling will depend only on relative position in a space, distances summarize all the information in the scaling. Focusing on distances not only eliminates arbitrary scale and rotational differences between true and recovered spaces but also reduces assessment to a single criterion, rather than looking at one dimension at a time. Second, we cross-tabulated the yea-nay predictions from the scaling with the yea-nay predictions from the "true" spatial representation. The percentage of matches is a good measure of fit. Comparing simulated predicted to "true" predicted is better than comparing simulated predicted to "true" actual because the scaling is designed to recover the systematic, spatial aspect of voting, not the errors. So we want to know how well D-NOMINATE scaling noisy data would predict voting if the noise were removed.

The simulation results are presented in Table A.3. An immediate observation is that the two-dimensional world is not recovered as well as the one-dimensional world. This is to be expected. The number of "observations" is identical in every design condition, but the parameter space is doubled in moving to two dimensions.

The first column of the table shows how well the recovered distances correlate with the "true." It can be seen that D-NOMINATE is very robust with respect to variability in noise across roll calls. Recovery of senators is totally insensitive to whether the "true" world has a fixed β across all roll calls or one with considerable variability. On the other hand, fit does decline with dimensionality. Raising the error level forces only a moderate deterioration in fit.

To some degree the lack of higher measures of fit in two dimensions reflects the constraints in D-NOMINATE. The constraints force estimates into an ellipse when estimation is restricted to a single Congress. In one dimension, this has no impact, but in two dimensions the "true" coordinates come from a square. The impact of the constraints can be seen in comparing the last column of Table A.3 to the first. The last column reports correlations between the recovered solutions. In one dimension these correlations are slightly less than the correlations of the recovered with the true distances. In two dimensions the pattern reverses. As the distorting constraints tend to get invoked for the same senators in all recoveries, the recovered points, particularly at low error levels, tend to be very similar.

The impact of the constraints is also seen in the second column. The distances between senators are more precisely estimated, in two dimensions, when the error level is high. With a high error level, legislators near the periphery of the space have some "noise" in their voting patterns, and the constraints are invoked less frequently. (The higher percentage of precise estimates in two-dimensional high as compared to one-dimensional high reflects the fact that the average distance is greater in the two-dimensional space than in the one-dimensional space, so ratios of true distances to standard errors tend to be greater.)

Actual vote predictions are less sensitive to whether the constraints are invoked. The third column of the table shows results that appear most clearly to indicate the high quality of the recovery. At a low level of error, the implications for predictions of actual choices from the spatial model are nearly identical between the true space and the recovered space. There is only a slight deterioration (94% vs. 96%) when two dimensions must be estimated. The fit is still good, but less than perfect, at (very) high levels of error. In all cases the standard errors are extremely small, demonstrating that our simulation results are insensitive to the set of random numbers drawn in any one of the 200 simulations.

Tests using real data but with random starting coordinates. We have seen that the D-NOMINATE algorithm reliably recovers a "true" spatial configuration. We also are guaranteed, were there no constrained parameters, that D-NOMINATE is an ascending algorithm—the likelihood is improved at every step of the alternating procedure. Nonetheless, the likelihood function is not globally convex. Either the lack of global convexity or the constraints problem could result in the

Table A.3. Monte Carlo Results for Simulated Data

Error Level	β	Pearson R with True Config.	Percent Precisely Estimated	Proportion Correct Prediction	Stability of Solution
One-dimensional experiments:					
Low[f] (14%)	Fixed	.995[a]	90.8[b]	.962[c]	.992[d]
		(.001)[e]		(.003)	(.004)
	Variable	.995	91.0	.963	.977
		(.001)		(.003)	(.006)
High (30%)	Fixed	.970	75.8	.910	.947
		(.001)		(.011)	(.013)
	Variable	.976	77.0	.918	.958
		(.003)		(.004)	(.009)
Two-dimensional experiments:					
Low (14%)	Fixed	.894	71.7	.942	.955
		(.022)		(.009)	(.024)
	Variable	.896	69.6	.945	.962
		(.021)		(.003)	(.022)
High (30%)	Fixed	.847	77.8	.885	.851
		(.026)		(.017)	(.031)
	Variable	.841	81.5	.880	.841
		(.019)		(.005)	(.017)

[a] A Pearson correlation was computed between the 5,050 unique pairwise distances generated by the estimated coordinates for the 101 legislators and the true pairwise distances. One correlation was computed for each of the 25 simulations used in each design condition. The number reported in the table is the average of these 25 correlations.

[b] Each true unique pairwise distance between legislators (n = 5,050) was treated as a mean, and a standard error was computed around this mean using the 25 estimated distances. The entries show the percentage of true distances that are twice this standard error. In other words, the percentage that have a "pseudo-t" statistic greater than 2.0.

[c] Each predicted choice was compared to the true choice that would have been made had there been no stochastic term in the utility function. A percentage correct was computed for each of the 25 trials. This number is the mean of these 25 percentages.

[d] This number measures the stability of the estimated legislator coordinates. Pearson correlations were computed between each unique pair of the 25 estimated configurations, and this number is the average of those Pearson correlations. The n is 300.

[e] Standard deviations are shown in parentheses.

[f] This number measures the noise level in the 25 trials. Holding the legislator configuration, the roll call midpoint, and then β for the roll call constant, the percentage of times legislators who do not vote for the closest alternative varies inversely with the distance between the alternatives. The error level is this percentage. Distances between alternatives were scaled to achieve the low and high levels.

recovery being potentially sensitive to the starting values. Perhaps even (slightly) better recoveries would result if a different starting procedure were used. D-NOMINATE can break down either if the eigenvectors from the agreement matrix provide poor starts and Poole's procedure is sensitive to starts or if D-NOMINATE is sensitive to minor differences in the output from Poole's procedure. Thus, the results we report are a joint test of the sensitivity of the two procedures.

The test we carried out was to scale the 85th House and the 100th Senate (not included in our 99-Congress data set), replacing the eigenvector starts with random coordinates as input to Poole's procedure. The coordinates were again generated uniformly on [− 1, + 1]. Both the 85th House and 100th Senate were estimated in one, two, and three dimensions, with 10 simulations for each dimension.

Again we assessed fit by averaging the (45) pairwise correlations between the distances generated by the 10 simulations. We also computed the average percentage of agreement in predicted choices over the 45 comparisons of the 10 simulations. The results appear in Table A.4. The recoveries are virtually identical. The fits do become less stable as the dimensionality is increased, but this reflects only the general statistical principle that adding collinear parameters can reduce the precision of estimation. (N.B.: The roll call parameters are also estimated in the simulations.) Thus, the fit deteriorates more rapidly for the House, since there were only 172 roll call votes there as against 335 for the Senate.

These results show that D-NOMINATE combined with Poole's metric scaling procedure performs well in the joint test we carried out. We stress that both are essential to accurate recovery of the space. Particularly with two or more dimensions, D-NOMINATE does a better job of

Table A.4. Recovery from Random Starts

Dimensions	Average Correlation of Pairwise Distances from Recovered Configurations	Proportion Agreement of Predicted Choices Recovered Configurations
85th House:[a]		
1	.997	.992
	(.002)	(.003)
2	.983	.965
	(.009)	(.008)
3	.886	.902
	(.041)	(.015)
100th Senate:[b]		
1	.998	.991
	(.002)	(.002)
2	.984	.956
	(.006)	(.009)
3	.937	.918
	(.017)	(.007)

Note: Numbers in table based on 45 pairwise comparisons between 10 replications. Standard deviations are shown in parentheses.

[a]Based on 441 representatives, 172 roll calls. Number of legislators exceeds size of House because of deaths or resignations.

[b]Based on 101 senators, 335 roll calls.

recovery than the output of metric scaling. On the other hand, D-NOMINATE itself does very poorly if it begins with random starts. While the metric scaling results are not as accurate as D-NOMINATE, they are good enough to allow D-NOMINATE to converge to a solution close to the true configuration.

The "twist" problem. The above experiment showed that with little "missing" data, our procedure is insensitive to the starts. In actual practice a legislator votes on only a small slice of all roll calls in the history of a House of Congress, so there is very substantial "missing" data. "Missing" data is not a problem so long as there is, as in modern times, substantial overlap in careers. But when the membership of either House shifts very rapidly, the results become sensitive to the starts. The problem is greatest for the House in the nineteenth century. With large amounts of missing data, Poole's procedure provided poor starts to D-NOMINATE.

Our approach to this problem was to watch animated videos of the scaling results. When rapid movement induced a "twist" in the position of senators, we investigated multiplying second dimension starts for certain years (in the nineteenth century only) by -1—thereby flipping polarity. The result was to have a very slight improvement in the overall geometric mean probability and to substantially reduce the magnitudes of estimated trend coefficients in the period in question. In other words, when there is little overlap to tie the space together, it is difficult to identify the parameters of spatial movement. The results reported in the paper reflect the highest *gmp*s we have been able to achieve; they also have lower trend coefficients than solutions with slightly lower *gmp*s. (Our use of changed starts explains why readers familiar with our work may see minor differences between results here and those presented in conference papers. Experiments with different starts are a standard procedure in the estimation of nonlinear maximum likelihood models.)

REFERENCES

Bernhardt, M. Daniel, and Daniel E. Ingberman. 1985. "Candidate Reputation and the Incumbency Effect." *Journal of Public Economics* 27:47–67.

Berndt, E. K., Bronywn H. Hall, Robert E. Hall, and Jerry Hausman. 1974. "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement* 3–4:653–66.

Bullock, Charles S., III. 1981. "Congressional Voting and Mobilization of a Black Electorate in the South." *Journal of Politics* 43:663–82.

Clausen, Aage. 1973. *How Congressmen Decide: A Policy Focus.* New York: St. Martin's.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent,* ed. David Apter. New York: Free Press.

Cox, Gary W., and Mathew D. McCubbins. 1989. "Parties and Committees in the U.S. House of Representatives." Manuscript, University of California, San Diego.

Efron, Bradley. 1979. "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics* 7:1–26.

Enelow, James, and Melvin Hinich. 1984. *The Spatial Theory of Voting.* Cambridge: Cambridge University Press.

Fiorina, Morris. 1989. *Congress: Keystone of the Washington Establishment.* New Haven: Yale University Press.

Haberman, Shelby J. 1977. "Maximum Likelihood Estimation in Exponential Response Models." *Annals of Statistics* 5:815–41.

Krehbiel, Keith. 1986. "Unanimous Consent Agreements: Going along in the Senate." *Journal of Politics* 48:541–64.

Koford, Kenneth. 1989. "Dimensions in Congressional Voting." *American Political Science Review* 83:949–62.

Ladha, Krishna K. 1987. "Decision Making in the United States Congress." Ph.D. diss., Carnegie Mellon University.

Lord, F. M. 1975. *Evaluation with Artificial Data of a Procedure for Estimating Ability and Item Characteristic Curve Parameters*. Princeton, NJ: Educational Testing Service.

Lott, John R. 1987. "Political Cheating." *Public Choice* 52:169–86.

Macdonald, Stuart E., and George Rabinowitz. 1987. "The Dynamics of Structural Realignment." *American Political Science Review* 81:775–96.

McKelvey, Richard D., and William Zavoina. 1975. "A Statistical Model for the Analysis of Ordinal Level Dependent Variables." *Journal of Mathematical Sociology* 4:103–20.

Maddala, G. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Mendenhall, William, Richard Scheaffer, and Dennis Wackersley. 1986. *Mathematical Statistics with Applications*. Boston: Ducksbury.

Morrison, Richard J. 1972. "A Statistical Model for Legislative Roll Call Analysis." *Journal of Mathematical Sociology* 2:235–47.

Ordeshook, Peter C. 1986. *Game Theory and Political Theory*. Cambridge: Cambridge University Press.

Peltzman, Sam. 1984. "Constituent Interest and Congressional Voting." *Journal of Law and Economics* 27:181–200.

Poole, Keith T. 1990. "Least Squares Metric, Unidimensional Scaling of Multivariate Linear Models." *Psychometrika* 55:123–49.

Poole, Keith T., and R. Steven Daniels. 1985. "Ideology, Party, and Voting in the U.S. Congress, 1959–1980." *American Political Science Review* 79:373–99.

Poole, Keith T., and Howard Rosenthal. 1984. "The Polarization of American Politics." *Journal of Politics* 46:1061–79.

———. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29:357–84.

———. 1987. "Analysis of Congressional Coalition Patterns: A Unidimensional Spatial Model." *Legislative Studies Quarterly* 12:55–75.

———. 1988. "Roll Call Voting in Congress, 1789–1985: Spatial Models and the Historical Record." In *Science at the John von Neumann National Supercomputer Center: Annual Research Report Fiscal Year 1987*. Princeton, NJ: John von Neumann Center.

———. 1989a. "Color Animation of Dynamic Congressional Voting Models." Working Paper 64-88-89, Carnegie Mellon University.

———. 1989b. "The D-NOMINATE Videos." VHS Tape of Computer Animations. GSIA, Carnegie Mellon University.

Poole, Keith T., and Stephen E. Spear. 1990. "Statistical Properties of Metric Unidimensional Scaling." Working Paper 88-89-94, Carnegie Mellon University.

Smith, Steven S., and Marcus Flathman. 1989. "Managing the Senate Floor: Complex Unanimous Consent Agreements." *Legislative Studies Quarterly* 14:349–74.

Torgerson, W. S. 1958. *Theory and Methods of Scaling*. New York: Wiley.

Van Doren, Peter. 1990. "Can We Learn the Causes of Congressional Decisions from Roll Call Data?" *Legislative Studies Quarterly* 15:311–40.

Weisberg, Herbert F. 1968. "Dimensional Analysis of Legislative Roll Calls." Ph.D. diss., University of Michigan.

Wright, B. D., and G. A. Douglas. 1976. "Better Procedures for Sample-Free Item Analysis." Research Memorandum no. 20, Statistical Laboratory, Department of Education, University of Chicago.