ECONOMICS AND POLITICS 0954-1985
Volume 6 July 1994 No. 2

DIMENSIONAL SIMPLIFICATION AND ECONOMIC THEORIES OF LEGISLATIVE BEHAVIOR

KEITH T. POOLE AND HOWARD ROSENTHAL*

> In many many places, the political parties and political contests seem to array themselves along a single dimension. If you look at their programs in detail it always turns out that it's much more complicated than that. But it would appear that in the minds of the voters they are literally frequently arrayed on a single continuum which, as a result of the organization of the French chamber of deputies, is usually referred to as a left/right.
>
> Tullock, 1991, p. 127

IN 1985, we (Poole and Rosenthal, 1985a) issued a working paper with the deliberately provocative title "The Unidimensional Congress". We argued that in the modern period from 1919 to 1984 over 80 percent of Congressional roll call voting decisions, even on close votes, could be accounted for by a simple unidimensional "liberal-conservative" or "left-right" model. While the world may be complicated, there is, as Tullock suggests, a high degree of "dimensional simplification".

Our view of dimensional simplification has enjoyed some acceptance (see, for example, Kiewiet and McCubbins, 1990, and Snyder, 1991). Critics have acknowledged that "Without much debate, researchers have accepted the view that a single liberal/conservative dimension accounts for much of congressional-roll call voting" (Wilcox and Clausen, 1991, p. 393), even though we proposed "an interpretation greatly at odds with earlier understandings of the *complex* fabric of congressional policy making" (Wilcox and Clausen, 1991, p. 394, emphasis ours).

The most extensive criticism of our viewpoint has come from Kenneth Koford (1989, 1991). His work has included a recent (Koford, 1990) piece in this journal. In this article (but see also Poole and Rosenthal, 1991b), we deal with the body of argument he has developed.

In general form, the spatial model of roll call voting claims that utility on a roll call can be represented as a negative monotonic function of the metric $(z-x)'A(z-x)$ where x is an s-vector representing the legislator's ideal point, z an s-vector representing one of the roll call alternatives being voted on, and s the dimensionality of the space. On each roll call, there are two alternative points, one corresponding to a "Yea" vote and another to the "Nay" vote. The s by

*Rosenthal is grateful for financial support provided by National Science Foundation grant # BNS-8700864 during his stay at the Center for Advanced Study.

s A matrix represents variations in salience (diagonal elements) and correlation (off-diagonal elements) across dimensions. The model can be broadened, as is done in our empirical work, to allow for an additive stochastic component to utility.

Koford has always argued on the basis of a non-stochastic model of *perfect* spatial voting. First, within the context of equal salience weights (A = I for all legislators) he claimed that a one dimensional model would classify well even if the true space were of a higher but low (2 to 7) dimensionality simply because one dimensional projections will classify well. Second, he claimed that we understated the importance of higher dimensions by only reporting the *incremental* classification gain brought about by using an additional dimension. He argued that projecting onto these dimensions would show substantial classification power, similar to what we had obtained with one dimension.

After developing a unidimensional scaling algorithm and writing "The Unidimensional Congress" in the early to mid eighties, we developed a multidimensional algorithm that permitted investigation of these first two points. (See Poole and Rosenthal, 1991a, 1991b). We found that even with as many as 50 dimensions, there are only feeble improvements over the original 80 + percent classification. Consequently, a low dimensional model of *perfect* (errorless) spatial voting, as proposed by Koford, just didn't fit empirical observations for Congress. These results disposed of his first argument, that strong one dimensional classifications were obtained as a result of the ability to project from a low (2 to 7) dimensional world characterized by perfect spatial voting. Moreover, projection onto our empirical second dimension gave classifications barely better than the average size of majorities (Poole and Rosenthal, 1991a, Figure 4). We thus also concluded that the second critique was inconsistent with empirical observations.

Our initial claim of "unidimensionality" was in fact reinforced by the multidimensional analysis. Although a second dimension has been important in various periods of American history, particularly 1830–1850 and 1937–1970, the first dimension has always been dominant, except for two brief pre-Civil War episodes where the spatial model fails entirely in any dimensionality. We allowed the new findings to modify our earlier conclusions, somewhat tongue in cheek, by describing the space as having "1.5 dimensions" (Poole and Rosenthal, 1991b, p. 232).

These multidimensional findings disposed of Koford's initial critique. He acknowledged our points (Koford, 1991, p. 962) but proposed reconciling the empirical observations with a *perfect* spatial voting model that allowed for individual A's. We deal here with the variable salience weight approach and other aspects of the dimensionality problem.

In addition to raising the dimensionality issue, Koford also claimed a greater role for economic voting on constituency interests than we do. In other research (Poole and Rosenthal, 1985b, 1991c), we have demonstrated that the spatial model outperformed other attempts to explain the same votes, or subsets of votes, with limited dependent variable regressions using a panoply of variables that purported to measure the economic preferences of constituents. There is little to add here from our previous exchange, except for some concluding remarks.

1. SALIENCE WEIGHTS AND DIMENSIONALITY

Our D-NOMINATE procedure assumes that all legislator decisions depend on the unweighted Euclidean distances between the legislator and the roll call outcomes. A more general model would allow for these decisions to be a function of weighted Euclidean distance, each legislator having an individual A (see Ordeshook, 1986, pp. 24–25). Koford claims that our result – a good fit in one dimension but not much better fit with many dimensions – could result from differential salience of issues, expressed formally as weighted Euclidean distance. But he supports that claim solely on the basis of a highly stylized example.

Clearly, our recovery will be distorted by differential salience weights. But the fundamental question is whether the bias is sufficient to distort our assessment of the dimensionality of the issue space. We conducted a simulation to investigate this issue. We generated errorless data for 440 legislators voting on 220 roll calls in a true two dimensional environment. The legislator ideal points and roll call outcomes were randomly and independently drawn on each dimension from $[-1,+1]$ and the salience weights on each dimension for each legislator were randomly and independently drawn from $[0.5,1.5]$. The off-diagonal elements of the A's were all zero.

To motivate the $[0.5,1.5]$ range, we asked ourselves what would be the nature of the low dimensional space that Koford envisaged. It could not be a "what's good for my district" space. For such a space we would need many dimensions, perhaps even 434 for the House of Representatives. Clearly, the issues must be broader than "build a breeder reactor on the Clinch" or "dig a ditch from the Tennessee to the Tombigbee." They in fact should be close to the issues in the five dimensional space posed by Clausen (1973), broad issues like social welfare, government management, foreign affairs and defense, civil liberties, and agriculture. We thought it unlikely that a representative could, say, care only about domestic policy and place no weight on foreign affairs. Consequently, we allow the salience weights to range only over a factor of 3. The range of ideal points is similar to that produced by D-NOMINATE for actual roll call data.

If there were truly errorless voting in a two-dimensional environment with unweighted distance, D-NOMINATE, which assumes voting with error, would produce an imperfect fit and then, similar to standard probit and logit analysis, "blow up" as it reaches perfect classification in two dimensions (Poole and Rosenthal, 1991b, p. 956–957).

What happens when there is weighted Euclidean distance?

We ran three replications of the simulation. As would be expected from the large "N" of 440 by 220 simulated observations, the results were highly stable across replications. A one-dimensional D-NOMINATE scaling correctly classified

80 to 81 percent of the individual decisions. A two dimensional model correctly classified 96 to 97 percent, not perfect given the bias introduced but certainly far greater than the roughly 85 percent (Poole and Rosenthal, 1991a, Table 1) found with Congressional data. Estimating a four dimensional model only raised classification by another percentage point. Data analysts would surely see a two dimensional space in these results. In other words, if the true Congressional data were errorless spatial voting, D-NOMINATE should find the correct dimensionality, even if salience weights are variable.[1] The fact that extra dimensions make only slight improvements to classification of Congressional data cannot be a matter of differential salience. In any event, implementing weighted Euclidean distance would not be parsimonious science, since it would, even if no "off-diagonal" weights were estimated, double the parameter space.[2]

While the dimensionality D-NOMINATE recovers should be robust to variation in saliency, we believe that the saliency approach is misguided. In 1985, for example, the House of Representatives took 439 roll call votes. By the time the first 34 had occurred, the topics covered included: election of the speaker, disputed elections, the budget, famine in Africa, cash awards to cost-saving federal employees, farm aid, interstate highways, domestic hunger, narcotics abuse, fisheries, internal funding of the house, and MX missiles. True, in this enormous substantive diversity, topics can vary in saliency. But if the topics can be grouped in a way that leads to a small number (more than "1.5" but less than seven) of dimensions being a useful model, all the dimensions are likely to contain some topics of importance to the congressmen. Low dimensionality, be it just one or as many as Koford's seven, would appear to be incompatible with strong variations in salience.

2. WHAT DO DIMENSIONS MEAN?

In addition to questioning whether the space is as low dimensional as we claim, a further concern of Koford's (1990, p. 61) is that we do not recover "true" dimensions. The scenario he has in mind is one where there is a true two dimensional world and substantive, orthogonal issue axes are known a priori. Assume that, as a result of agenda control processes, all votes partitioned the space via cutting lines at an angle of 135 degrees to the first dimension (see

[1] D-NOMINATE's ability to recover the true dimensionality should be robust to the range of variation of salience weights. What will be sensitive to the range is the percentage correctly classified. The simulations reported here were performed in the fall of 1989. At the present time, we no longer have access to a Cyber supercomputer that runs our multidimensional code. The unidimensional code has been rewritten and tested on VAX mainframes and 386 and 486 machines running under OS2. A documented copy is available by sending a formatted 3.5" high density diskette to Keith T. Poole, GSIA, Carnegie-Mellon University, Pittsburgh, PA 15213. Work on a generic multidimensional program is in progress. Stop press. The multidimensional program is available.

[2] In this sense, correcting for variable salience is similar to correcting for heteroscedasticity in regression models. While homoscedasticity is a very restrictive assumption, using up degrees of freedom with a heteroscedastic model is unwarranted unless one believes, a priori, that the heteroscedasticity is "very substantial".

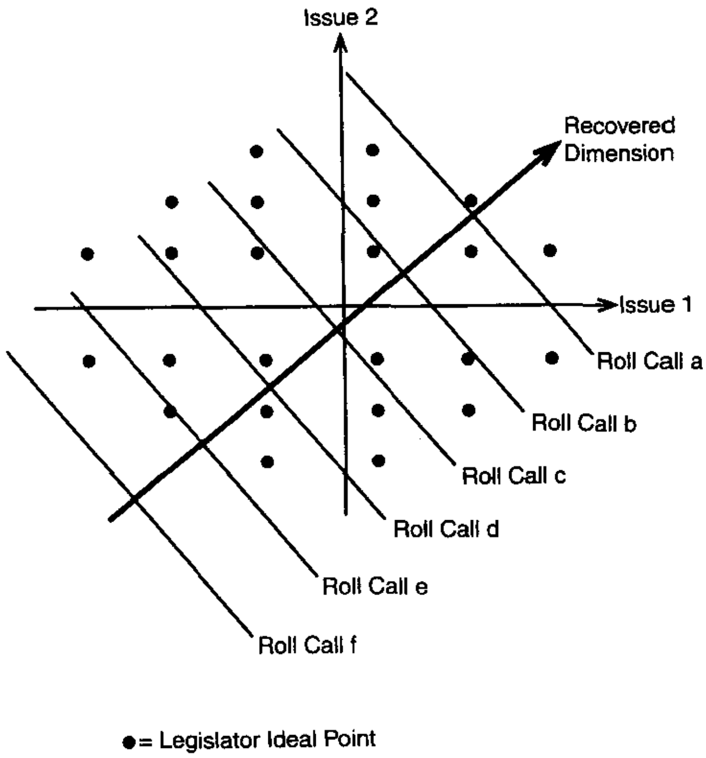

Figure 1. The recovered dimension when roll calls are parallel.

Figure 1). Then we would recover a one dimensional space at an angle of 45 degrees to the first dimension. The recovered dimension would not be a true dimension.

This point is hardly interesting. Like eigenvalue/eigenvector extraction, D-NOMINATE, which has to operate without such a priori information, tries to do as much business as possible on the first dimension. Finding that a single empirical dimension accounts for the data is an important regularity, one that merits much further research. What D-NOMINATE recovers is a reduced form, after agenda control, log rolls, gatekeeping, and other factors have whittled away a complex world. The recovered dimensionality will always be the minimum of the dimensionality of the roll calls and the legislators. In our stylized example, the roll calls are truly unidimensional. D-NOMINATE does not underestimate the dimensionality of the space that is politically relevant for roll call voting. Moreover, consider a slightly more complicated world, one with a preponderance of 135 degree cutting lines and other votes more or less evenly spread out from 0 to 180 degrees. In this case, D-NOMINATE will find that voting is two dimensional but the recovered dimensions will not be the "true" ones. This is a trivial "problem". The "true" dimensionality is recovered, and external, a priori information can be used for an arbitrary rotation of the recovery.[3]

3. THE CORRECT NULL HYPOTHESIS

While Koford's models of projection and variable salience are not good candidates for explaining the finding of low dimensionality, he does raise the issue of what is the relevant null model for assessing fit. Previously, we had argued for a null model of an infinite number of "equal" dimensions while Koford had argued (1990, p. 63) for "a null hypothesis close to the hypothesis being tested". With our model, there would be a benchmark error rate of 50%, with his it would be as much as 75%. Clearly, 80+ % looks better against 50 than against 75.

Nonetheless, Poole, Sowell, and Spear (1992) developed a null model of errorless multidimensional spatial voting much in the spirit of Koford's and found that the observed pattern of classification errors from D-NOMINATE is consistent with a low dimensional voting space. In particular, they were able to solve for the projection of perfect voting onto one dimension and the resultant classification error. This enabled them to derive the probability density function over the classification errors which in turn can be used to calculate the likelihood that a sample of votes was drawn from an N-dimensional hypersphere. Since *perfect* voting in one dimension would imply zero classification errors, the presence of a single error would make the likelihood zero. Obviously, we reject the joint hypothesis of unidimensionality and perfect voting. For dimensionality greater than one, the likelihood function peaks at two dimensions in all Congresses since 1853. This result shows that, following Koford's reasoning about null models to its logical limit, perfect voting cannot be occurring in spaces above two dimensions.

Koford's use of perfect voting null models to assess our work is, in any event, misplaced, since (i) the D-NOMINATE model does not assume errorless voting and (ii) as mentioned above, there is no empirical evidence for errorless voting in any reasonable dimensionality.

Moreover, benchmarks are usually defined on the basis of empirical distributions and not the theoretical ones used by Koford (1989) and Poole, Sowell, and Spear (1992). Consider the basic regression model, where the standard null model is $\beta = 0$. The benchmark then becomes the empirical variance of the dependent variable. The analogy in classification would be the empirical distribution of minorities. Corresponding to "predict the mean of the dependent variable", the null prediction would be "predict everyone votes with the majority on the issue". Since, for the roll calls in our analysis, average minorities

[3] Indeed, this is a decades old technique in multidimensional scaling analysis. For examples and a bibliography, see Kruskal and Wish (1978).

are around 40%, our roughly 20% one dimensional error rates are reasonably impressive. What is also important is that error rates go down by only about 3 percent when a second dimension is added. About 15 percent of the individual votes cannot be accounted for by Euclidean models of reasonable dimensionality.

4. WHAT PRODUCES LOW DIMENSIONALITY?

Our empirical finding of low dimensionality does contrast with the rich diversity of substantive issues faced by legislatures. Attempts to explain this result include:

(1) It is an artifact. Scalings always show a high fit in one dimension. This is disproved by the fact that D-NOMINATE does not fit the data for several pre-Civil War Congresses (Poole and Rosenthal, 1991a, Figure 5).
(2) Bias due to the "incidental parameters" problem. (The incidental parameters are the additional Euclidean coordinates estimated for every additional legislator and roll call.) Our simulation results show that the recovery of D-NOMINATE is excellent when data are generated according to the assumed statistical model (Poole and Rosenthal, 1991a, pp. 273–276).
(3) It results from projection from a true higher dimensional space (Koford, 1989, 1990; Poole, Sowell, and Spear, 1992). See above.
(4) Weighted Euclidean distance (Koford, 1991). See above.
(5) Agenda control exercised by Congressional committees. This explanation was proposed by Snyder (1992). While gatekeeping may indeed exclude certain issues from voting, our rebuttal to Snyder, like the one to Koford, is based on the inability of his model to account for the data (see Rosenthal, 1992 for full details).
(6) Coalition formation, logrolling (Ferejohn, 1986), and avoidance of transaction costs (Koford, 1990). In particular, the role of committees in coalition formation is perhaps at least as important as the gatekeeping function. Specific constituency benefits which, if voted upon, would lead to high dimensionality, are typically written into the details of legislation in committee, while roll call votes tend to be restricted to packages. All these factors are undoubtedly important, but we as yet do not have formal models that lead to testable predictions. In particular, Koford (1990) is not helpful on this score.
(7) Voters themselves, at least in the aggregate, also exhibit low dimensionality in their political preferences. This reconciles the constituency theory of voting with the ideological theory. This very important observation was made by Snyder (1991). Snyder applied NOMINATE to the California Assembly and found, echoing our results for Congress, that a one dimensional model explained 90 percent of the decisions. He also found that a low dimensional model accounted for aggregate voting on initiatives, even though the initiatives exhibited a wide diversity of economic content. The scaling of assembly districts' voting patterns correlated highly with the positions calculated from the roll call votes of their representatives.

5. WHAT PRODUCES ERROR?

Future research must also explain why a significant fraction of roll call votes are not accounted for by the Euclidean model. Possibilities include:

(1) Strategic voting. This may not be a problem with certain binary amendment trees in a complete information setting. Voters still vote along Euclidean lines, except that they replace the locations of the ostensible alternatives with the locations of the sophisticated equivalents.
(2) Vote trading. For example, Groseclose (1991) has identified the Byrd Amendment to the Clean Air Act as one where there was competitive bidding for votes. Our model has only modest success with this roll call.
(3) Perceptual error/differential information. Lack of information on issues by legislators is the focus of work by Gilligan and Krehbiel (1987). Ladha (1991) has developed a method for estimating spatial locations when legislators vary in their perceptions of roll call outcomes. Similar in spirit to the NOMINATE model, his model has a more appealing theoretical basis for the error term. Unfortunately, his model cannot be extended beyond one dimension.

6. CONSTITUENCY VOTING AND IDEOLOGICAL VOTING

Economists frequently analyze roll call votes using aggregate constituency measures. (Ample references and examples can be found in Poole and Rosenthal 1985b, 1991c, and Koford, 1990, 1991.) If some combination of coalition formation, agenda control, and cognitive limits of voters forces roll call voting to appear "as if" it were ideological, economists have been using severely misspecified models. The problems are exacerbated if the low dimensional pattern is disrupted by vote trading, strategic voting, and cognitive and informational limits of legislators.

Vote trading is a particularly important consideration in dealing with economic models of roll call voting. Even if we grant that legislators are mainly motivated, through the reelection constraint, by constituency economic interests, the rational pursuit of these interests can lead to trades. Thus, measuring economic interests solely in terms of the substance of the issue of the vote, is, in a fully rational context, a potential misspecification. On the other hand, the D-NOMINATE scores, since they are based on the legislator's votes throughout his career, at least partially reflect trades.

Moreover, to repeat an old point, the relevant constituency may hardly be captured by aggregate variables. The California Desert Protection Act has languished in Congress for years. The legislation concerns only California. Thus the state's senators are unlikely to have perceptual problems on the issue or to have their votes for sale. Yet the legislation supported by Senator Cranston was always opposed by his Republican cosenator, first Wilson and later Seymour.

Unless economists greatly refine their measurements (except for the political party dummy variable cop out), they will fail to capture these intraconstituency variations in roll call voting.

The specification problems we have outlined suggest why "economic models" of roll call votes typically have low explanatory power once "ideology" has been taken into account.

In summary, dimensional simplification, where a complex world gets boiled down into relatively simple voting patterns, is a strong empirical regularity that should not be ignored.

KEITH T. POOLE
Carnegie-Mellon University
5000 Forbes Avenue
Pittsburgh
Pennsylvania 15213
USA

HOWARD ROSENTHAL
Princeton University
Princeton
New Jersey 08544
USA

REFERENCES

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

Ferejohn, John, 1986, Logrolling in an Institutional Context: The Case of Food Stamps in Leroy Resielbach et al., eds., *Congress and Policy Change* (New York: Agathon Press).

Gilligan, Thomas and Keith Krehbiel, 1987, Collective Decision-Making and Standing Committees: An Information Rationale for Restrictive Amendment Procedures. *Journal of Law, Economics and Organization* 3, 287–335.

Groseclose, Timothy, 1991, The Market for Favors and Votes in Congress: A Test of Two Theories (manuscript Graduate School of Business, Stanford University).

Kiewiet, D. Roderick and Matthew McCubbins, 1990, *The Logic of Delegation: Congressional Parties and the Appropriations Process* (Chicago: University of Chicago Press).

Koford, Kenneth, 1989, Dimensions in Congressional Voting. *American Political Science Review* 83, 949–962.

Koford, Kenneth, 1990, Dimensions, Transactions Costs and Coalitions in Legislative Voting. *Economics and Politics* 2, 59–82.

Koford, Kenneth, 1991, On Dimensionalizing Roll Call Votes in the U.S. Congress (controversy with Keith T. Poole and Howard Rosenthal). *American Political Science Review* 85, 955–975.

Kruskal, Joseph B. and Myron Wish, 1978, *Multidimensional Scaling* (Beverly Hills: Sage Publications).

Ladha, Krishna K., 1991, A Spatial Model of Legislative Voting with Perceptual Error. *Public Choice* 68, 151–174.

Ordeshook, Peter C., 1986, *Game Theory and Political Theory: An Introduction* (New York: Cambridge University Press).

Poole, Keith T. and Howard Rosenthal, 1985a, The Unidimensional Congress: 1991–84 (presented at the APSA annual meeting, August).

Poole, Keith T. and Howard Rosenthal, 1985b, The Political Economy of Roll-Call Voting in the 'Multi-Party' Congress of the United States. *European Journal of Political Economy* 1, 45–58.

Poole, Keith T. and Howard Rosenthal, 1991a, Patterns of Congressional Voting. *American Journal of Political Science* 35, 228–278.

Poole, Keith T. and Howard Rosenthal, 1991b, On Dimensionalizing Roll Call Votes in the U.S. Congress (controversy with Kenneth Koford). *American Political Science Review* 85, 955–975.

Poole, Keith T. and Howard Rosenthal, 1991c, The Spatial Mapping of Minimum Wage Legislation in Alberto Alesina and Geoffrey Carliner, eds., *Politics and Economics in the 1980s* (Chicago: University of Chicago Press).

Poole, Keith T., Fallaw B. Sowell, and Stephen F. Spear, 1992, Evaluating Dimensionality in Spatial Voting Models. *Mathematical and Computer Modeling* 16, 85–101.

Rosenthal, Howard, 1992, The Unidimensional Congress Is Not the Result of Selective Gatekeeping. *American Journal of Political Science* 36, 31–35.

Snyder, James M. Jr., 1991, The Dimensions of Constituency Preferences: Voting on California Ballot Propositions, 1974–1990 (manuscript, Chicago: University of Chicago).

Snyder, James M. Jr., 1992, Committee Power, Structure-Induced Equilibria, and Roll Call Votes. *American Journal of Political Science* 36, 1–30.

Tullock, Gordon, 1991, Duncan Black: The Founding Father. *Public Choice* 71, 125–128.

Wilcox, Clyde, and Aage Clausen, 1991, The Dimensionality of Roll-Call Voting Reconsidered. *Legislative Studies Quarterly* 16, 393–406.