Dimensionality in Congress: Alternative Sources of Low-Dimensionality in Scaled Roll-Call Analyses*

John H. Aldrich
Department of Political Science
Duke University

Jacob M. Montgomery†
Department of Political Science
Washington University in St. Louis

David Sparks
Department of Political Science
Duke University

July 28, 2011

ABSTRACT

There exists a general consensus that the preferences of political elites can be represented as existing in a low-dimensional space, where the primary dimension is the left-right or liberal-conservative spectrum. However, the statistical evidence used to support this claim is largely driven by restrictive assumptions that precede estimation. If these assumptions are not met, scaling procedures may be biased in favor of finding a small number of dimensions. In this paper, we conduct simulation experiments altering assumptions regarding the dimensionality of the policy space, the distribution of preferences, and the voting behavior of legislators. We scale these simulated roll-call matrices and compare the output to analyses the roll-call record for the U.S. Senate. We find that the estimation of a low-dimensional space does not necessarily imply a truly low-dimensional world. Rather, modest polarization of preferences and quasi-stable distributive coalitions will both generate strong and misleading evidence in favor of low-dimensionality.

*A previous version of this paper was presented at the 2009 Annual Meeting of the Southern Political Science Association in Atlanta, GA and the 2010 Annual meeting of the American Political Science Association in Washington, D.C. We are grateful for comments from Frances Lee, John Patty, and helpful audiences at Duke University and Washington University in St. Louis.

†Corresponding author. Address: Campus Box 1063, One Brookings Drive, St. Louis, MO, USA, 63130-4899. Email: jacob.montgomery@wustl.edu

1. INTRODUCTION

One of the accepted truisms of American politics research is that we can effectively describe public policy with only one or two dimensions. Indeed, scholars and pundits routinely conceptualize policies and political figures as fitting onto a single underlying liberal-conservative continuum. Only on occasion is this is supplemented with a second dimension (e.g., "social issues"). Indeed, with several notable exceptions discussed below, the bulk of contemporary research on American politics implicitly or explicitly accepts Poole and Rosenthal's famous conclusion that "one-and-a-half" issue dimensions adequately encapsulates every era of the nation's political history and that the contemporary Congress is virtually or actually unidimensional (Poole and Rosenthal 1984, 1997, 2007; Wilcox and Clausen 1991).

There are good reasons for conceptualizing American political competition as low-dimensional. First, the low-dimensional account provides tremendous theoretical and descriptive leverage. If we accept that policies, voters, candidates, and incumbents all exist in a simple space, we can construct a vivid unifying account of America's political history and its current debates. This one simplifying assumption provides the basis for drawing a straight line between the political choices of the public with those of candidates, parties, and officeholders. We can construct both narratives and mathematical models for how the preferences of voters are translated in predictable ways into policy outcomes. Moreover, without this assumption, no such clear narrative is possible. In a high-dimensional world, the upward aggregation of voter preferences through elections, parties, and legislators into public policy is necessarily complex and potentially chaotic.[1]

Second, analyses of roll calls in the U.S. Congress have reinforced the low-dimensionality assumption. Scholars have applied various statistical scaling methods to the roll-call record and almost universally interpreted their results as supportive of this conjecture. In their foundational work, for instance, Poole and Rosenthal apply their particular scaling procedures and analyze all of American political history as an inherently low-dimensional phenomenon (1997; 2007; McCarty,

[1]Indeed, we can make *no* clear predictions in a high-dimensional setting without the addition of numerous – often contentious – assumptions (McKelvey 1976, 1979; Schofield 1978).

Poole and Rosenthal 2005). However, this is *not* particular to the Poole and Rosenthal technology. Rather, nearly *every* scaling procedure applied to the broader congressional roll-call record has lead scholars to similar conclusions (e.g., Clinton, Jackman and Rivers 2004; Poole 2000, 2005; Bonica 2011).[2] Indeed, this widespread agreement across methodologies has lead many to claim that "one-and-a-half" dimensions is a feature American politics that statistical analysis uncovers.

In this paper, we challenge this low-dimensionality conjecture – or at least the statistical evidence used to justify it. We use Monte Carlo simulations to show that the statistical results used to back these claims do not, in fact, support such strong assertions. While it is true that *if* the data generating process behind roll-call votes is one- or two-dimensional, then widely used scaling procedures will reproduce that low dimensionality. However, we further show that this relationship is not reciprocal. That is, it is also true that if the data are generated from a high-dimensional political world, then it is possible that even the best scaling procedures can falsely produce results suggesting a low-dimensional space. In addition, we show these misleading low-dimensional findings are not only possible but likely under many plausible circumstances.

In essence, our point is that *low-dimensional scaling results do not imply low dimensionality in the actual political world*. We show that there is a sufficient relationship. That is, if the world is actually low dimensional, then the scaling results will be low dimensional. However, we show that there is no necessity (i.e., only if) let alone necessity and sufficiency (i.e., if and only if). Moreover, in what follows, we go several steps farther than simply demonstrating that high-dimensional data *can* result in a low-dimensional scaled space. This is no "just so" story. Instead, we demonstrate that under many likely circumstances we would *expect* scaling procedures to yield erroneous and misleading low-dimensional statistical results.

We begin with brief discussions of past research on the dimensionality of Congress and methods for establishing dimensionality. In Section 3, we present the details of our Monte Carlo simulations, and we explain our results in Section 4. We conclude with a discussion of the broader implications of our findings.

[2]However, Heckman and Snyder (1997) develop a scaling procedure that typically suggests larger numbers of dimensions.

2. DERIVING DIMENSIONALITY FROM ROLL-CALL SCALING

For nearly a generation, congressional research has advanced empirically on estimates of member "ideologies" generated from scaling analyses of the roll-call record. Indeed, one of the first systematic statistical advances in political science was the creation of the Rice index (1925). More recently, various multidimensional scaling techniques have been invented, with Keith Poole and Howard Rosenthal's (PR) extensive work, yielding the W-NOMINATE procedure among others, being the most well known and most extensively used (Poole and Rosenthal 1997, 2007). Their procedure represents a significant advance that fully deserves its popularity and wide utilization. Yet, PR's scaling techniques and their numerous cousins[3] are no more appropriate for answering every question in congressional research than are, say, Likert scales appropriate for answering every question in political behavior (VanDoren 1990). Like all measurement techniques, W-NOMINATE and its kin are based on specific assumptions that condition the scope of their applicability.[4]

Our claim in this paper is that the task of identifying the true number of dimensions in the Congress lies outside the scope of W-NOMINATE and other widely applied scaling procedures. Or, to be more precise, the observation of a small number of dimensions resulting from the application of such scaling procedures is insufficient to support the inference that the true number of dimensions is actually small. As we show, a strong bias towards low dimensionality is sometimes present even when the stringent behavioral assumptions of the measurement models are met (e.g., when preferences are bimodally distributed). More importantly, under many other plausible assumptions (e.g., distributive coalitions), the techniques are likely to produce very misleading results biased in favor of low dimensionality.

[3]See also Heckman and Snyder (1997), Clinton and Meirowitz (2001), Martin and Quinn (2002), Clinton, Jackman and Rivers (2004), Bafumi et al. (2005), Poole (2005), Peress (2009), and Bonica (2011).

[4]Because they are based on a specific data source (recorded votes cast on the floor), the data themselves also limit their range of applicability and the nature of inferences they support.

2.1. *One dimension or many?*

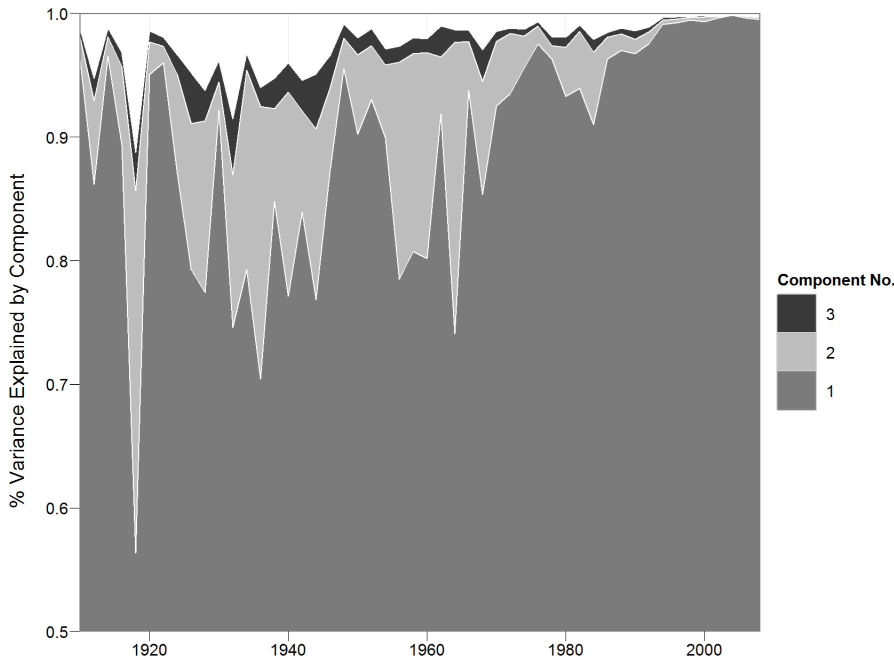
Our challenge to the low-dimensionality claim may seem quixotic to some scholars of Congress because, at first blush, the evidence in favor of a simple political space appears so compelling. Consider, for example, the data displayed in Figure 1. The figure reports the percentage of variance in the Senate roll-call record explained by the first, the first two, and the first three dimensions of W-NOMINATE for each year from 1910 to 2010. Note that these three dimensions cumulatively account for over 90% of the variance in *every* Congress. Indeed, for most of this period, even the third dimension seems superfluous. It rarely explains more than a few percentage points of the variance. After World War II, the first two dimensions always combine to account for at least 95%. Moreover, while the second dimension adds substantially to the variance explained in the 1950s and 1960s (about 15%), it explains little after the 1970s. Indeed, only in a few years does the first dimension account for less than about 85% of total variance in the roll-call record, and it accounts for at least 95% in each of the last 15 years. Thus, the pattern *seems* to support the conclusion that American public policy can be adequately characterized by one or one-and-a-half dimensions.

However, recent scholarship has raised serious doubts about the low-dimensionality of the roll-call record.[5] Crespin and Rohde (2010) and Roberts, Smith and Haptonstahl (2009), for instance, analyze roll calls in specific issue areas and uncover substantial evidence in favor of a much larger number of dimension. Jenkins (1999; 2000) uses data from the Confederate Congress to show that, in the absence of strong political parties, the structure, stability, and low-dimensionality of the roll-call record evaporates. Similar findings exist for state legislators without strong two-party systems (Welch and Carlson 1973; Wright and Schaffner 2002).[6] In a comprehensive review of the content of roll calls in the U.S. Senate, Lee (2009) argues that much of the structure of scaling estimates – including the evidence supporting low dimensionality – is actually a result of partisan

[5]Roberts, Smith and Haptonstahl (2009) provide additional discussion of recent scholarship challenging the low-dimensional assumption. See also Heckman and Snyder (1997), Hurwitz, Moiles and Rohde (2001), Talbert and Potoski (2002), and Wright and Schaffner (2002).

[6]Lee (2009, p 43) also notes that W-NOMINATE performs very poorly in the two historical periods when there were not two functioning political parties.

Figure 1: Percent of overall variance in the Senate roll-call record explained by the first three components of W-NOMINATE by year from 1910-2010.



Most of the variance in each year seems to be explained by only one or two dimensions. *Note that the proportions explained by the first component as these data-points will be reproduced in subsequent plots to provide empirical reference points.*

"teamsmanship" and has nothing (or very little) to do with members' ideological positions.

This work parallels a growing body of research that investigates how violations of NOMINATE's stringent assumptions might result in systematic patterns of errors and misclassifications of roll-call votes.[7] These findings suggest that NOMINATE scores and related measures are not detecting member preferences so much as providing a summary of each member's observed behavior. The origins of that behavior are manifold, and might include member preferences but also voting procedures, pressure from party leaders, and the influence of outside actors.[8]

In total, this stream of scholarship has not so much raised doubts about the value of scaling procedures or their "correctness" as it has suggested that caution is needed in interpretation. In this

[7] In addition to those cited above, an incomplete list of recent empirical work in this vein would include Snyder and Groseclose (2000), Ansolabehere, Snyder and Stewart (2001), Cox and Poole (2002), Roberts (2007), Smith (2007), Masket (2007), Patty (2008), and Brady and Rohde (2009).

[8] For a similar argument in the context of judicial voting see Ho and Quinn (2010).

article, however, we go several steps farther. We do not analyze subsets of roll calls, suggest alternative methods for detecting underlying dimensions, or focus on the interpretation of recovered scale estimates. Rather, we show that the patterns shown in Figure 1, the patterns most commonly used to justify the low-dimensionality conjecture, are themselves consistent with the true number of latent policy dimensions being either small or large. In addition, we show that this will occur under many plausible – indeed theoretically and empirically likely – circumstances.

2.2. *Establishing dimensionality*

In Sections 3 and 4, we describe and present the results of Monte Carlo simulations that support our claims. Before moving on, however, it is worth stepping back to briefly consider some of the difficulties inherent in establishing the dimensionality of any given dataset.

Dimensionality is empirical, not theoretical: One possible method for examining the dimensionality of the roll-call record would be to draw on theory to specify conditions under which a certain number of dimensions should affect voting. To the best of our knowledge, however, none of the widely applied scaling procedures (and virtually or actually none of the research that uses their estimates) takes the number of dimensions as a parameter to be derived from theory. Instead, it is only a maintained assumption that preferences are defined over a space with a finite number dimensions.[9]

PR, for example, embed their scaling in the spatial model of legislative behavior that has been the workhorse of positive theory since Black (1948), Downs (1957), and Enelow and Hinich (1984). They follow a standard set of assumptions ranging from the substantive (e.g., sincere voting) to the technical (e.g., preferences are defined over a policy space measured via a Euclidean metric).[10] However, *nothing* is assumed about the dimensionality of the space other than that there

[9]Moreover, the spatial model is no help in defining the dimensions. The *meaning* of the dimensions must also be assumed, asserted, or derived in some other fashion.

[10]The Euclidean metric is the most common with utility functions defined as strictly Euclidean, weighted-Euclidean, or a monotonic transformation of Euclidean. A more general Minkowski metric is sometimes employed.

are p dimensions, where $p \in [0, 1, 2, \ldots, \infty)$.

Our point here is not only to remind the reader of these well known considerations, but to remind the reader that the theory upon which most applications of scaling rely is not a source for addressing these questions. The theory is one of choice, not a theory of the nature of preferences.

Yet, these hypothesized policy dimension are of central importance to spatial theory. We typically assume that citizens have preferences over policies that inform their voting choices. Candidates take positions in that same policy space to attract support. Once elected, officeholders vote for legislation – defined as points in the same space – to alter the status quo. Thus, this ill-defined policy space is the very stage upon which the drama of democratic politics is played out. To understand politics in terms of spatial models, it is crucial to get the space right. Yet, its shape does not itself fall from the theory.

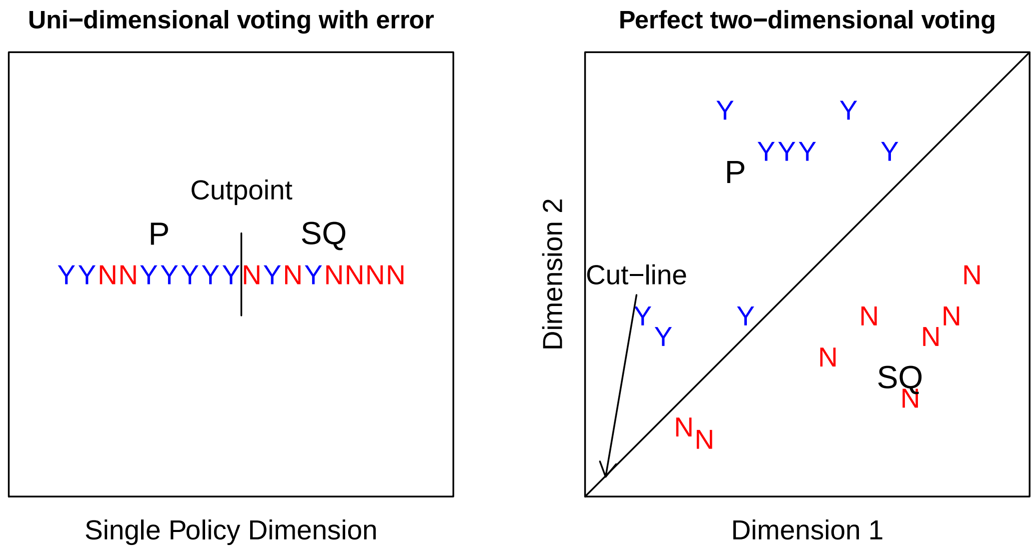
Empirical evidence on dimensionality is interpretive, not objective: Given that dimensionality is not something that flows from the spatial models behind most scaling procedures, how do we – how does anyone – know how many dimensions are appropriate? There are two main methods for determining the appropriate number of dimensions, both of which suffer from a common problem; the number of dimensions is a subjective judgement by the researcher.

One approach is to scale the data under a number of different dimensionality assumptions and to compare the results. An inference, to the extent that this can be said to be an inference, is then made by comparing the estimates from a model with a maintained assumption that there is a single dimension to one that assumes there are exactly two dimensions, and those, in turn, to the model that assumes there are exactly three dimensions, and so on. Thus, inferences rely in essence on statistics similar to those reported in Figure 1. The scholar fits multiple alternative models, examines the results, and determines which model seems the most "adequate."

However, it is important to realize that this decisions is a judgment call.[11] The adequacy of the

[11] Indeed, for many procedures (c.f., W-NOMINATE) there are not even any formal statistical tests (but see Poole, Sowell and Spear 1992). For other methods (c.f., item response models), it may be possible to fit nested models and calculate Bayes factors or likelihood ratio statistics that implicitly or explicitly penalize model complexity. However, it remains true that it is always possible to improve model fit by adding additional dimension, and the tradeoff between complexity and model fit is a judgment of the researcher.

Figure 2: Misclassifications as error or additional dimensions?



The panels shows the location of a hypothetical status-quo position (SQ) and a proposed alternative (P). The left-hand panel shows the one-dimensional preferences of members (estimated from previous roll calls) and their votes on some new roll call. The Y's represent yea votes and N's represent nay votes. The right-hand panel shows these same members plotted on two dimensions. By adding a second dimension, it is possible to draw a line that perfectly divides yea and nay votes while keeping their ordering on the first dimension constant. In general, it is always possible to add dimensions to better predict the data.

model is determined by the number of observations we are comfortable with describing as external to the model or as "random." However, it is *always* possible to account for more of the data within the spatial model by adding additional dimensions. Indeed, we can perfectly account for all of the observed data in the roll-call record for M legislators if we use exactly $M - 1$ dimensions. It is in this sense that the question of how many dimensions best describe the data is a judgment.

This point is illustrated in Figure 2. The left hand panel shows that, for this hypothetical roll call (given previously derived member preferences), if the voting data are placed onto a single dimension, then the model yields a significant number of errors. The right-hand panel, on the other hand, shows that by expanding the dimensionality it is possible draw a single line that yields no errors at all and does not require altering member positions on the first dimension.

A second approach is to use one of several heuristics in the literature to identify the appropriate number of dimensions. The two most widely used are Kaiser's (1960) eigenvalue-greater-than-one rule and the so-called "elbow-test" proposed by Cattell (1966). Each of these heuristics is designed to help scholars make a *judgment* regarding whether “enough” of the structure of the data has been explained by a specific number of dimensions. The remaining errors are again attributed to noise or some other unknown source. But eigenvalue and elbow rules are not tests in any strict sense. They merely provide guidance for researchers regarding when adding additional dimensions will reduce the number of errors “sufficiently.”

3. SIMULATING ROLL-CALL RECORDS

With this discussion behind us, we now turn to the Monte Carlo simulations themselves. In each, we generate ideal points of members from a known distribution with a known number of dimensions. We generate observations from these ideal points by having members vote according to known rules and generate a roll-call record.[12] In Section 4, we then scale this data to answer two questions. First, under what circumstances (i.e., for what parameter settings) do we recover a small number of dimensions from the simulated data? Second, under what circumstances do we recover evidence in favor of low-dimensionality concomitant with that produced in analyses of real-world roll calls?

3.1. *Spatial voting*

We begin with simulations that are perfectly consistent with assumptions behind the W-NOMINATE procedure and in which preferences are distributed unimodally. We then alter the assumed distribution of member ideal points to consider the effect of bimodal distributions on the recovered dimensionality.

[12]It would be possible to include additional hypothesized complications such as majority-party agenda control (Cox and McCubbins 1993, 2005) or bill events (Clinton and Meirowitz 2001; Roberts, Smith and Haptonstahl 2009). However, the simple adjustments to the basic spatial and distributive voting models that we make below are sufficient for making our point, and we attempt to keep our simulations as simple as possible.

Case 1 (Spatial voting with unimodal distributions): Preferences of legislators are strict, linear transformation of Euclidean distance in R^p, where $p \in [0, 1, \ldots, 100]$ denotes the assumed number of dimensions.[13] Member preferences, denoted $\mathbf{x}_i$, are drawn from a multivariate normal distribution $\mathbf{x}_i \sim N(\mathbf{0}_p, \mathbf{I}_{p\times p})$. In each simulation, we generate N observations (i.e., votes) for each of the $M = 100$ member.[14] That is, we ask members to cast a vote comparing a single status quo, $\mathbf{a}_j$, and a single proposal, $\mathbf{b}_j$. Members vote for the alternative that minimizes their squared error loss. Thus, the vote of member i on roll call j is $y_{ij} = I\left[||\mathbf{x}_i - \mathbf{a}_j||^2 > ||\mathbf{x}_i - \mathbf{b}_j||^2\right]$, where $I[.]$ is the usual indicator function.[15]

Case 2 (Spatial voting with bimodal distributions): Preferences of legislators are again linear transformation of Euclidean distance in R^p. However, the ideal points are now distributed bimodally. The two modes are a distance, D, apart in the policy space along each separating dimension. As D increases from zero, the two clusters of legislators become increasingly distinct.[16] We might think of this as a form of polarization. As D increases, the legislature consists of increasingly polarized Democrats and Republicans, or of increasingly polarized liberals and conservatives, or both. Note, however, that nothing here distinguishes Democrats from Republicans or liberals from conservatives other than their policy preferences.

More formally, we assume that $\frac{m+1}{2}$ members are distributed $x_i \sim N(\mu_p, \mathbf{I}_{p\times p})$ while the remainder are distributed $x_i \sim N(-\mu_p, \mathbf{I}_{p\times p})$. One possibility we consider is that the populations may not be polarized on *every* dimension simultaneously. Thus we add an additional parameter, $p_D \in [0, 1, \ldots, p]$, that indicates the number of dimensions on which the distributions are separated

[13]More precisely, the vector of parameters included in these simulations is
$p = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100)$.

[14]We set $M = 100$ to be identical to the U.S. Senate.

[15]To offer a fair test, it is necessary to generate status quo points that result in cut-points (or separating hyperplanes) spread evenly throughout the space occupied by members. We therefore use the following procedure. First, we randomly select (with replacement) one member to be the proposer, whom we assume proposes her own ideal point. Second, we randomly draw a "cut-point", c_j, from the distribution $c_j \sim N(\mathbf{0}_p, \frac{1}{2}\mathbf{I}_{p\times p})$. We then project across this cut-point to specify a status quo. That is, the status quo position on dimension p is chosen as
$a_{jp} = c_{jp} - (|c_{jp} - b_{jp}|)I\left[b_{jp} > c_{jp}\right] + (|c_{jp} - b_{jp}|)I\left[b_{jp} < c_{jp}\right]$. In the bimodal case below, we alter the draw of cut-points to match changes in the distribution of members.

[16]Technically, Case 1 is the special case of Case 2, where $D = 0$.

by the distance D.[17] Thus, if $p_D = 2$ and $p = 4$, then $\mu_p \equiv (\frac{D}{2}, \frac{D}{2}, 0, 0)$.

3.2. *Distributive voting*

So far we have focused on simulations where the assumed behavior of members is closely aligned with what is commonly assumed by PR and other scaling procedures. However, we also considered a very different set of assumptions for member behavior – distributive voting.

In the version the Patient Protection and Affordable Care Act that first passed in the Senate, Ben Nelson (D - NE) was virtually the pivotal voter in its passage. However, he gave his support only after receiving special attention for his state.

> The Senate bill would expand Medicaid to people below 133 percent of the poverty level. And up until 2017, the federal government will pick up the tab for the added cost that will place on state governments. After that, states will have to start sharing the cost. The provision added in the amendment will exempt Nebraska from that sunset, however, meaning that the federal government would forevermore pick up all of the expense of expanded eligibility to Medicaid in that state. (Farley 2009)

Such concentrated benefits are not typically the focus of political scientists, however it is important to realize that this is also a sort of policy option. That is, pork barrel policies are still policies.

Distributive politics imply a different set of assumptions about how members make choices. Instead of choices over continuous policy options, the dimensions of choice are as much about who gets the "policy" as what they get. From a legislator's perspective, what matters is whether he or she brings home the bacon. Thus, each policy exists in a continuous space, but member preferences are binary. Legislators excluded from the distributive coalition have zero utility. If they are included, their utility is unity. Moreover, no legislator gets any value over benefits that goes to another district.[18]

[17] We ran simulations in which the two distributions differ on 0,1,2,3, and p dimensions.

[18] This is obviously a simplified description. There are, of course, costs to each project. These costs are ordinarily assumed to be equally distributed, while the project is targeted to particular districts and more valuable than the increase in their share of costs. The short-hand version in the text is in effect a normalization equivalent to one often used. See Baron and Ferejohn (1989) and Shepsle and Weingast (1994) for details.

How are such coalitions constructed? If Black's (1948) spatial model is the quintessential example of how to analyze a policy space of the kind modeled in the first two cases, Baron and Ferejohn (1989) is the quintessential model of distributive politics. In this model, members form randomly selected minimally winning coalitions to pass bills. We also consider the possibility that members are more likely to share particularistic benefits with co-partisans (see Aldrich 1995; Cox and McCubbins 1993, 2005).

Case 3 (Distributive voting with randomly selected coalitions) :

We suppose, consistent with a Baron-Ferejohn design, one legislator is selected at random to make a proposal by offering a particularistic benefit to form at least a minimal winning majority ($\frac{M+1}{2}$). These chosen members are included in the division of resources and vote 'yea', and the remainder vote 'nay.' More formally, we implement this random coalition formation as follows. For each roll call j, a proposing member k is chosen at random. Each member $i \neq k$ is assigned a random number, r_{ij}, drawn from the standard normal distribution

$$r_{ij} \sim N(0, 1). \tag{1}$$

The proposer then ranks members based on their value of r_{ij} and awards the highest $\alpha \equiv \frac{M}{2}$ members a distributive benefit. Thus, $y_{ij} = I\left[r_{ij} > rank_{\alpha}(r_{ij})\right] \; \forall \; i \neq k$ and $y_{ij} = 1 \; \forall \; i = k$.[19] Thus, although legislators and policies exist in some policy space with a defined number of dimensions, this space does not directly affect member behavior.

Case 4 (Distributive voting with a partisan bias) : In our final scenario, each legislator is either a Democrat ($\tau_i = 1$) or a Republican ($\tau_i = 0$). We suppose that the party label has at least a small value such that the proposer slightly favors his or her co-partisans. For example, the probability of including a co-partisan in the minimal winning coalition may be 0.6 while the probability and of including opposing partisans may be 0.4.

[19] In additional simulations (Aldrich, Montgomery and Sparks 2010), we find that the size of the winning coalition does not affect our results. We therefore assume minimal winning coalitions.

Formally, we assign each member a new trait $\tau_i \in \{0, 1\}$ that indicates party. There are M_{Dem} members in the Democratic caucus and $M_{Rep} = (M - M_{Dem})$ in the Republican.[20] For each roll call j we randomly choose one member k to form a distributive coalition. In this case, we alter (1) such that r_{ij} is assigned to each member where $i \neq k$ from the distribution

$$r_{ij} \sim N\left(\gamma I[\tau_k = \tau_i] - \gamma I[\tau_k \neq \tau_i], 1\right), \tag{2}$$

where γ is the parameter controlling intra-party bias.[21] Higher values of γ mean that members are more likely to form distributive coalitions with co-partisans. Just as in Case 3, the members are ranked based on their value of r_{ij}, and the highest α members are included in the coalition. Members vote based solely on their inclusion or exclusion.

4. SIMULATION RESULTS

For each parameter setting, we generate a hypothetical congress and ask members to vote on N different roll calls according to the procedures described above.[22] We vary only five parameters shown in Table 1. As noted above, several other parameters (e.g., the number of members, denoted M) were found to have no effect and were set to a constant value. In total, we ran 69,000 simulations. Each simulation results in a roll-call matrix. We analyzed these matrices using the W-NOMINATE package in R (Poole et al. 2011). We compare these results with those from an identical analysis of votes in the actual U.S. Senate. Our focus is on the percent of total variance explained by the first dimension alone, which is shown for the U.S. Senate from 1910-2010 in Figure 1.[23] No substantive differences are observed by considering the variance explained by additional dimensions.

[20] The size of the majority makes no difference to our results, nor does the size of the coalition being formed. We set $M_{Dem} = 51$ in all simulations.

[21] Technically, Case 3 is a special case of Case 4 where $\gamma = 0$.

[22] The computer code to generate the datasets is shown in Appendix B.

[23] We have also analyzed these data using the eigenvalue-greater-than-one test and reached identical conclusions (Aldrich, Montgomery and Sparks 2010).

Table 1: Parameters values for Monte Carlo simulations.

Parameter symbol	Interpretation	Simulated values
p	# of dimensions	$1, 2, \ldots, 10, 20, 30, \ldots, 100$
N	# of roll calls	$101, 800, 2000$
D	distance between modes on each separating dimension	$0, 0.25, 0.5, 1, 2, 4, 6$
p_d	# of dimensions on which modes are separated	$0, 1, 2, 3, p$
γ	partisan bias in distributive coalitions	$0, 0.25, 0.5, 1, 2$

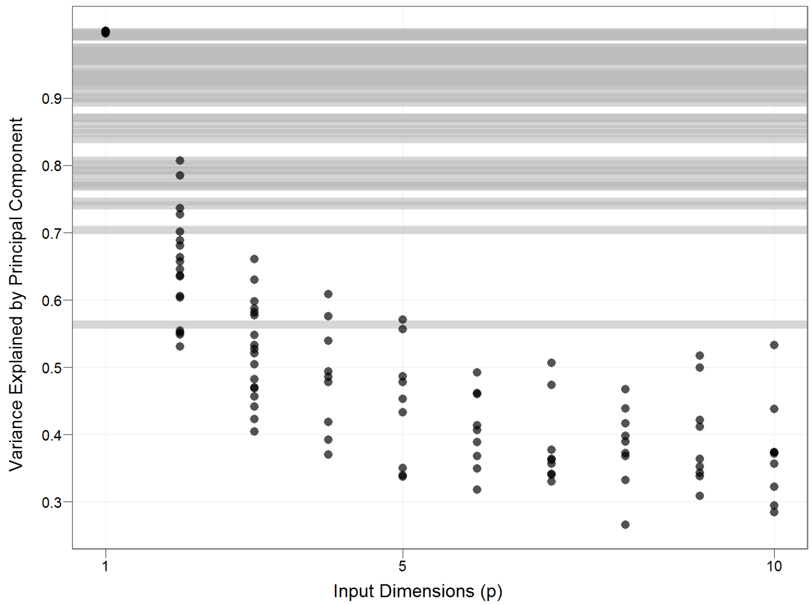
We emphasize that our focus on W-NOMINATE reflects the prominence of this scaling procedure in the literature rather than any flaw or fault inherent to this particular method. All of the findings here have been replicated using simple principal component factor analysis (Aldrich, Montgomery and Sparks 2010). Moreover, using a reduced parameter sweep, the results have been reproduced using both ideal point estimation (Clinton, Jackman and Rivers 2004) and optimal classification (Poole 2000, 2005) (results available upon request).

4.1. *Exemplar simulation results*

In our analyses, we present evidence to answer two questions. If the simulated data is generated from a world where there are p dimensions, would the W-NOMINATE procedure lead us to conclude there are instead one or two dimensions? Second, are the results of these scaling procedures as supportive of the low-dimensionality conjecture as the real-world results shown in Figure 1?

A very simple example of the kind of analysis we will be conducting is displayed in Figure 3. It shows results for a subset of the simulations in Case 1 – those in which the assumed number of dimensions is between one and ten. The horizontal axis shows the true number of dimensions (p) used in our simulations. The vertical axis shows the percent of the overall variance in the roll-call matrix explained by the first dimension. The circles are from our analysis of the simulated roll-call matrices. The gray horizontal lines, show the percent of variance explained by the first dimension for the empirical roll-call matrices for each year in the U.S. Senate from 1910-2010 (see Figure 1).

Figure 3: Exemplar plot of real number of policy dimensions versus percent of variance explained by one dimension.



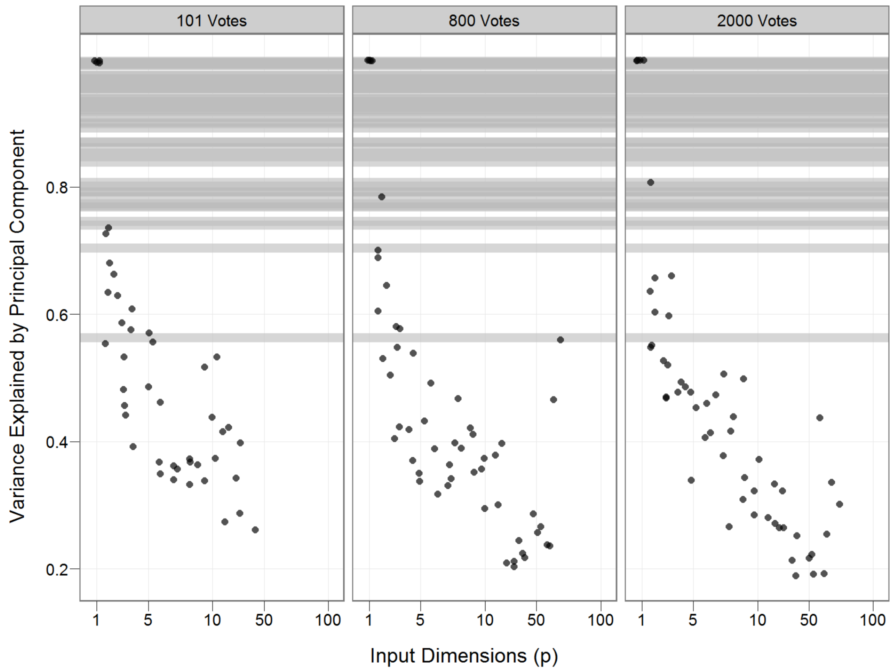
The plots compares W-NOMINATE analyses of simulated and observed roll-call matrices, where the observed data comes for the U.S. Senate. The horizontal axis in each plot shows the "true" number of policy dimensions (p) assumed in the simulations. The vertical axis shows the percent of total variance explained by the first dimension when the-call matrix is analyzed using W-NOMINATE. The circles show the results from our simulations. The horizontal gray bars show the percent of total variance explained by a single dimension for each year of the U.S. Senate from 1910-2010 (see Figure 1). Notice that the dots appear in the region covered by the gray horizontal bars only when the true number of dimensions is small. This shows that – for these parameter settings – W-NOMINATE leads researchers to the correct conclusion regarding dimensionality.

The plot shows quite clearly that, under these parameter settings, W-NOMINATE is very likely to lead researchers to a correct conclusion. Only when the true number of dimensions are *actually* very low (between one and two) does the percent of variance explained match the results we obtain from analyses of the empirical roll-call record. That is, the circles are in the region covered by the gray horizontal bars only when the true number of dimensions is one or two.

4.2. *Results for basic spatial voting (Case 1)*

With this example in mind, we now turn to the broader set of simulations. In Case 1, we vary only two parameters. First, we vary the number of dimensions, p, from 1 to 10 by increments of 1 and

Figure 4: Real number of policy dimensions versus percent of variance explained in the spatial voting model with a unimodal distribution of preferences (Case 1).

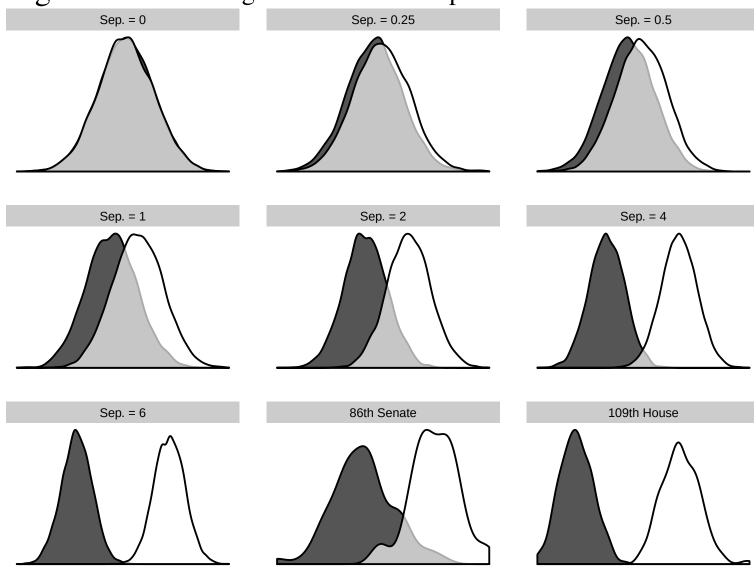


The plots compares W-NOMINATE analyses of simulated and observed roll-call matrices, where the observed data comes for the U.S. Senate. The horizontal axis in each plot shows the "true" number of policy dimensions (p) assumed in the simulations. The vertical axis shows the percent of total variance explained by the first dimension when the matrix is analyzed using W-NOMINATE. The circles show the results from our simulations. The horizontal gray bars show the percent of total variance explained by a single dimension for each year of the U.S. Senate from 1910-2010 (see Figure 1). Under the assumptions of Case 1, W-NOMINATE leads researchers to the correct conclusion regarding the number of underlying dimensions for any plausible number of roll calls. The circles only appear in the areas shaded by the gray bars when the true number of dimensions is small.

then from 10-100 by increments of 10. Second, we vary the number of roll calls. We consider simulations with 101, 800, and 2000 votes. The U.S. Senate averages fewer than 1,000 roll-calls per year, and this last represents far more data than we actually observe for most cases.

The results for the full set of simulations for Case 1 are shown in Figure 4. Each panel in the figure plots the real number of dimensions against the percent of the variance explained for the simulated roll-call matrices. The plot shows that W-NOMINATE does quite well under the assumptions of Case 1. Regardless of the size of the roll-call matrix, the procedure points to a low-dimensional world only when there are actually one or two dimensions.

Figure 5: Visualizing simulated and empirical bimodal distributions.



The first seven panels show a random draw of 20,000 observations from the bimodal distribution used in our simulations. The final two panels show the distribution of 1^{st} dimensional W-NOMINATE scores for the 86^{th} Senate and 109^{th} House, as examples of low and high levels of empirically-observed polarization. The level of polarization in our simulations is no greater than what we observe in the empirical roll-call record.

4.3. *Results for bimodal distribution (Case 2)*

In Case 2, we consider the effect of allowing for bimodal member preferences. We do this to mimic the real-world polarization of Democrats and Republicans (or liberals and conservatives). Figure 5 shows examples (in one dimension) of these bimodal distributions for several parameter settings of D. The final panels of Figure 5 also show the distribution of the members of the 86^{th} Senate and 109^{th} House as estimated by the first dimension of W-NOMINATE. These demonstrates that the level of bimodality we consider in our simulations is no greater than what we observe in the actual roll-call record.

What is the effect of such bimodality on the statistical scaling of roll-calls? Figure 6 plots the percent of variance explained by just one dimension against the true number of simulated dimensions (p) for varying settings of the separation parameter D. In our simulations, we consider values of $D = (0, 0.25, 0.5, 1, 2, 4, 6)$. For the moment, we assume that the modes are separated on every dimension (i.e., that $p = p_D$), although this assumption is relaxed below. The figure shows that larger values for D have a significant impact on the ability of W-NOMINATE to recover the correct dimensionality. When the two modes are even slightly separated, the circles appear in the area shaded by the gray bars for almost any value of p. This is true even when the actual number of dimensions is fairly low (e.g., $p = 5$) and even if the separation of the modes is quite modest (e.g., $D = 1$).

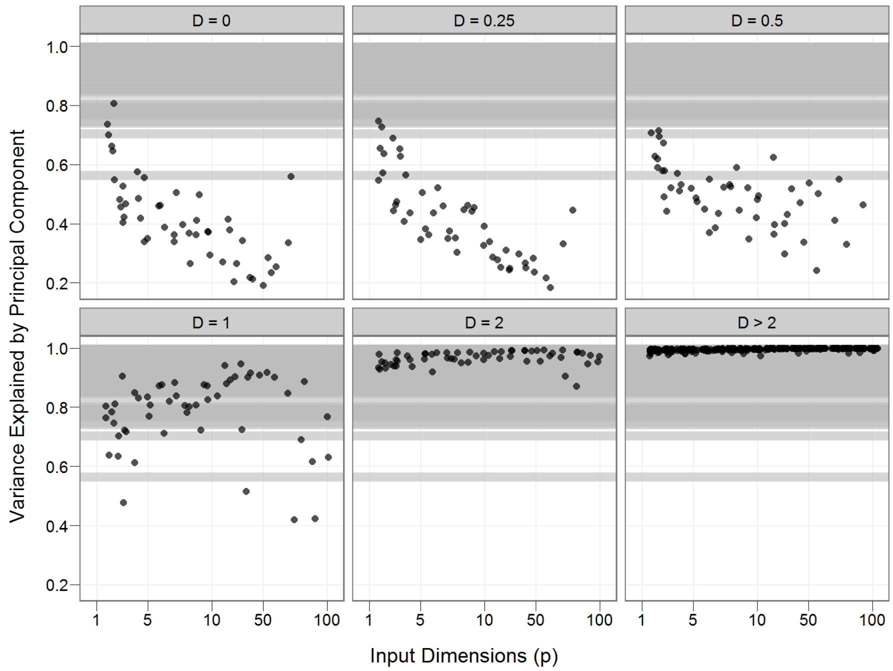
This suggests that if the distribution of ideal points is bimodal, then roll-call scaling will strongly (and potentially incorrectly) suggest that the true space is very small regardless of the actual underlying state. This is most especially true when the degree of polarization is similar to what we observe in the contemporary House and Senate (e.g., $D > 2$). In such an environment (shown in the bottom right panel of Figure 6), the low-dimensionality result is universal and unavoidable.

Thus far, however, we have only considered simulations where the two modes are separated on all dimensions ($p_D = p$). Figure 7 shows the results when $p_d = (0, 1, 2, 3, p)$, where party separation is set to $D = 2$. That is, it shows simulation results when the modes are separated on only a subset of possible dimensions. The figure shows that misleading low-dimensional results are *still* likely even when the modes are only modestly separated and even when this separation takes place on only a fraction of the total number of underlying dimensions. For instance, the circles are in the space shaded by the gray bars even when the parties are separated by only two dimensions, but the true number of dimensions is $p = 20$.

4.4. *Results for distributive cases (Cases 3 and 4)*

In the pure distributive model with random coalitions, members care whether they personally receive a portion of the pie and care nothing about what other members are included in the coalition. From the perspective of a standard scaling technique this is random voting. Our expectation is that W-NOMINATE will find no structure to the data. And, indeed, the leftmost panel of Figure 8 shows exactly that. W-NOMINATE (or any scaling procedure) is unable to account for any of

Figure 6: Real number of dimensions versus variance explained in the basic spatial voting model with bimodal preference distributions, *where the modes are separated on every dimension* (Case 2).

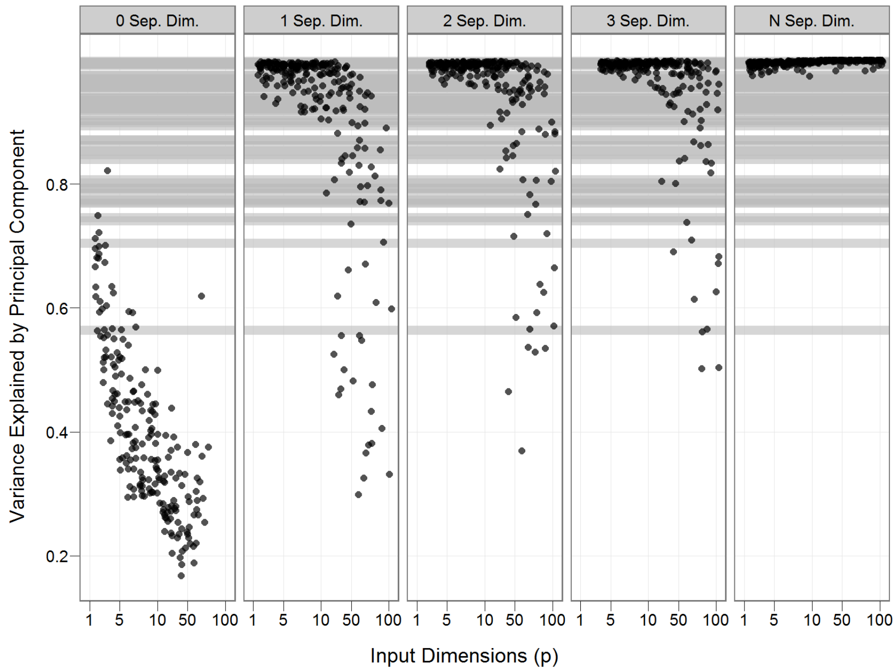


The plots compares W-NOMINATE analyses of simulated and observed roll-call matrices, where the observed data comes for the U.S. Senate. The horizontal axis in each plot shows the "true" number of policy dimensions (p) assumed in the simulations. The vertical axis shows the percent of total variance explained by the first dimension when the matrix is analyzed using W-NOMINATE. The circles show the results from our simulations. The horizontal gray bars show the percent of total variance explained by a single dimension for each year of the U.S. Senate from 1910-2010 (see Figure 1). The presence of the circles in the regions shaded by the gray bars indicates that, under the assumptions of Case 2, W-NOMINATE leads researchers to incorrectly conclude there are a small number of dimensions even when the actual underlying dimensionality is large.

the variance with a single dimension when voting is purely distributive and coalitions are formed randomly.

However, the same is not true if we allow for even modest partisan bias in the formation of distributive coalitions. Figure 8 shows the results when we allow for there to be partisan bias. Recall that the bias parameter (γ) controls the level of correlation between co-partisan affiliation and the probability of being selected to join minimally-winning distributive coalitions. Again, the left panel shows the percent of the total variance in the simulated roll-call record explained by one dimension when there is no partisan bias. The panels to the right show this same information for increasing values of γ. As can be seen, the simulated results swiftly move into the regions shaded

Figure 7: Real number of dimensions versus variance explained in the basic spatial voting model with bimodal preference distributions, *where the modes are separated on a varying number of dimension* (Case 2).

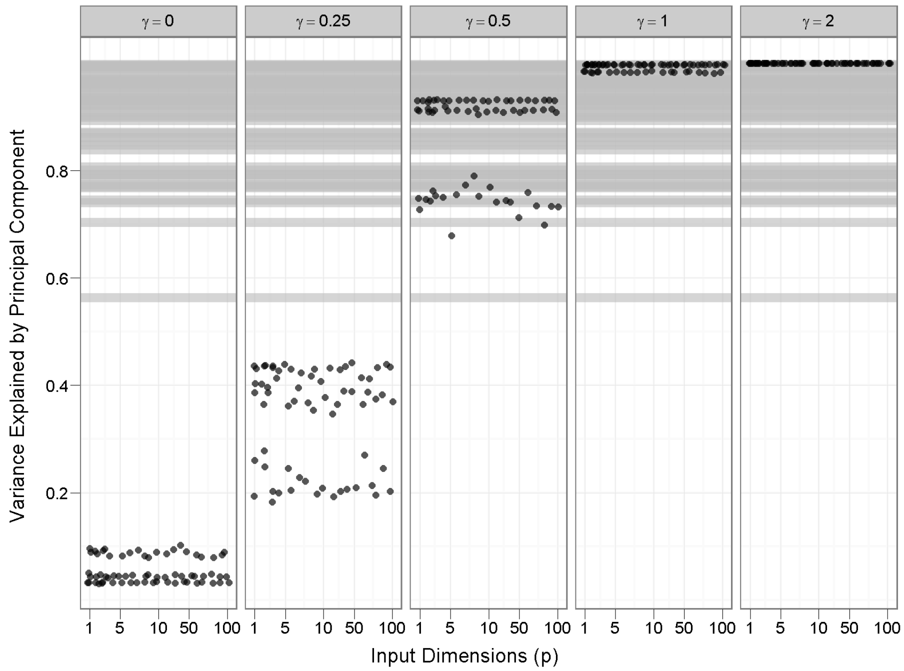


The plots compares W-NOMINATE analyses of simulated and observed roll-call matrices, where the observed data comes for the U.S. Senate. The horizontal axis in each plot shows the "true" number of policy dimensions (p) assumed in the simulations. The vertical axis shows the percent of total variance explained by the first dimension when the matrix is analyzed using W-NOMINATE. The circles show the results from our simulations. The horizontal gray bars show the percent of total variance explained by a single dimension for each year of the U.S. Senate from 1910-2010 (see Figure 1). The presence of the circles in the regions shaded by the gray bars indicates that, under the assumptions of Case 2, W-NOMINATE leads researchers to incorrectly conclude there are a small number of dimensions even when the actual underlying dimensionality is quite large and even when the modes are only separated on a small subset of dimensions.

by the gray bars as the value of partisan bias (γ) increases. For instance, when γ is equal to only 1, W-NOMINATE *always* suggests that there is only one underlying dimension. That is, nearly all of the variance is explained by just one dimension.

Substantively, this means that when minimally winning coalitions consist of only 80% co-partisans on average (compared with the 50% baseline), W-NOMINATE will always suggest a single dimension. Thus, while it is true that distributive votes cannot be "represented on spatial or left-right policy dimensions" (Cox and McCubbins 2005, p.46) in a formal sense, it is nonetheless true that scaling techniques will map strictly distributional voting patterns into an "ideological"

Figure 8: Real number of policy dimensions versus percent of variance explained in the distributive voting model (Cases 3 and 4) for various degrees of own-party bias in the formation of distributive coalitions (γ).



The plots compares W-NOMINATE analyses of simulated and observed roll-call matrices, where the observed data comes for the U.S. Senate. The horizontal axis in each plot shows the "true" number of policy dimensions (p) assumed in the simulations (which are irrelevant in determining votes). The vertical axis shows the percent of total variance explained by the first dimension when the matrix is analyzed using W-NOMINATE. The circles show the results from our simulations. The horizontal gray bars show the percent of total variance explained by a single dimension for each year of the U.S. Senate from 1910-2010 (see Figure 1). When there is no bias ($\gamma = 1$), W-NOMINATE is unable to recover any dimensions. For higher values of γ, W-NOMINATE suggests a low-dimensional solution.

space if there are somewhat stable coalitions, and, if there are two quasi-stable coalitions, this space will be one dimensional.

5. DISCUSSION AND CONCLUSION

We have shown that, under a wide range of plausible circumstances, the finding that a low number of estimated dimensions can explain a large proportion of the variance in roll-call voting does not imply that the process which generate the roll calls is itself low-dimensional. If something like partisanship, or a bimodal distribution of ideology exists, the two modes need not be far removed before we observe low-dimensional estimates across the entire range $(1 - 100)$ of true dimensional states. Further, under a distributive model of politics, moderate increases in the partisan bias with which coalitions are built will produce statistical results that wrongly suggest a low dimensional world.

These findings are not merely methodological. There are at least two broader implications of this work for research on Congress and legislative bodies more generally. To begin with, these results speak to the need for expanded attention to models of politics that are robust to assumptions about the number of dimensions. There is a considerable difference in what spatial models say about politics if the space is or is not *exactly* one-dimensional. In one dimension there is a median voter. If, however, the space is perturbed even infinitesimally away from a pure single dimension, there is no median, and a great many results evaporate (Kramer 1973).[24]

Yet, models that are exceptionally fragile to dimensionality assumptions continue to proliferate in the literature. In many cases, these assumptions are justified either implicitly or explicitly via references to the W-NOMINATE results discussed above. If, as we suggest, the inference of low-dimensionality from these scaling analyses of the roll-call record are unjustified, then there is a great need for expanding the class of dimensionality-robust models.

Thus far, we have focused exclusively on the number of dimensions. However, there are additional features of the basic space of political competition. One question many who use scaled roll-call estimates would like to answer is what the dimensions actually mean? In the case of W-NOMINATE scores, it is quite common to assert that the dimensions are ideological, and that the first dimension is what we ordinarily mean by the liberal-conservative or ideological dimension.

This may be true, but the results above indicate that the evidence we have at hand cannot

[24] While not strictly requiring a single dimension, nearly all applications of Romer-Rosenthal agenda setting are also based on exacting unidimensionality assumption for the simple reason that they nearly always require a median voter to exist (Romer and Rosenthal 1978). Pivot point models are the same category (e.g., Krehbiel 1998). Moreover, many derivations of Duvergerian-style results (Palfrey 1989), prominent models of elections and government formation under proportional representation (Austen-Smith and Banks 1988), informational models of Congress (Gilligan and Krehbiel 1989; Krehbiel 1991), Persson-Tabellini models (Persson and Tabellini 2000), and others (e.g., Iversen and Soskice 2001) require a very exacting form of unidimensionality. Many, if not all, of their derivations simply collapse if the assumption fails to the slightest possible degree (Kramer 1973). It is even the case that many of the results used to study n-dimensional policy spaces are built on repeated application of median voter logic (Shepsle and Weingast 1987; Laver and Shepsle 1990).

justify such a conclusion. It could be that the major dimension estimated is a liberal-conservative dimension, but it could just as easily be something else. Specifically, the dimension may simply be picking up the effect of party pressures and "teamsmanship" (Lee 2009).

This has potentially broad implications. Consider the often made claim that the first dimension is a liberal-conservative ideology, and that it is those preferences that are the most important *causes* of voting choices (Poole and Rosenthal 2007). From this standpoint, the very clear pattern of increased polarization in the Congress is interpreted as ideology leading to greater party polarization and higher levels of party voting. While that set of causal claims is consistent with the observed patterns in the scaled roll-call voting data, so is the opposite claim that parties have become stronger and more unified, and they have led their members to vote more often along party lines. Our results suggest that this would result in scaled member "ideologies" that appear to become increasingly polarized and unidimensional.

Indeed, interpreting increased polarization and unidimensionality in the roll-call record as a *result* of increasingly powerful parties is consistent with other findings that many roll calls with no apparent ideological content (e.g., distributive votes) map neatly onto the single left-right dimension (Lee 2009). It also explains why issues that are historically unrelated to the main left-right dimension come to be mapped so completely onto it once the issue comes to divide Democrats from Republicans (Karol 2009; Lee 2009). The main dimension of conflict is that which political elites have chosen to reveal in their public discussions and on the floor of Congress. It did not used to include Civil Rights, but now it does (Carmines and Stimson 1989; Poole and Rosenthal 2007). It did not used to include abortion, but now it does (Adams 1997; Karol 2009). Our results suggest that these changes may simply be a result of the changing position of the *parties* rather than any more fundamental change in the relationship between these policies in the minds of members, the public, or anyone else. As Lee (2009) notes, "Any issue on which members of the two parties take opposing stands, whether or not is has any ideological content, will map on the first dimension of NOMINATE" (p 52).

Thus, rather than capturing the reality of an increasingly unidimensional world, the low dimensionality of W-NOMINATE (and similar techniques) captures the enhanced coherence of Democrats voting with Democrats, Republicans voting with Republicans, and the estimated first dimension is the line of cleavage along which the parties divide. What political science (and the media) call liberal and conservative may be whatever divides the parties and nothing more.

References

Adams, Greg D. 1997. "Abortion: Evidence of Issue Evolution." *American Journal of Political Science* 41(3):718–737.

Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Political Parties in America*. Chicago: The University of Chicago Press.

Aldrich, John H., Jacob M. Montgomery and David Sparks. 2010. "Drawing (Inferences) Outside the Lines: Dimensionality in Congress." Paper presented at the Annual Meeting of the American Political Science Association.

Ansolabehere, Stephen D., James M. Snyder and Charles Stewart. 2001. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science* 45(1):136–159.

Austen-Smith, David and Joseph Banks. 1988. "Elections, Coalitions, and Legislative Outcomes." *American Political Science Review* 82(2):405–422.

Bafumi, Joseph, Andrew Gelman, David K. Park and Noah Kaplan. 2005. "Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation." *Political Analysis* 13(2):171–187.

Bafumi, Joseph and Michael C. Herron. 2010. "Leapfrog Representation and Extremism: A Study of American Voters and Their Members in Congress." *American Political Science Review* 104(3):519–542.

Baron, David P. and John A. Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83(4):1181–1206.

Black, Duncan. 1948. "On the Rationale of Group Decision-Making." *Journal of Political Economy* 56(1):23–34.

Bonica, Adam. 2011. "A Day by Day Measure of Legislator Ideology: Non-Parametric Smoothing of Legislator Ideal Points with Optimal Classification." Working Paper. `files.nyu.edu/ajb454/public/files/Bonica_essay3_localOC.pdf` (Accessed on July 24, 2011).

Brady, Michael and David W. Rohde. 2009. "Moving Beyond Mavericks: An Exploration of Competing Conceptions of Individual Member Vote Misclassifications in the 80th to 110th Congresses:." Paper Presented at the Annual Meeting of the Southern Political Science Association.

Canes-Wrone, Brandice, David W. Brady and John F. Cogan. 2002. "Out of Step, Out of Office: Electoral Accountability and House Members' Voting." *The American Political Science Review* 96:127–140.

Carmines, Edward G. and James A. Stimson. 1989. *Issue Evolution: Race and the Transformation of American Politics*. Princeton, NJ: Princeton University Press.

Cattell, Raymond B. 1966. "The Scree Test for the Number of Factors." *Multivariate Behavioral Research* 1(2):245–276.

Clinton, Josh D. and Aadam Meirowitz. 2001. "Agenda Constrained Legislator Ideal Points and the Spatial Voting Model." *Political Analysis* 9(3):242–259.

Clinton, Joshua, Simon Jackman and Doug Rivers. 2004. "The Statistical Analysis of Roll Call Voting: A Unified Approach." *American Political Science Review* 98(2):355–370.

Cox, Gary W. and Keith T. Poole. 2002. "On Measuring Partisanship in Roll-Call Voting: The US House of Representatives, 1877-1999." *American Journal of Political Science* 46(3):477–489.

Cox, Gary W. and Matthew D. McCubbins. 1993. *Legislative Leviathan: Party Government in the House*. Berkeley: University of California Press.

Cox, Gary W. and Matthew D. McCubbins. 2005. *Setting the Agenda: Responsible Party Government in the US House of Representatives*. New York: Cambridge University Press.

Crespin, Michael and David W. Rohde. 2010. "Dimensions, Issues, and Bills: Appropriations Voting on the House Floor." *Journal of Politics* 72(4):976–989.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper and Row.

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

Farley, Robert. 2009. "Sen. Ben Nelson Says Medicaid Provision Added to Health Care Reform Is Not a Special Deal for Nebraska." *St. Petersburg Times,* Dec. 23.

Gilligan, Thomas W. and Keith Krehbiel. 1989. "Asymmetric Information and Legislative Rules with a Heterogeneous Committee." *American Journal of Political Science* 33(2):459–490.

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

Ho, Daniel E. and Kevin M. Quinn. 2010. "How Not to Lie with Judicial Votes: Misconceptions, Measurement, and Models." *California Law Review* 98:813–1075.

Hurwitz, Mark S., Roger J. Moiles and David W. Rohde. 2001. "Distributive and Partisan Issues in Agriculture Policy in the 104th House." *American Political Science Review* 95(4):911–922.

Iversen, Torben and David Soskice. 2001. "An Asset Theory of Social Policy Preference." *American Political Science Review* 95(4):875–93.

Jenkins, Jefery A. 1999. "Examining the Bonding Effects of Party: A Comparative Analysis of Roll-Call Voting in the U.S. and Confederate Houses." *American Journal of Political Science* 43(4):1144–1165.

Jenkins, Jeffrey A. 2000. "Examining the Robustness of Ideological Voting: Evidence from the Confederate House of Representatives." *American Journal of Political Science* 44(4):811–822.

Kaiser, Henry F. 1960. "The Application of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement* 20(1):141.

Karol, David. 2009. *Party Position Change in American Politics: Coalition Management*. New York: Cambridge University Press.

Kramer, Gerald H. 1973. "On a Class of Equilibrium Conditions for Majority Rule." *Econometrica* 41(2):285–297.

Krehbiel, Keith. 1991. *Information and Legislative Organization*. Ann Arbor: University of Michigan Press.

Krehbiel, Keith. 1998. *Pivotal Politics: A Theory of U.S. Lawmaking*. Chicago: University of Chicago Press.

Laver, Michael and Kenneth A. Shepsle. 1990. "Coalitions and Cabinet Government." *American Political Science Review* 84(3):873–890.

Lee, Frances E. 2009. *Beyond Ideology: Politics, Principles, and Partisanship in the U.S. Senate*. Chicago: University of Chicago Press.

MacRae, Duncan. 1970. *Issues and Parties in Legislative Voting*. New York: Harper and Row.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10(2):134–153.

Masket, Seth E. 2007. "It Takes an Outsider: Extralegislative Organization and Partisanship in the California Assembly, 1849-2006." *American Journal of Political Science* 51(3):482–497.

Mayhew, David R. 1974. *Congress: The Electoral Connection*. New Haven: Yale University Press.

McCarty, Nolan, Keith T. Poole and Howard Rosenthal. 2005. *Polarized America: The Dance of Political Ideology and Unequal Riches*. Cambridge, Mass.: MIT Press.

McKelvey, Richard. 1976. "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control." *Journal of Economic Theory* 12(3):472–482.

McKelvey, Richard. 1979. "General Conditions for Global Intransitivities in Formal Voting Models." *Econometrica* 47(5):1085–1112.

Palfrey, Thomas R. 1989. A Mathematical Proof of Duverger's Law. In *Models of Strategic Choice in Politics*, ed. Peter C. Ordershook. Ann Arbor: University of Michigan Press pp. 69–91.

Patty, John W. 2008. "Equilibrium Party Government." *American Journal of Political Science* 52(3):636–655.

Peress, Michael. 2009. "Small Chamber Ideal Point Estimation." *Political Analysis* 17(3):276–290.

Persson, Torsten and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Boston: The MIT Press.

Poole, Keith T. 2000. "Nonparametric Unfolding of Binary Choice Data." *Political Analysis* 8(3):211–237.

Poole, Keith T. 2005. *Spatial Models of Parliamentary Voting*. New York: Cambridge University Press.

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

Poole, Keith T. and Howard Rosenthal. 1984. "The Polarization of American Politics." *Journal of Politics* 46:1061–1079.

Poole, Keith T. and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.

Poole, Keith T. and Howard Rosenthal. 2007. *Ideology & Congress*. New Brunswick: Transaction Publishers.

Poole, Keith T., Jeffrey Lewis, James Lo and Royce Carroll. 2011. "Scaling Roll Call Votes with WNOMINATE in R." *Journal of Statistical Software* 42(14):1–21.

Rice, Stuart A. 1925. "The Behavior of Legislative Groups: A Method of Measurement." *Political Science Quarterly* 40:60–72.

Roberts, Jason M. 2007. "The Statistical Analysis of Roll-Call Data: A Cautionary Tale." *Legislative Studies Quarterly* 32(3):341.

Roberts, Jason M., Steven S. Smith and Stephen R. Haptonstahl. 2009. "The Dimensionality of Congressional Voting Reconsidered." Paper presented at the Duke University Conference on Bicameralism.

Romer, Thomas and Howard Rosenthal. 1978. "Political Resource Allocation, Controlled Agendas, and the Status Quo." *Public Choice* 33(4):27–43.

Schofield, Norman. 1978. "Instability in Simple Dynamic Games." *Review of Economic Studies* 45(3):575–594.

Shepsle, Kenneth A. and Barry R. Weingast. 1987. "Institutional Foundations of Committee Power." *American Political Science Review* 81(1):85–104.

Shepsle, Kenneth A. and Barry R. Weingast. 1994. "Positive Theories of Congressional Institutions." *Legislative Studies Quarterly* 19(2):149–179.

Smith, Steven S. 2007. *Party Influence in Congress*. New York: Cambridge University Press.

Snyder, James M. and Tim Groseclose. 2000. "Estimating Party Influence in Congressional Roll-Call Voting." *American Journal of Political Science* 44(2):193–211.

Talbert, Jeffrey C. and Matthew Potoski. 2002. "Setting the Legislative Agenda: The Dimensional Structure of Bill Cosponsoring and Floor Voting." *The Journal of Politics* 64(03):864–891.

Truman, David B. 1951. *The Governmental Process*. New York: Knopf.

VanDoren, Peter M. 1990. "Can We Learn the Causes of Congressional Decisions from Roll-Call Data?" *Legislative Studies Quarterly* 15(3):311–340.

Welch, Susan and Eric H. Carlson. 1973. "The Impact of Party on Voting Behavior in a Nonpartisan Legislature." *The American Political Science Review* 67(3):854–867.

Wilcox, Clyde and Aage Clausen. 1991. "The Dimensionality of Roll-Call Voting Reconsidered." *Legislative Studies Quarterly* 16(3):393–406.

Wright, Gerald C. and Brian F. Schaffner. 2002. "The Influence of Party: Evidence from the State Legislatures." *American Political Science Review* 96(2):367–379.

A. MIXING DISTRIBUTIVE AND SPATIAL MOTIVATIONS

It may be that both the spatial and distributive models of politics are by themselves incomplete. Scholars have long argued that many factors simultaneously affect the voting decisions made by members of Congress (c.f., MacRae 1970; Mayhew 1974; Truman 1951). Senators' votes on the Patient Protection and Affordable Care Act, for example, were certainly informed by ideological preferences. Yet, there was also a distributive component, seen most clearly in the federal reimbursement for expenses received by the state of Nebraska as part of the Democrats' attempts to gain Senator Ben Nelson's vote. In these supplementary simulations, we examine the effect that mixing these factors may have on estimates of dimensionality in roll-call data.

We specify a voting heuristic for members that cleanly combines both spatial and distributive motives. We introduce a new parameter, $\kappa \in [0, 1]$, that controls the degree to which roll-call votes are made based on each concern. Just as in the main text, each roll-call j is associated with a randomly selected status quo, $\mathbf{a}_j$, and proposal point $\mathbf{b}_j$. In addition, for each roll-call there is a distributive element. Members are chosen at random to receive a distributive benefit from the proposed bill, and these coalitions may be biased towards co-partisans.[25]

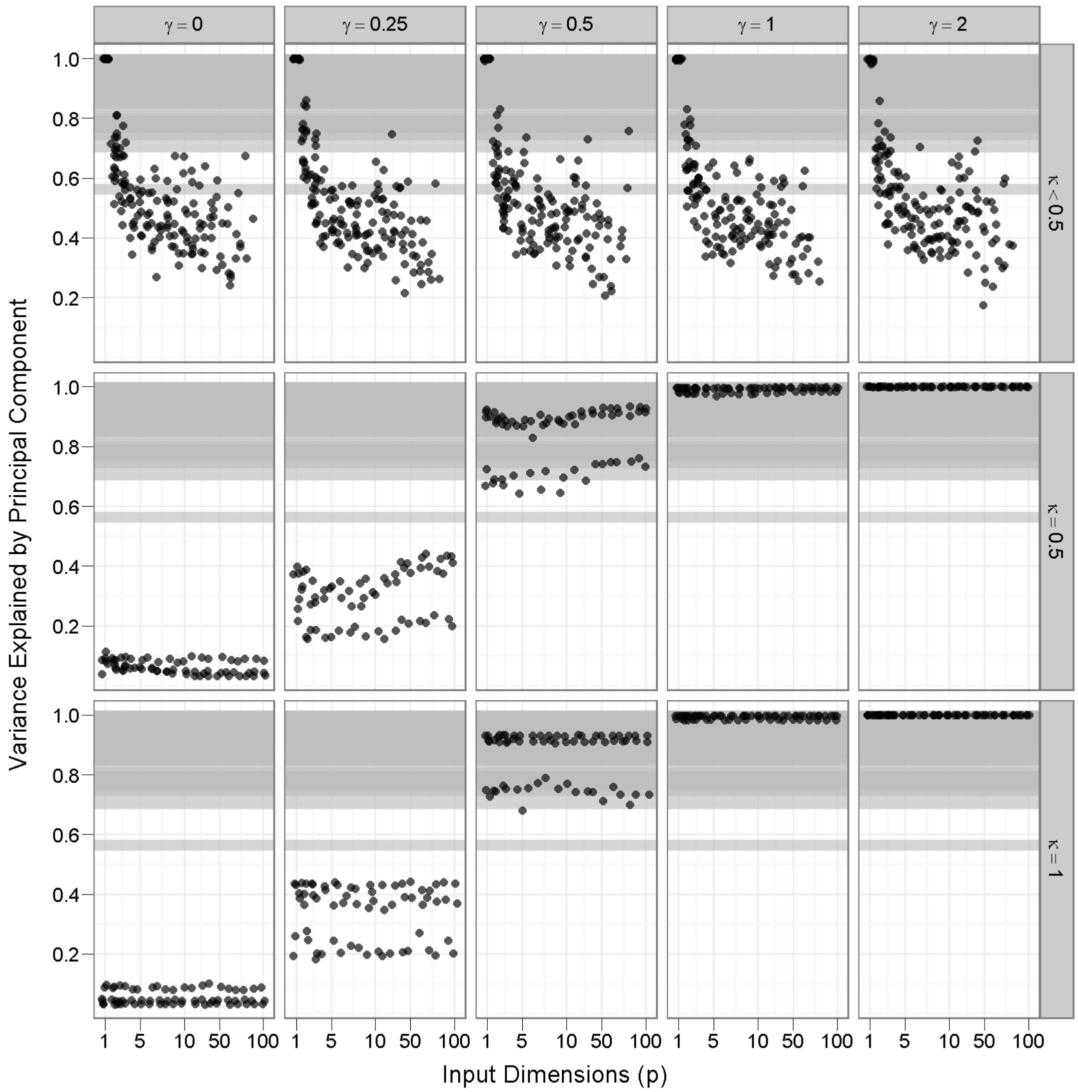
To combine these two competing forces, we re-scale the spatial component using the cumulative distribution function of the standard normal distribution, $\Phi(.)$, similar to a probit link function. We can then use the κ parameter to weight between distributive and spatial concerns. Formally, this is denoted as:

$$y_{ij} = \begin{cases} 0 & \text{if } (1-\kappa)\Phi\left(\|x_i - a_j\|^2 - \|x_i - b_j\|^2\right) + \kappa I[r_{ij} > rank_\alpha(r_j)] < 0.5 \\ 1 & \text{if } (1-\kappa)\Phi\left(\|x_i - a_j\|^2 - \|x_i - b_j\|^2\right) + \kappa I[r_{ij} > rank_\alpha(r_j)] > 0.5 \end{cases} \quad (3)$$

We generate roll-call matrices using this heuristic with five values of $\kappa = (0, 0.125, 0.25, 0.5, 1)$. We also set $D = 2$ and $p_D = p$. Obviously, any values for κ of 0.5 and greater are largely equiva-

[25] In these simulations, we assume that any separation in the bimodal distribution is strictly partisan. That is, the two groupings are assumed to also belong to separate parties for the purposes of forming distributive coalitions.

Figure 9: Real number of dimensions versus percent of variance explained by one estimated dimension for member voting with mixed spatial and distributional motives.



The horizontal axis in each plot shows the "true" number of dimensions (p) assumed in the simulations. The vertical axis shows the percent of total variance explained by the first dimension when the resulting roll-call matrix is analyzed using W-NOMINATE. The circles show the results from our simulations. The horizontal gray bars show the percent of total variance explained by a single dimension for each year of the U.S. Senate from 1910-2010 (see Figure 1). Each plot shows results when simulated distributions are separated by $D = 0.5$.

lent to pure distributive voting. The results of these simulations are shown in Figure 9. The figure shows that for values of κ below the 0.5 threshold, the number of recovered dimensions is not affected by the addition of distributive politics.

B. ABRIDGED SIMULATION CODE

```
####################
# Define Functions #
####################

PreferenceGenerator <- function(Cuts = FALSE){
  Party <- c(rep(1, MajorityPartySize), rep(-1, NObservations - MajorityPartySize))
  Alpha1 <- c(rep((PartySeparation/2) * 1, NSeparateDimensions),
   rep(0, NDimensions - NSeparateDimensions))
  Alpha2 <- c(rep((PartySeparation/2) * -1, NSeparateDimensions),
   rep(0, NDimensions - NSeparateDimensions))
  if (Cuts == FALSE){
    NormalDataPrefs2 <- rmvnorm(NObservations - MajorityPartySize, Alpha2,
     diag(NormalVariance, NDimensions))
    NormalDataPrefs1 <- rmvnorm(MajorityPartySize, Alpha1,
     diag(NormalVariance, NDimensions))
    }
  if (Cuts == TRUE){
    NVotes.Maj = floor(NVotes*MajorityPartySize/NObservations)
    NVotes.Min = NVotes - NVotes.Maj
    NormalDataPrefs1 <- rmvnorm(NVotes.Maj, Alpha1,
     diag(NormalVariance/Cut.Squeeze, NDimensions))
    NormalDataPrefs2 <- rmvnorm(NVotes.Min, Alpha2,
     diag(NormalVariance/Cut.Squeeze, NDimensions))
    }
  NormalData <- data.frame(rbind(NormalDataPrefs1, NormalDataPrefs2))
  if(Cuts == FALSE){ NormalData$party <- Party }
  return(NormalData)
  }

DistanceDifferencer <- function(x){
  Temp1 <- (rowSums((as.matrix(IdealPointArray) - matrix(c(Proposals[x,]),
   nrow=NObservations, ncol=NDimensions, byrow=T)) ^ 2))
  Temp2 <- (rowSums((as.matrix(IdealPointArray) - matrix(c(StatusQuoPoints[x,]),
   nrow=NObservations, ncol=NDimensions, byrow=T)) ^ 2))
  return(Temp1 - Temp2)
  }

Distribute <- function(x){
  Give.Temp <- rnorm(NObservations, IdealPoints$party *
   DistributivePartyBias * ProposerParty[x] ,1)
  return((rank(Give.Temp) <= NDist)*1)
```

```
  }

##############################
# Set of Parameters to Sweep #
##############################

# For example:

Sweep.NDimensions <- c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10,
                       20, 30, 40, 50, 60, 70, 80, 90, 100)
Sweep.NVotes <- c(400, 800, 2000)
Sweep.NObservations <- c(100)
Sweep.NormalMean <- 0
Sweep.NormalVariance <- 1
Sweep.NSimulations <- 1:1
Sweep.NSeparateDimensions <- c(0, 1, 2, 3, 999)
Sweep.PartySeparation <- c(0, 1/4, 1/2, 1, 2, 4)
Sweep.MajorityPartySize <- c(51, 60)
Sweep.Gamma <- c(0, 1/8, 1/4, 1/2, 1)
Sweep.DistributivePartyBias <- c(0, 1/2, 1, 2)
Sweep.NDist <- c(51, 60)
Sweep.Cut.Squeeze <- c(2)  # Aids in narrowing of SQ space

# ...Expand all combinations of parameters to
# a matrix, SweepParameters...

####################################
# Iterate over the parameter space #
####################################

for (i in 1:n.sims){

# ...Assign parameters to local variables...

  # Generate ideal points
  IdealPoints <- PreferenceGenerator()
  IdealPointArray <- as.matrix(IdealPoints[, 1:NDimensions])

  # Select a random proposer, their ideal point, and party
  WhoProposes <- sample(c(1:NObservations), NVotes, replace=T)
  Proposals <- as.matrix(IdealPoints[WhoProposes, 1:NDimensions])
  ProposerParty <- IdealPoints$party[WhoProposes]

  # Draw cut-points from the same distribution as the voters.
  CutPoints <- PreferenceGenerator(Cuts=T)[,1:NDimensions]

  # Given the proposer’s ideal point and the cutpoint,
  # we can infer backwards to the status quo position.

  StatusQuoPoints <- (CutPoints - abs(CutPoints - Proposals)*(Proposals > CutPoints) +
      abs(CutPoints - Proposals)*(Proposals < CutPoints))
  StatusQuoPoints <- as.matrix(StatusQuoPoints)
```

```
  # Ideological component of voting probability function
  IdeologyPart <- (1 - pnorm(DifferenceinDistances))

  # Distributive
  DistributivePart <- sapply(c(1:NVotes), Distribute)

  # Combined Voting Probability Function

  VoteProbability <- (1 - Gamma) * IdeologyPart + Gamma * DistributivePart
  VoteProbability[VoteProbability > 1] <- 1
  VoteProbability[VoteProbability < 0] <- 0
  Votes<-matrix(VoteProbability, nrow=NObservations, ncol=NVotes)


# ...Perform any analyses / save any output...
  }
```