Political Analysis Advance Access published February 5, 2014

Political Analysis (2014) pp. 1–22
doi:10.1093/pan/mpt048

Polarization and Ideology: Partisan Sources of Low Dimensionality in Scaled Roll Call Analyses

John H. Aldrich
Department of Political Science, Duke University
e-mail: aldrich@duke.edu

Jacob M. Montgomery
Department of Political Science, Washington University in St. Louis, Campus Box 1063, One Brookings Drive, St. Louis, MO 63130-4899
e-mail: jacob.montgomery@wustl.edu (corresponding author)

David B. Sparks
Department of Political Science, Duke University
e-mail: dsparks@gmail.com

Edited by Jonathan Katz

In this article, we challenge the conclusion that the preferences of members of Congress are best represented as existing in a low-dimensional space. We conduct Monte Carlo simulations altering assumptions regarding the dimensionality and distribution of member preferences and scale the resulting roll call matrices. Our simulations show that party polarization generates misleading evidence in favor of low dimensionality. This suggests that the increasing levels of party polarization in recent Congresses may have produced false evidence in favor of a low-dimensional policy space. However, we show that focusing more narrowly on each party caucus in isolation can help researchers discern the true dimensionality of the policy space in the context of significant party polarization. We re-examine the historical roll call record and find evidence suggesting that the low dimensionality of the contemporary Congress may reflect party polarization rather than changes in the dimensionality of policy conflict.

1 Introduction

For nearly a generation, congressional research has advanced empirically using estimates of member "ideologies" generated from scaling analyses. Two questions that many who use scaled roll call estimates would like to answer are: "How many dimensions are there?" and "What do the dimensions actually mean?" The standard answer is that there are between one and two dimensions (and today only one), and that the first dimension is a left-right ideological dimension that structures most of congressional politics, or at least roll call voting.

In this article, we challenge this low-dimensionality conjecture and with it the further claim that multidimensional preferences map down to just one liberal-conservative dimension.[1] We show that moderate to high levels of bimodality in the distribution of legislators' preferences necessarily leads scaling procedures to suggest a single dimension—regardless of the true dimensionality of the policy

Authors' note: A previous version of this article 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, DC. We are grateful for comments from Jeff Gill, Frances Lee, Gary Miller, Brendan Nyhan, John Patty, Jon Rogowski, and helpful audiences at Duke University and Washington University in St. Louis. Finally, we thank Keith Poole and Howard Rosenthal for making their roll call data publicly available. Supplementary materials for this article are available on the *Political Analysis* Web site.

[1]Throughout, we use the term "preference" to indicate the point in policy space that each member acts to achieve through roll call voting. We are agnostic as to whether these ideal points derive from personal beliefs, constituency and electoral pressures, or both. This might also be termed each member's "induced ideal point." See Section 3.1 for additional discussion of this issue.

space. That is, party polarization on the order of that found in the contemporary US Congress obscures the true dimensionality of the policy space.

To demonstrate how party polarization downwardly biases estimates of dimensionality, we conduct Monte Carlo experiments varying, first, the true dimensionality of the policy space, and second, the distribution of legislators' preferences. We scale these simulated roll call matrices and show that, if the two parties polarize sufficiently along even a few policy dimensions, scaling procedures will estimate just one or two dimensions, whether there truly are one, fifteen, or anything in between. Thus, finding that a low number of dimensions can explain a large proportion of the variation in the roll call record does not necessarily imply that legislators are making decisions based on preferences arrayed in a low-dimensional space. At the macro level, these results suggest that existing analyses of the roll call record provide ambiguous evidence in favor of low dimensionality. At the micro level, they raise questions as to how scaling estimates should be interpreted and utilized.

We then show through simulation that analyzing the dimensionality of roll calls within each of the two party caucuses provides information about the true dimensionality of legislators' preferences even in the presence of significant party polarization. Indeed, it is precisely in the presence of significant polarization that focusing on each caucus separately is most informative. This suggests a plan of attack in which the scaling of interparty roll call votes can be combined with scaling intraparty roll call votes. The first yields insight into the extent of polarization between the parties on partisan issues, and the second illustrates the true complexity of the policy space.

We therefore re-examine the dimensionality of the US Senate, scaling only the intraparty roll call record, and show that there is less evidence in favor of the low-dimensional conjecture than is commonly supposed. Indeed, examining each caucus separately, the evidence in favor of the low-dimensionality conjecture is relatively weak.

1.1 *Implications for Theories of Political Conflict*

Although these results may seem narrowly methodological, we believe that they are nonetheless of significant theoretical and practical importance for two main reasons. First, at the broadest level, the low-dimensional conjecture is a critical assumption that informs many important theories, formal or otherwise, in the literatures on elections, legislative institutions, and interbranch relations. Our results speak to the need for expanded attention to theoretical models of politics robust to assumptions about the number of dimensions. For instance, there is a dramatic 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).[2] Yet, models that are exceptionally fragile to dimensionality assumptions continue to proliferate in the literatures on elections, Congress, and interbranch relations. In many cases, these assumptions are justified implicitly or explicitly via references to the scaled roll call analyses discussed below. Brady and Volden (2006), for instance, state that:

> In addition to the above reasons to focus on the main policy dimension despite the possibility of multiple dimensions, there is strong empirical support for the existence of a main policy dimension for a number of issues. Poole and Rosenthal (1997) address the history of roll call voting in the Congress and find that preferences along a single dimension can account for about three-fourths of the votes of members of Congress on a wide range of issues (Brady and Volden 2006, 9).

[2]Although not strictly requiring a single dimension, nearly all applications of Romer-Rosenthal agenda setting are also based on exacting unidimensionality assumptions for the simple reason that they nearly always require a median voter to exist (Romer and Rosenthal 1978). Pivot point models are in 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 1992), and others (e.g., Iversen and Soskice 2001; Persson and Tabellini 2000) require a very exacting form of unidimensionality. Many, if not all, of their derivations simply collapse if the unidimensionality 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).

1.2 *Implications for the Interpretation of Scaled Estimates*

Second, our results challenge common substantive interpretations of ideal points estimates produced for each member of Congress. Although combining all roll call behavior into a single unidimensional score is a useful data reduction technique, the resulting estimates may not correspond well with the substantive "ideological" meaning with which they are sometimes ascribed. We show that increased interparty polarization on even a few policy dimensions can lead to a dramatic distortion of the recovered space. Basically, scaled analyses of the entire record emphasize issues that divide the parties, while underemphasizing those issues that divide one or both parties internally. This means that when scaling a legislative body that is deeply divided on just one or two issues, preferential distinctions on less polarized dimensions are swamped by the larger interparty conflict. This leads to a conflation of dimensions such that legislator preferences in multiple distinct issue areas appear to map onto a single dimension—even when preferences on each dimension are distinct and uncorrelated.

Thus, relying on single-dimensional scaled roll call scores may obscure the true nature of member preferences and political conflict itself. The estimated first dimension will represent the issues that divide the parties at any given moment with no necessary "ideological" meaning. What political science (and the media) call liberal and conservative may be whatever divides the parties and nothing more. Single-dimensional scores, therefore, will represent the positions of legislators on those few polarizing issues, while preferences on less partisan dimensions will be obscured.

Substantively, this seems unsatisfactory. It leads to a confusing interpretation of ideology in Congress. No matter whether parties separate on a single policy dimension, a few, many, or all dimensions, it will all be lumped haphazardly in the term "ideology." Ideology therefore means something different if the partisan cleavage happens to involve only economic and welfare policies, or includes civil rights, or includes abortion and family values, or even incidental policies or pork-barrel measures designed by leaders to serve as the basis for running for re-election.

1.3 *Discussion*

Our interpretation of roll call estimates stands in contrast with the usual view that the first dimension is a liberal-conservative ideology, and that it is those preferences that are the most important *causes* of vote choices. From this standpoint, the very clear pattern of decreased dimensionality in recent Congresses is interpreted as ideology becoming more central to all decision-making, leading to high levels of partisanship and polarization. Although that set of causal claims is consistent with the observed patterns in the scaled roll call voting record, in this article we provide an alternative. Party polarization, even on a subset of underlying issue dimensions, results in scaled member preferences that appear increasingly unidimensional regardless of the true number of underlying dimensions.[3] Thus, as parties have become stronger and more unified on a subset of issues, this has led to the illusion of reduced dimensionality.

Indeed, interpreting unidimensionality in the roll call record as a result of increasingly powerful parties is consistent with other findings in the literature.[4] For instance, many roll calls with no apparent ideological content but that divide the parties (e.g., distributive votes) map neatly onto the single left-right dimension (Lee 2009). Moreover, this distortion of the policy space may explain why issues that were historically unrelated to the main left-right dimension map onto it completely, but only once those issues divide Democrats from Republicans in Congress (Karol 2009; Lee 2009). At one time, Civil Rights did not map onto the first dimension of conflict, but now it does (Carmines and Stimson 1989; Poole and Rosenthal 2007). The first dimension previously did not include abortion, but now it does (Adams 1997; Karol 2009). Our analysis suggests that these

[3]Thus, it is not possible to distinguish between (1) a data-generating process where many dimensions are mapped down to just one due to constraint (Enelow and Hinich 1984; Hinich and Munger 1994) and (2) a data-generating process where many dimensions *only appear* to map onto one as a result of partisan teamsmanship on a subset of roll calls (Lee 2009).

[4]By "increasingly powerful" we mean that co-partisans vote together increasingly often, in opposition to an (increasingly unified) opposition party for whatever reasons.

changes may be a result of the changing position of the *parties* rather than any more fundamental alteration 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 it has any ideological content, will map on the first dimension ..." (p. 52). Significant changes in party position from conflict replacement (Schattschneider 1960) or conflict extension (Layman and Carsey 2002) may be masked as distinct policy dimensions, and are subsumed into the broader "liberal-conservative" dimension once the parties divide sufficiently.

Thus, our results speak to the need for empirical scholars to increase attention to more aspects of the roll call record than one-dimensional scaling scores. In the presence of the high levels of polarization that characterize the contemporary Congress, analyses of the first dimension will reveal factors that cause Democrats to differ from Republicans—and little else. A one-dimensional-dominant result may reflect party "teamsmanship," pure left-right ideology, or anything in between. Using standard empirical techniques, we can tell only that parties are divided from one another, but not if they are divided on one issue, many issues, or even, in the sense of Lee (2009), none at all.

Theories that seek to uncover more fine-grained differences in legislative behavior must move beyond party and beyond explaining variation along the main Republican-Democratic axis. Yet many empirical analyses of legislative behavior and interbranch negotiations, our own included (e.g., Aldrich and Battista 2000), rely largely or exclusively on unidimensional scaling estimates (e.g., Krehbiel 1998; Cameron 2000; Cox and McCubbins 2005). This point applies equally to studies seeking to place voters and legislators on a single ideological dimension to explore linkages between voter attitudes and members' roll call votes (e.g., Jessee 2009; Bafumi and Herron 2010; Jessee 2010).

Nonetheless, we wish to emphasize at the outset that our aim is to introduce more caution and skepticism into the theoretical and empirical analysis of legislative bodies, not to attack any method of roll call scaling in particular or roll call scaling generally. Scholars have implemented various multidimensional scaling techniques, with Keith Poole and Howard Rosenthal's work being the most well known (Poole and Rosenthal 1997, 2007). Their procedure represents a significant advance that fully deserves its popularity and widespread adoption. Led by Poole and Rosenthal's own seminal research (e.g., Poole and Rosenthal 1991, 1997, 2007), the analysis of scaled roll call estimates has spawned hundreds of pathbreaking articles and books that have significantly advanced our understanding of Congress, interbranch relations, elections, and democracy itself.[5]

Yet, W-NOMINATE and its numerous cousins[6] are no more appropriate for answering every question in legislative research than are Likert scales appropriate for answering every question in political behavior. Like all measurement techniques, W-NOMINATE and its kin are based on specific assumptions that condition the scope of their applicability (e.g., Herron 2004; Clinton 2012). Our argument here is *not* that scaling procedures are "wrong," only that they are ill equipped to uncover the true dimensionality and nature of legislators' preferences in the context of polarization.[7]

This rest of this article proceeds as follows. We begin by discussing past research on the dimensionality of Congress and approaches to estimating dimensionality in general. In Section 3, we present the details of our Monte Carlo simulations and then analyze the results. In Section 5, we analyze the postwar US Senate, focusing on the intraparty roll call record.

[5]Poole and Rosenthal themselves have been careful to emphasize the limitations of a simple unidimensional model. Although their empirical analyses illustrate the dominance of the first dimension in explaining roll calls in the contemporary era, their broader treatment of the historical record has rested more firmly on the two (or "one-and-a-half") dimensional model (Poole and Rosenthal 2007).

[6]See also Heckman and Snyder (1997); Clinton and Meirowitz (2001); Martin and Quinn (2002); Clinton, Jackman, and Rivers (2004); Bafumi et al. (2005); and Poole (2005).

[7]For ease of exposition, we refer to W-NOMINATE and the analysis of statistics resulting from the application of W-NOMINATE synonymously (Poole and Rosenthal 2007). The W-NOMINATE procedure itself does not predict any specific dimensionality. Rather, the standard approach to analyzing the outputs of this procedure guides researchers to making a judgment about the true dimensionality. We expand upon these points below.

2 Deriving Dimensionality from Roll Call Scaling

Our claim in this article is that the task of identifying the "true" number of dimensions in the Congress lies outside the scope of widely employed scaling procedures—at least as they have been applied to date. Or, to be more precise, the observation of a small number of dimensions resulting from the application of such scaling procedure to the entire roll call record is insufficient to support the inference that the true number of dimensions is actually small. As we show, a surprisingly strong bias toward low dimensionality is present even when the stringent behavioral assumptions of the measurement models are met. When the preferences of two subpopulations are polarized on even a few dimensions, almost any number of true dimensions will appear to map onto just one dimension. Further, our simulations suggest that one approach to checking the true dimensionality of the space in the presence of polarized subpopulations is to examine roll calls *within* the polarized groups (i.e., within party caucuses). Before turning to our simulations and analysis, however, it is worthwhile to step back and examine common approaches to conceptualizing and evaluating dimensionality in Congress.

2.1 *Does the Basic Space of Congress Have 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 Fig. 1, which analyzes the empirical roll call record in the US Senate from 1945 to 2010.[8] The left panel shows the aggregate proportional reduction in error (APRE) associated with the first three dimensions of W-NOMINATE.[9] The right panel shows the difference in APRE for each dimension, reflecting the marginal gains in explanatory power as we add dimensions to the W-NOMINATE model. In essence, these plots show that one or two dimensions alone are sufficient to explain most of the variation in the empirical roll call record, especially for the contemporary Congress. The APRE for a one- or two-dimensional model is very high, approaching the maximum value of unity. Moreover, the marginal improvement in APRE for each additional dimension, shown in the right panel, is fairly modest throughout and shrinks toward zero. Indeed, as shown in the right panel, even adding the second dimension does not improve APRE in the contemporary Senate.

2.1.1 Recent challenges to the low-dimensionality conjecture

Despite this strong evidence, however, recent scholarship has raised serious doubts about the low dimensionality of the roll call record.[10] 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 larger number of dimensions. Norton (1999) shows the same when focusing on roll calls related to gender issues. This work parallels a growing body of research that investigates how violations of W-NOMINATE's stringent assumptions result in systematic patterns of errors and misclassifications of roll call votes.[11] These findings suggest that NOMINATE scores

[8]Replication materials for all of the results in this article are provided in the online dataverse archive associated with this article (Aldrich, Montgomery, and Sparks 2013).
[9]APRE is a common metric for evaluating model fit in roll call analyses. It ranges from zero to one, with larger numbers indicating superior fit. More precisely, the APRE for n roll calls is

$$\text{APRE} = \frac{\sum_{i=1}^{n} (\text{Minority vote size}_i - \text{NOMINATE classification errors}_i)}{\sum_{i=1}^{n} \text{Minority vote size}_i}.$$

See Poole and Rosenthal (2007, 36–37) for additional details.
[10]Roberts, Smith, and Haptonstahl (2009) provide additional discussion of recent scholarship challenging the low-dimensional assumption. See also MacRae (1958), Clausen (1973), Peltzman (1985), Wilcox and Clausen (1991), Heckman and Snyder (1997), Talbert and Potoski (2002), Wright and Schaffner (2002), and Dougherty, Lynch, and Modonna (2012).
[11]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 and Smith (2003); Roberts (2007); Masket (2007); and Patty (2008).

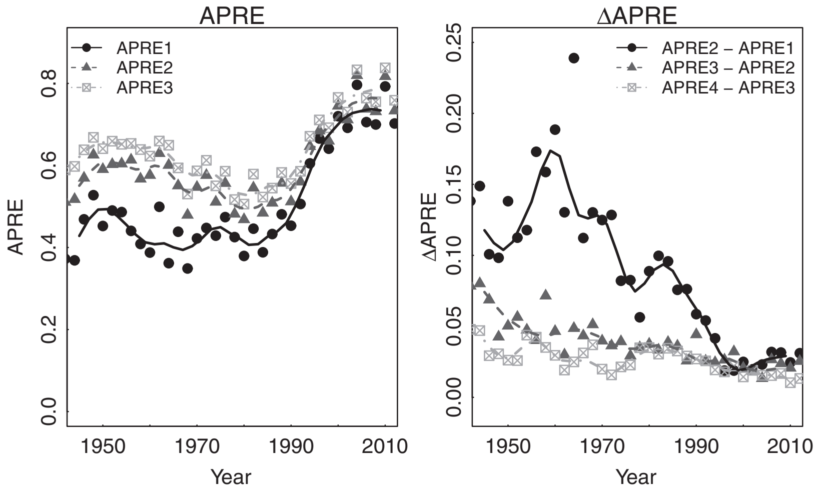

Fig. 1 Empirical APRE results for the US Senate (1945–2010). The left panel shows the APRE statistics for the first three dimensions of a W-NOMINATE analysis of the US Senate from 1945 to 2010. The right panel shows the marginal improvement in APRE as each dimension is added. The points show the estimates from each Senate, and the lines are loess curves. This evidence strongly suggests an increasingly low-dimensional Senate.

and related measures are not detecting member preferences so much as providing a summary of each member's observed behavior, which is "endogenous to the legislative context" (Shepsle and Weingast 1994). The origins of that behavior are manifold, and might include members' personal beliefs, but also voting procedures, pressure from party leaders, and the influence of outside actors. In total, this stream of scholarship does not cast doubt on the value of scaling procedures or their "correctness," but rather suggests that caution is needed in interpretation, since roll call behavior is not an unmediated reflection of a legislator's ideology (Snyder 1992a; Clinton 2012).

A particularly important strain of research has focused on the role that party institutions and party leaders may play in altering the data-generating process to obscure the true dimensionality of the policy space (Dougherty, Lynch, and Modonna 2012).[12] Hurwitz, Moiles, and Rohde (2001) show evidence of multidimensionality in voting on agricultural appropriations, where party pressure is less likely to play a role. 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 evaporate. Similar findings exist for state legislatures without strong two-party systems (Welch and Carlson 1973; Wright and Schaffner 2002). In a comprehensive review of the content of roll calls in the US Senate, Lee (2009) argues that much of the structure of scaling estimates—including the evidence supporting low dimensionality—is a result of partisan "teamsmanship" and may have little to do with members' ideology.

2.1.2 Preference distributions and dimensionality

In this article, we make a different and somewhat broader point. Our aim is *not* to show that party institutions, committee gatekeeping, or agenda control influence the estimated dimensionality of legislators' preferences. Rather, we want to show that scaling procedures will underestimate the true dimensionality of Congress in the presence of moderately strong party polarization *even when all institutional factors have no effect*.

First, we show that the patterns in Fig. 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. The number of dimensions will be severely underestimated in the

[12]For a similar argument focusing on committee gatekeeping, see Snyder (1992a).

presence of the significant party polarization that characterizes the contemporary era, even if the parties polarize on only a small subset of dimensions. More pervasive and significant polarization will almost necessarily suggest a single dimension.

Second, we re-examine the empirical roll call record stretching back nearly 70 years and focus on the record created *within* the two party caucuses. Our analysis of intraparty voting shows that it is possible that it is polarization rather than any fundamental change in the dimensionality of the political conflict that is responsible for the statistical patterns commonly interpreted as supportive of the low-dimensionality conjecture.

2.2 *Establishing Dimensionality*

We begin by briefly considering some of the difficulties inherent in establishing the dimensionality of any data matrix. 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 affect voting. To the best of our knowledge, however, none of the widely applied scaling procedures takes the number of dimensions as a parameter derived from theory. Instead, it is only a maintained assumption that preferences are defined over a space with a finite number of dimensions.

Poole and Rosenthal, 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). However, Poole and Rosenthal assume *nothing* about the dimensionality of the space other than that there are p dimensions, where p is some positive integer. To the best of our knowledge, the same is true for all other scaling procedures prominent in the literature. 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 about dimensionality. The theory is one about choice given preferences, not one about the nature of preferences.

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, both of which suffer from a common problem: the number of dimensions is a subjective judgment by the researcher.

2.2.1 Comparative model fit

One approach, and the method we will rely on below, is to scale the data under a number of different dimensionality assumptions and then compare the results. An inference, to the extent that this can be said to be an inference, is 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, our inferences rely on statistics similar to those reported in Fig. 1. We fit multiple models, examine the results, and determine which model seems "adequate."

However, it is important to realize that this decision is a judgment call. Scholars do not typically fit models of various dimensionalities and then conduct formal likelihood ratio tests or similar statistics. Indeed, for many procedures such formal tests either do not exist (e.g., W-NOMINATE) or are nearly impossible to calculate (but see Poole, Sowell, and Spear 1992). Poole and Rosenthal, for instance, primarily rely on comparative fit indices, such as the APRE statistics shown in Fig. 1. For other methods (e.g., item response models), it may be possible to fit nested models and calculate Bayes factors, although we are aware of no studies that have done so. Instead, researchers rely on comparative fit indices such as the Bayesian information criterion (BIC) or the Akaike information criterion (AIC), which penalize for model complexity (and hence added dimensions). However, although these metrics are in some sense "standard," the degree to which they penalize for model complexity is nonetheless arbitrary. In the end, the adequacy of any model is determined by the number of observations that we are comfortable describing as external to the model or as "random." Yet, it is *always* possible to account for more of the data within the spatial model by adding dimensions.[13] It is in this sense that the question of how many dimensions best describe the data is a judgment.

2.2.2 Additional heuristics

A second approach is to use one of several heuristics in the literature to identify the appropriate number of dimensions. The most widely used are the Kaiser (1960) eigenvalue-greater-than-one rule, the "elbow-test" proposed by Cattell (1966), and the parallel analysis test (Horn 1965).[14] Each of these heuristics is designed to help scholars make a *judgment* regarding whether "enough" of the data's structure is explained by a specific number of dimensions. The remaining errors are again attributed to noise. But the eigenvalue-greater-than-one rule and elbow tests do not allow for statistical inference in a strict sense; they merely provide guidance for researchers as to when adding dimensions will reduce the number of errors "sufficiently."[15] Moreover, these methods extract information from the roll call matrices themselves rather than on any output of the scaling methods. Thus, in our analyses below, we rely on the first approach, with a specific focus on the APRE statistics that have been the dominant focus in the literature (e.g., Jackman 2001; Poole and Rosenthal 2007).

3 Simulating Roll Call Records

With this discussion in mind, we now turn to the Monte Carlo simulations. In each, we generate ideal points (i.e., preferences) of members from a known distribution with a known number of dimensions. We simulate observations by having members vote according to known rules to create a roll call record. In Section 4, we then scale these simulated data.

Although it would be possible to include additional complications, such as majority-party agenda control (Cox and McCubbins 2005) or bill events (Clinton and Meirowitz 2001), the simple simulations below are designed to make our point as cleanly as possible. We are *not* attempting to faithfully replicate the "true" data-generating process of Congress (whatever that may be), and we readily acknowledge that there may be other causes of artificially low-dimensional estimates left unexplored. Rather, we aim to show that interparty polarization in the distribution of legislator preferences by itself can downwardly bias estimates of dimensionality.[16]

[13]Indeed, we can perfectly account for all of the observed data in the roll call record for M legislators if we use q dimensions, where $q \in [1, M-1]$ is the rank of the roll call matrix.

[14]See Brown (2006, Chap. 2) for additional discussion of these tests.

[15]A helpful contrast here is to compare exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) methods. As is well known, in EFA there are no formal tests to compare the dimensionality of the data and researchers must rely on the various heuristics discussed here to choose among possible solutions (Brown 2006).

In contrast, in CFA analyses, one can fit nested models of various dimensions and calculate formal likelihood ratio tests by relying on strong assumptions about the data-generating process that provide marginal likelihoods for different models. Such tests are standard in analyses of factor models with continuous outcomes, and some related techniques are available for latent trait models with binary indicators which would be appropriate for roll calls. Analogous Bayesian methods for comparing models with different dimensionalities using Bayes factors are also available (Ghosh and Dunson 2009). Thus, in principle, the statistics and psychometrics literatures *do* contain several ways that one could fit models of different dimensionalities and test between them based on an assumed data-generating process.

However, there are real and important practical impediments to following such a procedure. First, maximum likelihood routines for estimating latent trait models using roll call data (e.g., IRT models) are poorly behaved and almost never arrive at a proper solution due to the large parameter space. These issues are exacerbated for models with multiple dimensions (e.g., Bartholomew, Knott, and Moustaki 2011). Further, actually constructing Bayes factors in these kinds of high-dimensional models often proves to be computationally burdensome, and we are aware of no published roll call analyses that have done so. We discuss possible paths forward for developing better formal tests of dimensionality in the concluding section.

[16]The code used to generate our main results is provided in our replication materials. In our supplementary materials, we also show results from an alternative set of simulations, where the vote margin of the simulations exactly matches the empirical record (Appendix A). Appendix C briefly discusses how the simulated roll call matrices match the empirical record in terms of party unity votes.

3.1 *Simulation Details*

3.1.1 Distribution of member ideal points

Member ideal points, denoted $\mathbf{x}_i$, are drawn from a multivariate normal distribution $\mathbf{x}_i \sim N(\mathbf{0}_p, \mathbf{I}_{p\times p})$, where p denotes the stipulated number of dimensions. These points represent policy outcomes that members vote to achieve. We are agnostic as to whether these decisions are guided by personal beliefs, constituency pressure, or both. Each of the p issue dimensions represents a distinct area of public policy. Conceptually, we can think of these policy domains as being latent traits (e.g., support for intervention in foreign wars) that guide decisions on multiple specific policies or acts of government (e.g., support for the 1993 US intervention in Somalia). In our simulations, we consider low-dimensional worlds where these policy dimensions are broad (e.g., liberalism-conservatism), high-dimensional worlds where we imagine as many as fifteen more specific latent dimensions (e.g., civil liberties, national security, crime, gay rights, etc.), and multiple states in between. The fewer dimensions, the broader and more encompassing the latent policy traits and policy dimensions are assumed to be.

We hold the number of legislators constant at $M=101$ members.[17] To approximate party polarization, we assume that members come from two subpopulations whose mean locations 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. We might imagine that as D increases, the legislature consists of increasingly polarized Democrats and Republicans. Note, however, that nothing here distinguishes Democrats from Republicans other than their policy preferences.

More formally, we assume that 51 members are distributed $\mathbf{x}_i \sim N(\boldsymbol{\mu}_p, \mathbf{I}_{p\times p})$, whereas the remainder are distributed $\mathbf{x}_i \sim N(-\boldsymbol{\mu}_p, \mathbf{I}_{p\times p})$.[18] We allow that the populations may not be polarized on *every* dimension simultaneously. Thus, we add a parameter, $p_D \in [0,1,\ldots,p]$, that indicates the number of dimensions on which the distributions are separated. Thus, if $p_D=2$ and $p=4$, then $\boldsymbol{\mu}_p \equiv (\frac{D}{2},\frac{D}{2},0,0)$. Note that conceptually we are defining the degree of polarization on a given issue dimension to be equivalent to the distance between party means on that dimension.

3.1.2 Voting behavior

It is here that we mould our analysis around W-NOMINATE in particular, so we can be as faithful as is feasible to the exact assumptions of this procedure. Thus, the assumed behavior of members in our simulations is designed to be consistent with assumptions behind the W-NOMINATE procedure. In each simulation, we generate N observations (i.e., votes) for each of the M members. 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 have Euclidean preferences, and $d^{(a)}_{ijp} \equiv ||x_{ip} - a_{jp}||$. Members vote probabilistically, and we define the probability of member i voting for the proposal on roll call j as[19]

$$P_{ij} = \Phi\left[\beta\left\{\sum_{p=1}^{P} w_p\left(d^{(b)}_{ijp}\right)^2 - \sum_{p=1}^{P} w_p\left(d^{(a)}_{ijp}\right)^2\right\}\right], \quad (1)$$

where w_p indicates the weight of dimension p in member preference and Φ is the cumulative function of the standard normal distribution.[20]

[17]We set $M=101$ to be identical to the US Senate.
[18]In previous work, we found no significant effect for majority size (Aldrich, Montgomery, and Sparks 2010).
[19]This is the only major difference between our simulations and the assumptions of W-NOMINATE, which assumes that

$$P_{ij} = \Phi\left[\beta\left\{\exp\left(-\frac{1}{2}\sum_{p=1}^{P} w_p\left(d^{(b)}_{ijp}\right)^2\right) - \exp\left(-\frac{1}{2}\sum_{p=1}^{P} w_p\left(d^{(a)}_{ijp}\right)^2\right)\right\}\right].$$

In practice, we found simulating from the original formulation to be extremely difficult as the tail behavior from the exponentiation led to a significant amount of almost purely random voting.
[20]For simplicity, we assume in our simulations that $w_p = \frac{1}{p}\forall p$.

3.1.3 Generating random roll calls

To offer a fair test, it is necessary to generate proposals and status quo points spread evenly through the policy space. Therefore, rather than attempting to mimic the distribution of cut points[21] and majority sizes observed in the empirical record (e.g., Poole and Rosenthal 1991), our simulations generate roll calls randomly, an approach that is now common in the literature (e.g., Stiglitz and Weingast 2010; Hirsch 2011). This represents a more agnostic approach, and reduces the likelihood that our results are a product of how "nonrandom selection of roll calls may affect the ability to estimate ideal points that accurately reflect the preferences responsible for generating the observed votes" (Clinton 2012, 79). (We present results from an alternative approach to generating random roll calls in the supplementary materials.)

We use the following procedure to generate randomly distributed separating hyperplanes while still enabling us to easily calculate the voting probabilities shown in equation (1). First, we randomly select a proposal point from a p-dimensional hypersphere with radius r, where all points within the sphere are equally likely to be selected. Second, we randomly select a status quo point using the same approach. In each simulation, we then generate $N = 525$ roll calls such that more than 3% of all members are in the minority (i.e., we exclude all unanimous and near-unanimous votes). This number ($N = 525$) corresponds to the mean number of roll calls in the Senate during the 1945–2010 period.

3.1.4 Overview of simulations

For each parameter setting, we generate a hypothetical Congress and ask all M members to vote on 5000 different roll calls according to the procedures described above. We then randomly select $N = 525$ roll calls that are not unanimous or nearly unanimous. By holding constant several nuisance parameters (M, w_p, N), the latent parameter's space is not overly large. We vary only the five parameters as shown in Table 1. In total, we ran 3600 simulations.[22] Each simulation results in a roll call matrix. We analyzed these matrices using the W-NOMINATE package in R (Poole et al. 2011).

3.1.5 Discussion

Before moving on to our results, it is worth emphasizing 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 method. Several of the findings here have been replicated using simple principal component analysis (Aldrich, Montgomery, and Sparks 2010). Our use of W-NOMINATE should *not* be interpreted as either a critique of the method or an attempt to attribute the widely held belief in the unidimensionality of Congress to these authors.[23]

Finally, we emphasize that our method of generating random roll calls is not intended to reflect the "true" data-generating process in Congress or any other legislative body. We could, for instance, allow only some party "leader" to propose changes to the status quo, which would reflect the gatekeeping power of the majority leadership (Cox and McCubbins 2005). We might also restrict each roll call to a single dimension, reflecting the gatekeeping power of committees. However, relying on such variants to simulate a roll call record introduces selection biases that have been shown to alter estimates of ideal points (e.g., Snyder 1992b; Clinton 2007, 2012) and is likely to alter estimates of dimensionality. The scheme we implement here is a multidimensional equivalent to the random generation of roll call cut points for one-dimensional simulations (Stiglitz and

[21]For the sake of exposition, we use the term "cut point" to refer to the point of intersection between the separating hyperplane and the line segment connecting the proposal and status quo positions.

[22]We did not run simulations where the number of separating dimensions (p_D) would be larger than the number of dimensions (p).

[23]See Footnote 6. In any case, most alternative estimation techniques provide largely identical estimates and model-fit statistics (Clinton and Jackman 2009; Carroll et al. 2009).

Table 1 Parameter values for Monte Carlo simulations

Parameter symbol	*Interpretation*	*Simulated values*
p	No. of dimensions	1, 2, ..., 10, 15, 20
D	Distance between subpopulations on each separating dimension	0, 0.5, 1, ..., 7
p_d	No. of dimensions on which subpopulations are separated	1, 2, 3, 4
β	Scaling parameter for probabilistic voting	0.5, 1, 1.5
r	Radius of the hypersphere for random generation	9, 11

Weingast 2010; Hirsch 2011; Clinton 2012). We believe that our random proposal and status quo approach represents an agnostic assumption that will not artificially impose either a low-dimensional or high-dimensional solution.

4 Simulation Results for a Partisan Legislature

In this section, we present results from the simulations described in Section 3 to support our two major claims: (1) When scaling the entire roll call record, it *is not* possible to determine whether low-dimensional scaling results reflect true low dimensionality in policy preferences or high levels of party polarization; (2) It *is* possible to distinguish between these two sources of low dimensionality by analyzing the within-party record. In Section 5, we turn to testing the empirical implications of these simulations by analyzing the intraparty roll call record of the US Senate.

4.1 *Low Dimensionality in the Context of Party Polarization*

To introduce our method for presenting results, we begin by analyzing a subset of the simulations. We then more systematically support our major theoretical claims. Throughout, we use the empirical analysis of the actual US Senate (shown in Fig. 1) as a benchmark.

A very basic example of the kind of analysis we will be conducting is displayed in Fig. 2. Figure 2 shows results from simulations in which the true dimensionality is low ($1 \leq p \leq 10$) and there is no partisan polarization ($D = 0$). The horizontal axes show the true number of dimensions (p) used in our simulations. The vertical axis on the left panel shows APRE. The vertical axis on the right panel shows the difference in APRE as each new dimension is added.

As a reference, the shaded region shows the *range* of observed values of APRE1 and APRE3-APRE2 in the *empirical* roll call matrices in the US Senate from 1945 to 2010 (Fig. 1). Thus, the shaded rectangles indicate the boundaries of APRE statistics that we have actually observed in the real world, which are commonly interpreted as supporting a one- or two-dimensional policy space. (For the sake of clarity, we do not produce the estimates from every Congress separately.)

In interpreting this figure (and those below), we focus on two specific questions. First, in looking at the left panel, we ask: For what setting of p (the true number of dimensions) are the APRE results as high or higher than the empirical roll call record? Recall that high APRE scores indicate that just a few dimensions are able to explain most of the variation in the roll call record. In the case of Fig. 2, we see that only when the true number of dimensions are *actually* very low (between one and three) do the APRE statistics match the results we obtain from the analysis of the empirical record. That is, APRE1 quickly diminishes as p increases.

In looking at the right panel, we ask: For what settings of p are the differences in APRE scores as low or lower than the empirical record? Recall that small difference scores indicate that adding additional dimensions fails to improve the explanatory power of the model. In this case, APRE2-APRE1, APRE3-APRE2, and APR4-APRE3 begin quite small, but quickly increase to areas well outside the shaded region. Comparing these results to Fig. 1, we can see that only when the true dimensionality is low are the APRE differences comparable to the empirical

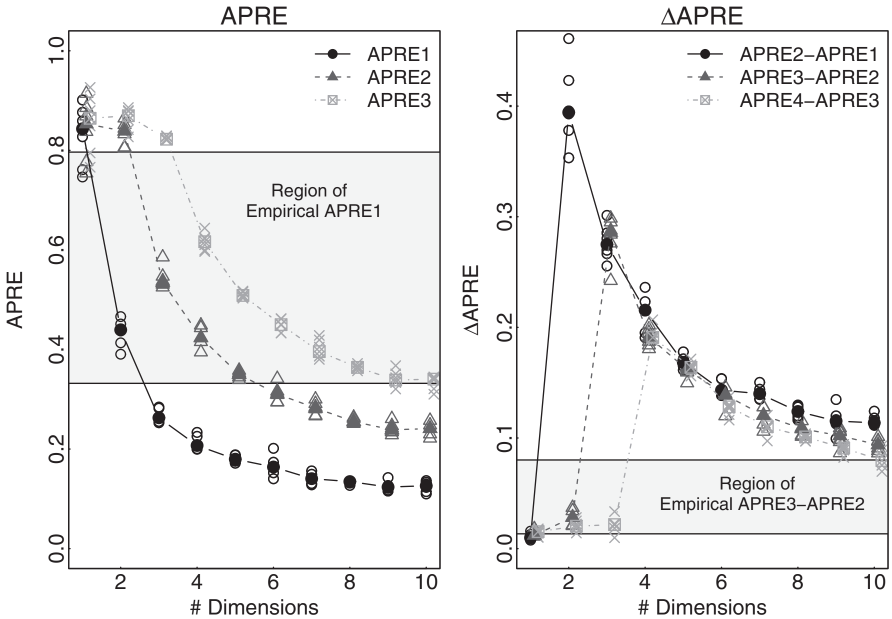

Fig. 2 APRE results with no party polarization. The left panel shows the raw APRE scores from simulated roll calls, whereas the right panel shows the difference in APRE values as each dimension is added. The bold points and lines show the median result at each parameter setting. The lighter points show the results from each simulation. The gray shaded region shows the *range* of observed values of APRE1 and APRE3-APRE2 in the *empirical* roll call matrices for the US Senate from 1945 to 2010 (Fig. 1). When polarization is low, analysts will identify a small number of dimensions only if there are actually few dimensions.

record. This again suggests that, in the absence of party polarization, analyses of APRE statistics resulting from the application of W-NOMINATE will only suggest a low-dimensional space when the underlying space is actually low-dimensional.

Conclusion 1: W-NOMINATE yields a close approximation to the true number of dimensions that actually generated the data when (1) the true number of dimensions is low, and (2) there is no [party] polarization.[24]

However, this positive finding—that scaling procedures only indicate a small number of dimensions in a truly low-dimensional world—does not hold when we allow member preferences to be distributed according to a mixture of normal distributions. We do this to mimic the real-world polarization of Democrats and Republicans. Figure 3 shows examples (in one dimension) of these distributions for several parameter settings of D. The final panels of Fig. 3 also show the distribution of the members of the 86th Senate and 109th House as estimated by the first dimension of W-NOMINATE. These plots demonstrate that the level of polarization we consider in our simulations is no greater (and generally less) than what we observe in the actual roll call record. For example, the 86th Senate (1959–1961) is perhaps the least polarized Senate in the postwar era, and it reflects a partisan separation similar to $D=3$ or $D=4$ in our simulations.

Figure 4 shows how APRE statistics vary as a function of polarization for simulated Congresses of different dimensionalities. Within each panel, the evidence in favor of low dimensionality becomes stronger as party polarization increases. That is, the amount of the roll call record explained by APRE1 goes up. In addition, the relative increase in APRE resulting from adding

[24]To be more precise, W-NOMINATE yields APRE and related statistics that lead researchers to this conclusion. See Footnote 6.

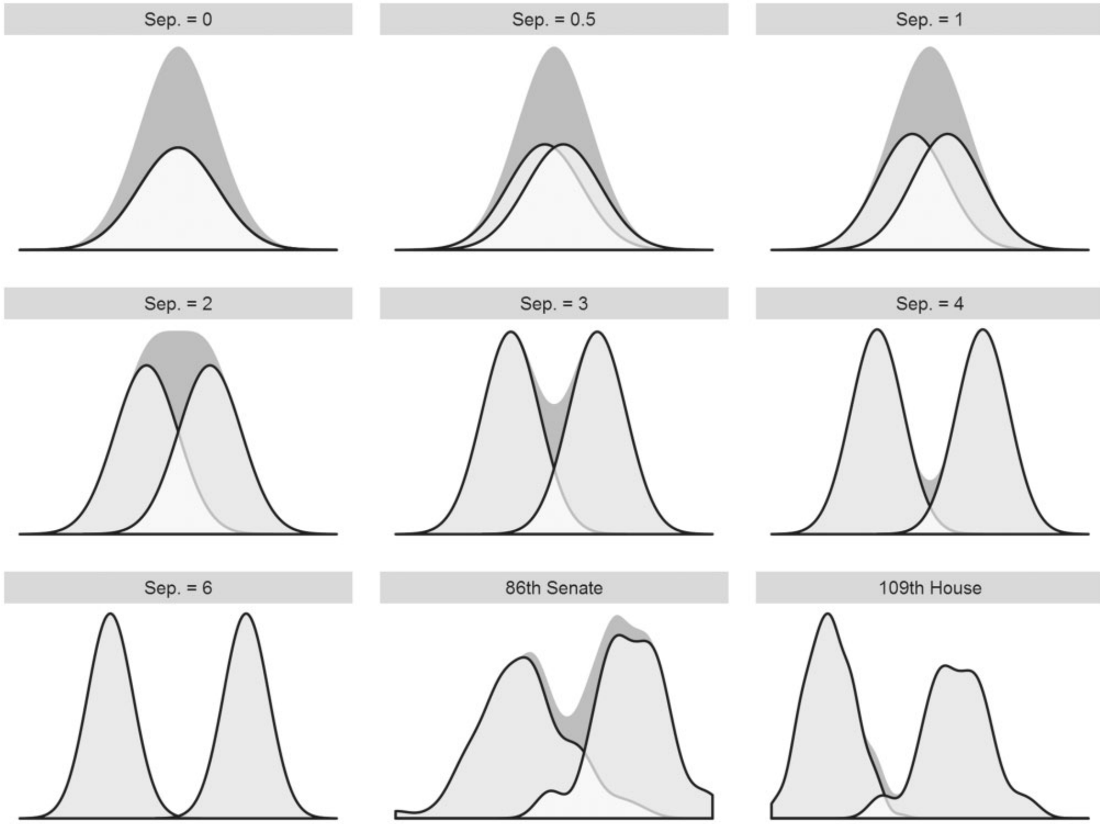

Fig. 3 Visualizing simulated and empirical mixture distributions. The first seven panels show a random draw of 20,000 observations from the mixture distribution used in our simulations. The final two panels show the distribution of first dimensional W-NOMINATE scores for the 86th Senate (1959–1960) and 109th House (2005–2006), 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.

each dimension falls toward zero, indicating that adding additional dimensions does not improve the explanatory power of the model.[25]

This pattern is consistent whether the true dimensionality in the simulations is one (the left panel) or fifteen (the right panel). Notably, for even quite modest levels of polarization (e.g., $D = 3$), these APRE results are in the region (shown in the gray shaded rectangles) that corresponds to the analyses of the actual US Senate and are displayed in Fig. 1. In addition, the results shown in Fig. 4 come with $p_d \in (1,2,3,4)$, meaning that the parties are differentiated along only a few dimensions. These patterns are even more stark when parties separate on more dimensions.

Conclusion 2: As polarization increases, the number of estimated dimensions goes to one, regardless of the true number of dimensions.

Before moving on to the next section, it is worth emphasizing the degree and severity of the underestimation of dimensionality that result from minor changes to the assumed distribution of member preferences. Scaling procedures are, after all, designed to summarize large amounts of data using a reduced number of parameters. A small amount of downward bias in the number of suggested dimensions would not be shocking and indeed is to be expected. However, these

[25]Some readers may notice some seeming outliers, where the APRE1 statistics fall significantly below the median observed values (the solid circles). Upon closer inspection, these points correspond to simulations where polarization occurs on only one dimension ($p_d = 1$).

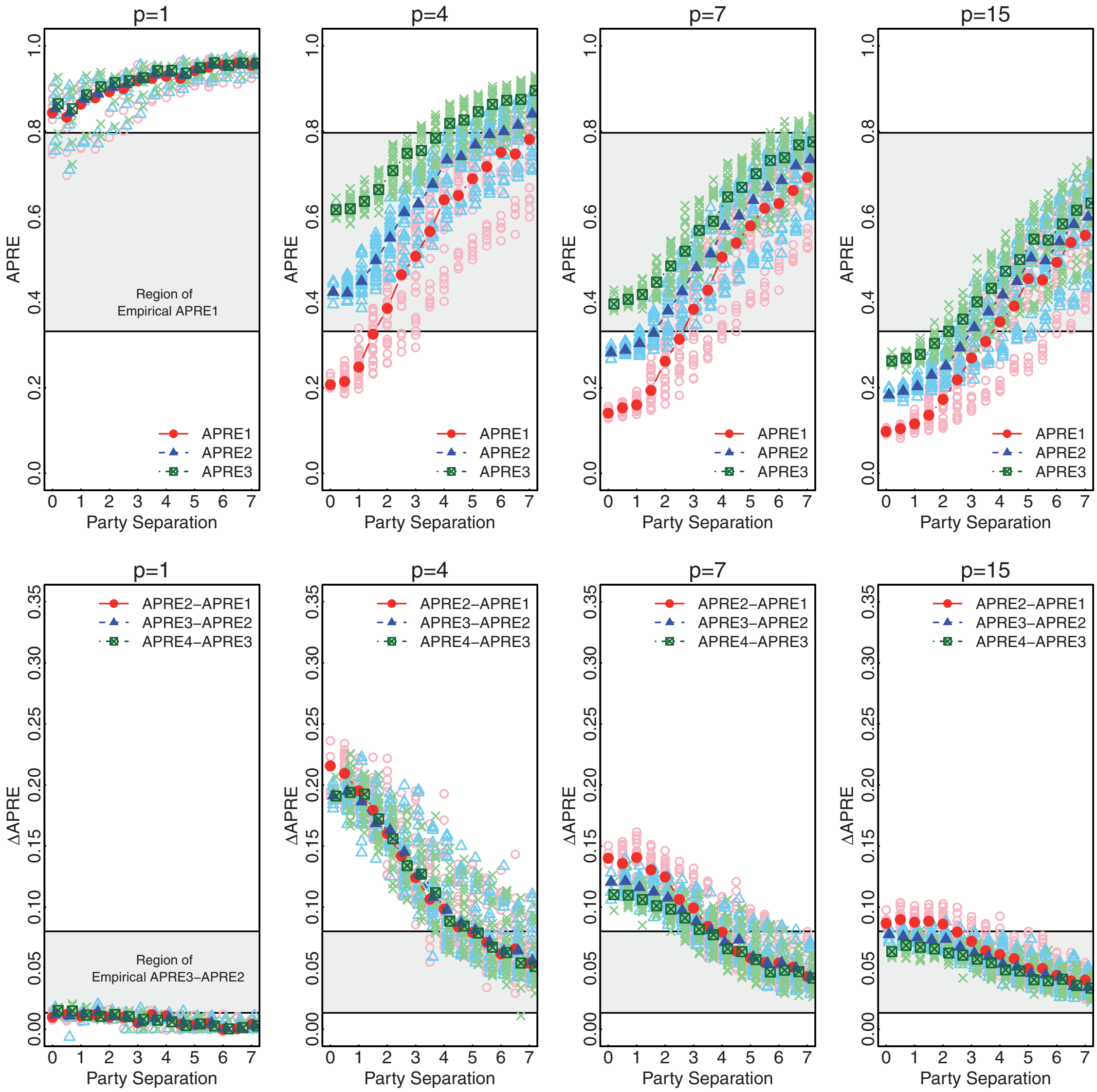

Fig. 4 APRE by values of party separation (D). The left panel shows results when the true dimensionality is one ($p = 1$). The remaining panels show APRE statistics for increasing numbers of dimensions. The bold points and lines show the median result at each parameter setting. The lighter points show the results from each simulation. In the context of significant polarization, W-NOMINATE will yield APRE and related measures that suggest a small number of dimensions regardless of the true value of p.

results indicate that if the parties are polarized on only a few issue areas—and even if this limited polarization is less than what we observe in the contemporary Congress—scaled roll calls will suggest low dimensionality whether there are actually one, seven, or fifteen dimensions.

4.2 *Intraparty Analysis*

The results presented thus far are relatively straightforward. When the preferences of legislators are bimodal—reflecting party polarization on a few issues—standard scaling procedures will suggest a low number of policy dimensions regardless of the "true" complexity of the policy space. Given the widespread acceptance and use of the unidimensionality assumption for analysis of congressional action, this is by itself an important result. Yet, this finding is limited in that it provides no guidance as to how to better estimate the dimensionality of the space in the context of significant polarization. Instead, we have shown that apparent low dimensionality may result *either* from voting in a low-dimensional setting or from bimodality. In this section, we go further by proposing one approach to grappling with dimensionality in the presence of polarization. Intuitively, the results above show that dramatic interparty differences along a subset of issue dimensions obscure distinctions in member preferences along less partisan dimensions. This suggests that we can gain additional leverage by scaling each of the two party caucuses separately.

Figure 5, therefore, shows intraparty APRE statistics for simulated roll call records analyzed above, but now analyzing each party caucus separately. Specifically, we take *all* 525 of the simulated roll calls and scale the responses from members of each party separately. In practice, we must drop all unanimous roll calls, and we also drop all near-unanimous votes in keeping with standard practices.[26] In contrast to Fig. 4, Fig. 5 shows that there is no relationship between the APRE statistics and the degree of party polarization caucus separately. That is, the flatness of each line shows that increased interparty polarization has relatively little influence on APRE statistics as measured within each party.[27] This confirms our notion that analyses of each party in isolation are less biased by party polarization.

Conclusion 3: Scaling intraparty observations approximates the true number of dimensions even in the context of significant (party) polarization.

5 Empirical Intraparty Scaling Analysis

In the previous section, we showed through Monte Carlo simulations that party polarization can interfere with correctly recovering the dimensionality of the "true" world from an observed roll call record. Further, one plausible approach to determining whether low-dimensional statistical estimates reflect true low dimensionality in the context of party polarization is to focus on the intraparty roll call records. With these findings in mind, we now turn to analyzing the claim that preferences in Congress have become more unidimensional over time.

We reanalyze the roll call record in the US Senate from 1945 to 2010, as we did in Fig. 1. However, we now calculate W-NOMINATE scores using only the intraparty roll call record.[28] This means that we generate two APRE estimates for each Congress—one for Democratic senators and one for Republicans.[29] The results are presented in the top panels of Fig. 6, which

[26]In this analysis, we drop all roll calls where fewer than 2.5% of members are in the minority.

[27]The one exception is for the APRE4-APRE3 statistic when $p = 4$, which decreases as a function of D. This illustrates that polarization can lead to a slight downward bias in the estimated dimensionality even using our intraparty approach. Specifically, the simulating suggests that APRE statistics in a highly polarized world will suggest a $p - 1$ dimensional world in the context of significant polarization.

[28]The major consequence of this procedure is the loss of many roll calls that divide Democrats and Republicans on near-party-line votes. This results from the fact that the NOMINATE procedure cannot be utilized for roll calls that are unanimous. Standard practice is also to drop roll calls that are nearly unanimous. In this case, we remove any roll call where fewer than 2.5% of members were in the minority. Since we are analyzing each caucus in isolation, roll calls in which the caucus is united are automatically dropped by the scaling software. However, beyond this, we have done nothing to select specific types of "divisive" roll calls.

One important implication of this is that our estimates are generated with fewer roll calls than in Fig. 1. However, we restrict our analysis to instances where there are more than fifty roll calls with sufficient variation within a given caucus. One possible criticism, however, is that our analysis may now be inaccurate due to the loss of important variance in the record. To make these results more comparable, in the supplementary materials (Appendix B), we scale the entire Congress using only roll calls included in each intraparty analysis. This approach provides nearly identical results.

[29]To be abundantly clear, this analysis differs significantly from past work that focused on distinguishing between estimates recovered when using only party-line and lopsided votes (e.g., Snyder and Groseclose 2000; Cox and Poole 2002). Here, we are *not* subsetting the data set by traits of the roll calls (e.g., whether or not they are lopsided), but by traits of the members. We scale members of each party entirely separately. This analysis makes no direct contribution to the debate on the usefulness of lopsided votes, and the differences between the results in Fig. 6 and those reported in previous studies reflect our entirely different empirical strategy.

The simulations show that it is by examining the party caucuses in isolation that we may better distinguish between low- and high-dimensional spaces in the context of high levels of interparty polarization. Of course, this prevents us from using the ideal points generated from these intraparty analyses in conjunction as they are not on the same scale. However, our focus throughout this article is the dimensionality of the policy space rather than on ideal point estimation per se. A brief discussion of the "accuracy" of the W-NOMINATE estimates when each party is analyzed in isolation is presented in Appendix D of the supplementary materials.

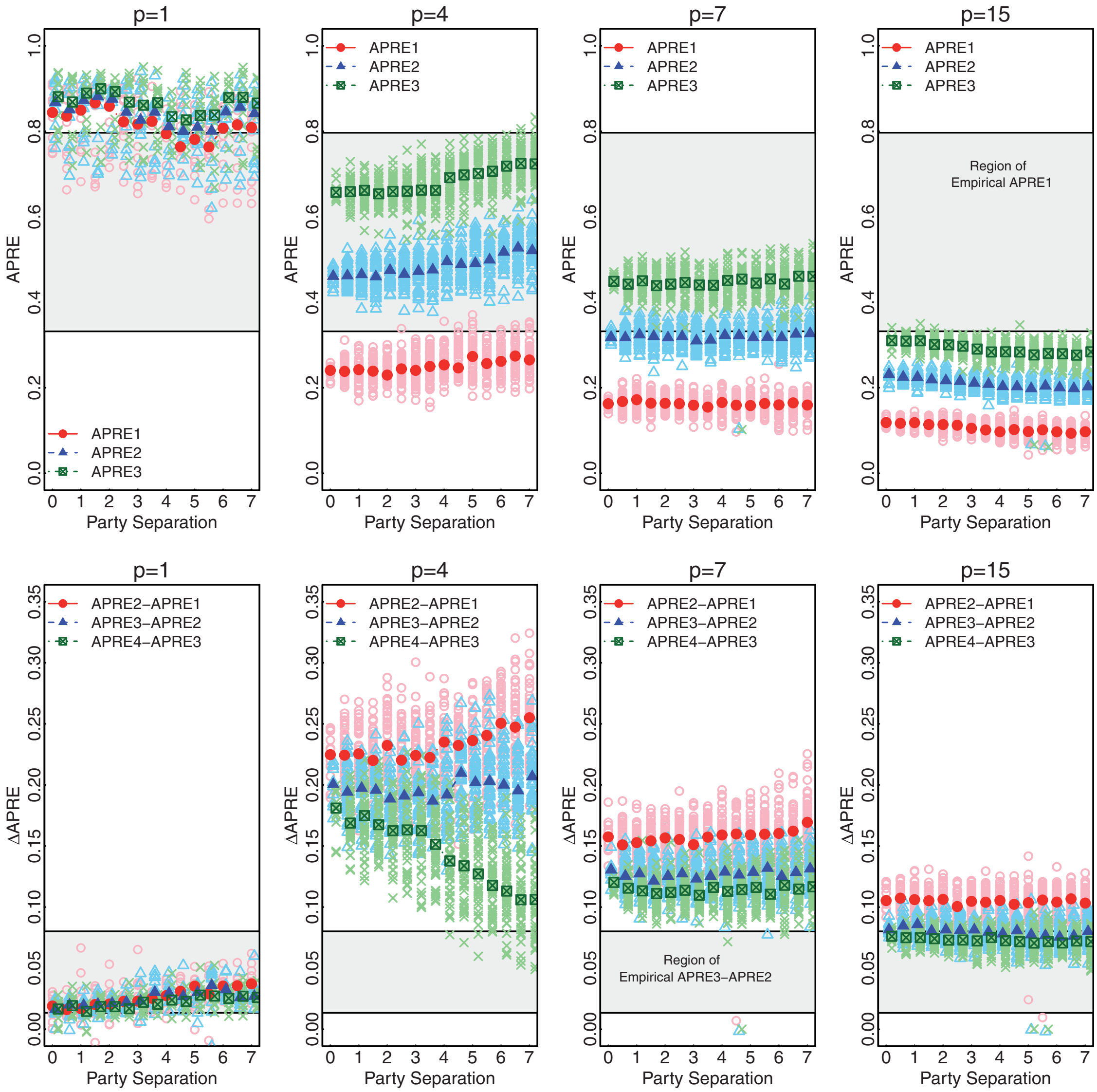

Fig. 5 Intraparty APRE for different values of party separation. The left panel shows results when the true dimensionality is $p = 1$. The remaining panels show APRE statistics for increasing numbers of dimensions. The bold points and lines show the median result at each parameter setting. The lighter points show the results from each simulation. In the context of significant polarization, W-NOMINATE analysis of the intraparty record only suggests a small number of dimensions when p is *actually* small.

show clearly that there is very little evidence of a secular trend toward low dimensionality over the course of the last nearly seventy years when we focus on the within-party roll call record. If anything, the APRE statistics suggest that there might be a slightly *larger* number of dimensions in recent years. Certainly, there is nothing like the dramatic evidence in support of low dimensionality that is revealed by analyzing both parties at once, which are shown in the bottom panels of Fig. 6.

To make our point a bit more explicit, Fig. 7 places the empirical record against simulations with either a low or a high number of dimensions. In each, the solid lines show APRE statistics for the *entire* record (either simulated or empirical), whereas the dashed lines show the APRE statistics as measured only within the two parties. Although this does not constitute a formal statistical test, Fig. 7 suggests that the actual record in the Senate is more in line with a high-dimensional world with increasing levels of polarization than with a mostly, and increasingly, unidimensional world.

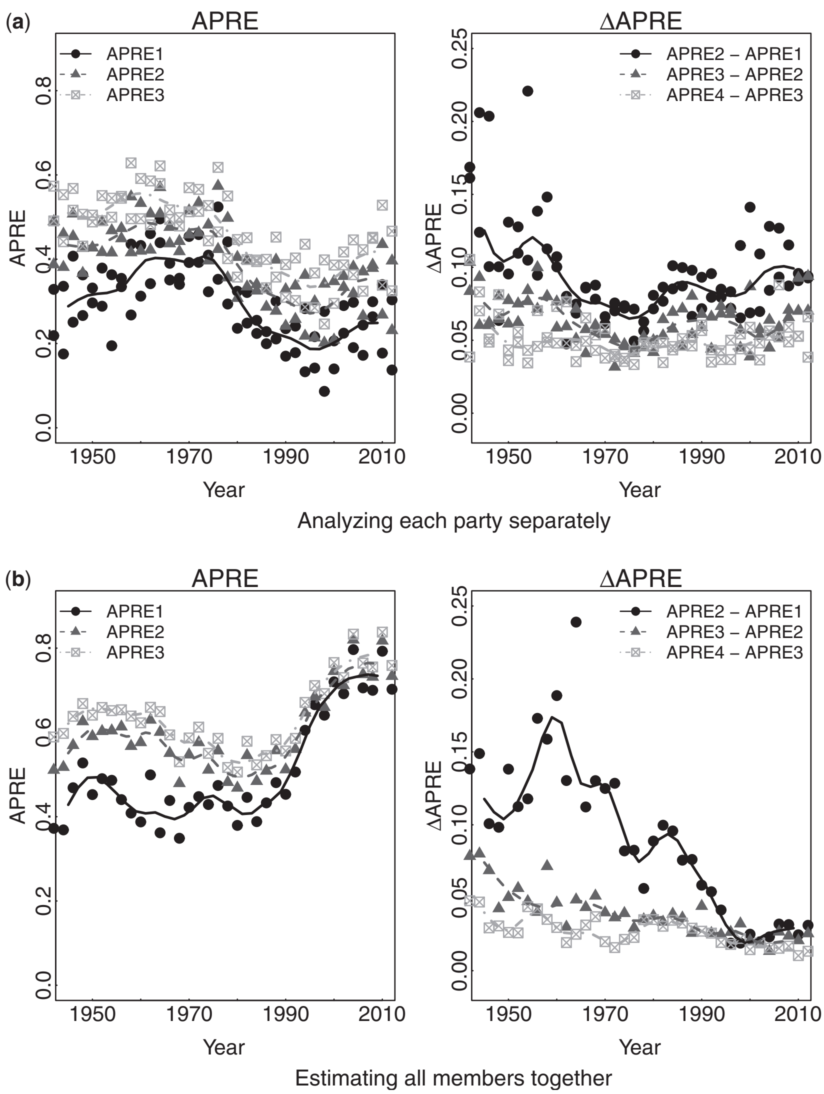

Fig. 6 Empirical APRE results for intraparty and interparty analysis of the US Senate (1945–2010). The left panels show the absolute proportional reduction in error (APRE) for the first three dimensions of a W-NOMINATE analysis of the US Senate from 1945 to 2010. The right panel shows the marginal improvement in APRE as each dimension is added. The results in the top panel are calculated for each party caucus separately. The bottom panels show estimates when all members are included together in the same measurement model. There is little evidence of decreasing dimensionality in the intraparty record.

That is, the APRE statistics for the entire roll call record have changed substantially since the 1960s, but there has been little change in the within-party record. This mirrors our simulation results and suggests that the dimensionality of Congress has not decreased over the last four decades, but rather the parties have polarized.

As party coalitions have drifted apart over time, the variance in roll call voting across parties has increased relative to the variance within parties. Within parties, the source of variation is less consistent and more likely to change over time, requiring a larger number of dimensions to explain the same proportion of variance, relative to an assessment of the entire chamber. Essentially, to the extent that other preference dimensions inform legislator vote choice, those considerations are "washed out" in the roll call record by the influence of partisanship.

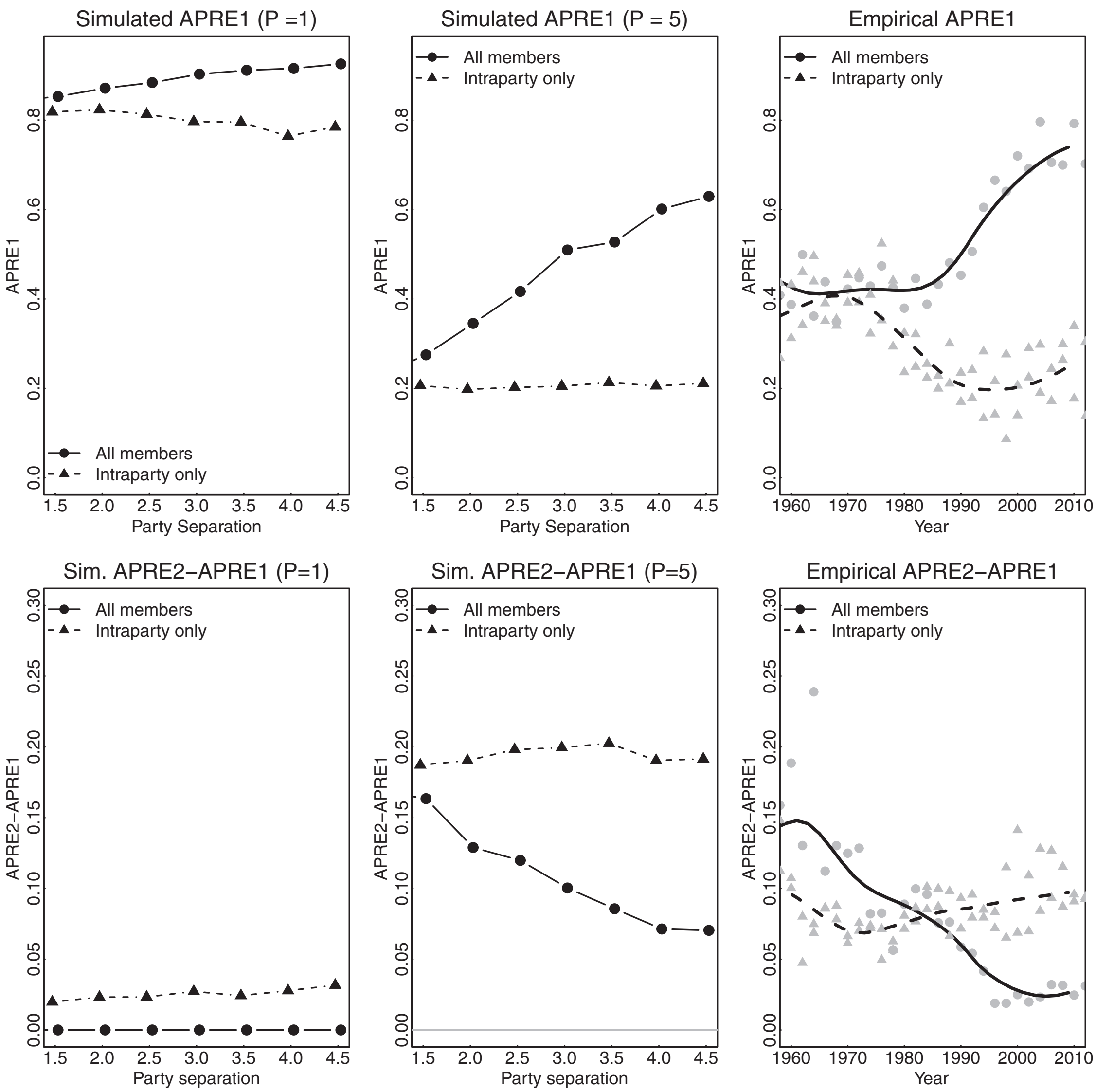

Fig. 7 Comparing inter- and intraparty APRE statistics of simulated roll call matrices with empirical APRE statistics of the US Senate (1960–2010). In the left and center panels, the points show the average APRE scores for analyses of simulated roll call records. The solid lines show results when the entire record is used, whereas the dashed lines show results when examining the intraparty record. The right panel shows the same APRE results for the US Senate from 1960 to 2010, a period of increasing party polarization.

Estimates based on each party separately remove this overwhelming partisan factor, allowing a clearer view of other, less pronounced, dimensions.

6 Conclusion: Taking Multiple Dimensions Seriously

One of the accepted truisms of American politics research is that we can accurately describe preferences over public policy with only one or two dimensions. Scholars and pundits conceptualize policies and political figures as fitting onto a single underlying liberal-conservative continuum. Only on occasion is this supplemented with a second dimension (e.g., "social issues"). Indeed, 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 encapsulate every era of the nation's political history, and that the contemporary Congress is virtually or actually unidimensional (Poole and Rosenthal 1991, 1997, 2007).

In this article, we developed a Monte Carlo procedure to generate roll call records in a multidimensional setting varying the true dimensionality of the policy space and the distribution of member preferences. By analyzing these data, we reached three fundamental conclusions. First, using W-NOMINATE as an exemplar, we showed that when there is no partisan polarization, scaling procedures yield low-dimensional results only if the true space is actually low-dimensional. Second, when the two parties sufficiently polarize on a subset of issue dimension, we found that scaling procedures yielded low-dimensional results whether the true space contained only one dimension, a few dimensions, or as many as fifteen. Thus, polarization on even a small subset of the true policy space can bias judgments about the dimensionality of the policy space. Third, we showed that precisely when interparty polarization obscured our ability to observe the true dimensionality of the policy space, scaling the intraparty roll calls improves our ability to observe the true dimensionality more accurately. Using these three conclusions, we then reexamined the roll call record of the Senate over the last seventy years and found that the evidence in favor of the low-dimensional conjecture is weaker than is commonly appreciated.

Before concluding, it is important to note the limitations of the above findings. To begin with, we emphasize that our goal here is to examine the degree to which it is possible to clearly establish the dimensionality of the roll call record in the context of strong party polarization. We argue above that examining comparative fit indices such as APRE for the entire roll call matrix will obscure the "true" dimensionality of the policy space. Yet, this does not change the fact that, as the empirical APRE results in Fig. 1 show, a great deal of legislative behavior can be predicted using only one or two dimensions. The genius of W-NOMINATE is that it is so well suited to generating "ideal point" estimates that can explain the largest number of roll calls using the fewest possible number of dimensions. However, as our simulations show, effectively summarizing a data matrix and recovering the "true" dimensionality of the data-generating process *can* be distinct tasks calling for different approaches to the data.

Second, the simulations above are far from exhaustive. Different assumed data-generating processes will certainly yield different results. This may be a particular concern for our intraparty analysis. We assume that votes that divide one or both of the parties appear randomly on the agenda. However, many theories of Congress specify particular conditions under which such roll calls may be expected to come to the floor (Cox and McCubbins 2005; Dougherty, Lynch, and Modonna 2012). Another possible set of simulations might assume that that party provides the basic nonstochastic structure for voting patterns, and that the remaining variance in behavior is simply random.[30] Under either of these alternative sets of assumptions, focusing exclusively on the intraparty roll call record could be misleading. Although the results above are sufficient for supporting our main argument—that it is *possible* to recover better dimensionality estimates in the context of polarization by looking within each party—further research might explore the degree to which focusing exclusively on the intraparty roll call record is appropriate under the various alternative assumptions.

Finally, the results above suggest several areas for further research for improving our understanding of the policy space. In particular, these results emphasize the need for exploring ways to directly compare models with different dimensionalities in a statistically informed manner. One plausible approach is suggested by Ghosh and Dunson (2009), who provide methods for computing Bayes factors for models of differing dimensionalities with uninformative priors. Another path forward may be to implement scaling procedures that allow multiple dimensions without assuming they affect all roll calls, similar to the constraints typical in confirmatory factor analysis (Brown 2006; Erosheva and Curtis 2011). These more subtle approaches to summarizing

[30]Such a world is not well supported by the empirical record. Legislators' positions in the ideological space appear to be quite consistent over time, even within parties. This indicates that there is some "signal" underlying the data-generating process besides party.

the roll call record promise to provide estimates that still summarize the data without confining all political disagreement to a single dimension.

References

Adams, Greg D. 1997. Abortion: Evidence of an issue evolution. *American Journal of Political Science* 41(3):718–37.

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.

Aldrich, John H., Jacob M. Montgomery, and David B. Sparks. 2013. Replication data for: Polarization and ideology: partisan sources of low-dimensionality in scaled roll-call analyses. http://dx.doi.org/10.7910/DVN/23248. IQSS Dataverse Network [Distributor] V1 [Version].

Aldrich, John H., and James S. Coleman Battista. 2000. Conditional party government in the States. *American Journal of Political Science* 46(1):164–72.

Ansolabehere, Stephen D., James M. Snyder, and Charles Stewart, III. 2001. Candidate positioning in U.S. House elections. *American Journal of Political Science* 45(1):136–59.

Austen-Smith, David, and Jeffrey Banks. 1988. Elections, coalitions, and legislative outcomes. *American Political Science Review* 82(2):405–22.

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):171–87.

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):519–42.

Bartholomew, David, Martin Knott, and Irini Moustaki. 2011. *Latent variable models and factor analysis: A unified approach*. Hoboken: John Wiley and Sons.

Black, Duncan. 1948. On the rationale of group decision-making. *Journal of Political Economy* 56(1):23–34.

Brady, David W., and Craig Volden. 2006. *Revolving gridlock: Politics and policy from Jimmy Carter to George W. Bush*. 2nd ed. Boulder, CO: Westview Press.

Brown, Timothy A. 2006. *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Cameron, Charles M. 2000. *Veto bargaining: Presidents and the politics of negative power*. Cambridge, UK: Cambridge University Press.

Carmines, Edward G., and James A. Stimson. 1989. *Issue evolution: Race and the transformation of American politics*. Princeton, NJ: Princeton University Press.

Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith Poole, and Howard Rosenthal. 2009. Comparing NOMINANT and IDEAL: Points of difference and Monte Carlo tests. *Legislative Studies Quarterly* 34(4):555–91.

Cattell, Raymond B. 1966. The Scree test for the number of factors. *Multivariate Behavioral Research* 1(2):245–76.

Clausen, Aage. 1973. *How congressmen decide: A policy focus*. New York: St. Martin's Press.

Clinton, Josh D., and Aadam Meirowitz. 2001. Agenda constrained legislator ideal points and the spatial voting model. *Political Analysis* 9(3):242–59.

Clinton, Joshua D. 2007. Lawmaking and roll calls. *Journal of Politics* 69(2):457–69.

———. 2012. Using roll call estimates to test models of politics. *Annual Review of Political Science* 15:79–99.

Clinton, Joshua D., and Simon Jackman. 2009. To simulate or NOMINATE? *Legislative Studies Quarterly* 34:593–621.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. The statistical analysis of roll call data. *American Political Science Review* 98(2):355–70.

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):477–89.

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):976–89.

Dougherty, Keith L., Michael S. Lynch, and Anthony Modonna. 2012. Partisan agenda control and the dimensionality of congress. Unpublished manuscript.

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.

Erosheva, Elena A., and S. McKay Curtis. 2011. Dealing with rotational invariance in Bayesian confirmatory factor analysis. Technical Report 589, University of Washington. http://www.stat.washington.edu/research/reports/2011/tr589.pdf.

Ghosh, Joyee, and David B. Dunson. 2009. Default prior distributions and efficient posterior computation in Bayesian factor analysis. *Journal of Computational and Graphical Statistics* 18(2):306–20.

Gilligan, Thomas W., and Keith Krehbiel. 1989. Asymmetric information and legislative rules with a heterogeneous committee. *American Journal of Political Science* 33(2):459–90.

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* 28:142–89.

Herron, Michael C. 2004. Studying dynamics in legislator ideal points: Scale matters. *Political Analysis* 12(2):182–90.

Hinich, Melvin J., and Michael C. Munger. 1994. *Ideology and the theory of political choice*. Ann Arbor: University of Michigan Press.

Hirsch, Alexander V. 2011. Theory driven bias in ideal point estimates—A Monte Carlo study. *Political Analysis* 19(1):87–102.

Horn, John L. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika* 30(2):179–85.

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):911–22.

Iversen, Torben, and David Soskice. 2001. An asset theory of social policy preferences. *American Political Science Review* 95(4):875–93.

Jackman, Simon. 2001. Multidimensional analysis of roll call data via Bayesian simulation: Identification, estimation, inference, and model checking. *Political Analysis* 9(3):227–41.

Jenkins, Jeffery 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):1144–65.

———. 2000. Examining the robustness of ideological voting: Evidence from the Confederate House of Representatives. *American Journal of Political Science* 44(4):811–22.

Jessee, Stephen A. 2009. Spatial voting in the 2004 presidential election. *American Political Science Review* 103(1):59–81.

———. 2010. Partisan bias, political information and spatial voting in the 2008 presidential election. *Journal of Politics* 72(2):327–40.

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):285–97.

Krehbiel, Keith. 1992. *Information and legislative organization*. Ann Arbor: University of Michigan Press.

———. 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):873–90.

Layman, Geoffrey C., and Thomas M. Carsey. 2002. Party polarization and "conflict extension" in the American electorate. *American Journal of Political Science* 46(4):786–802.

Lee, Frances E. 2009. *Beyond ideology: Politics, principles, and partisanship in the U.S. Senate*. Chicago: University of Chicago Press.

MacRae, Duncan. 1958. *Dimensions of congressional voting: A statistical study of the House of Representatives in the Eighty-First Congress*. Berkeley: University of California Press.

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):134–53.

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):482–97.

Norton, Noelle H. 1999. Uncovering the dimensionality of gender voting in Congress. *Legislative Studies Quarterly* 24(1):65–86.

Palfrey, Thomas R. 1989. A mathematical proof of Duverger's law. In *Models of strategic choice in politics*, ed. Peter C. Ordershook, 69–91. Ann Arbor: University of Michigan Press.

Patty, John W. 2008. Equilibrium party government. *American Journal of Political Science* 52(3):636–55.

Peltzman, Sam. 1985. An economic interpretation of the history of congressional voting in the twentieth century. *American Economic Review* 75(4):656–57.

Persson, Torsten, and Guido Tabellini. 2000. *Political economics: Explaining economic policy*. Boston, MA: MIT Press.

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(8–9):85–101.

Poole, Keith T., and Howard Rosenthal. 1991. Patterns of congressional voting. *American Journal of Political Science* 35:228–77.

———. 1997. *Congress: A political-economic history of roll call voting*. New York: Oxford University Press.

———. 2007. *Ideology & Congress*. New Brunswick, NJ: Transaction Publishers.

Poole, Keith T., Jeffrey B. Lewis, James Lo, and Royce Carroll. 2011. Scaling roll call votes with WNOMINATE in R. *Journal of Statistical Software* 42(14):1–21.

Roberts, Jason M. 2007. The statistical analysis of roll-call data: A cautionary tale. *Legislative Studies Quarterly* 32(3):341.

Roberts, Jason M., and Steven S. Smith. 2003. Procedural contexts, party strategy, and conditional party voting in the U.S. House of Representatives, 1971–2000. *American Journal of Political Science* 47(2):305–17.

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.

Schattschneider, E. E. 1960. *The semi-sovereign people: A realist's view of democracy in America*. New York: Holt, Rinehart, and Winston.

Shepsle, Kenneth A., and Barry R. Weingast. 1987. Institutional foundations of committee power. *American Political Science Review* 81(1):85–104.

———. 1994. Positive theories of congressional institutions. *Legislative Studies Quarterly* 19(2):149–79.

Snyder, James M. 1992a. Committee power, structure-induced equilibria, and roll call votes. *American Journal of Political Science* 36(1):1–30.

———. 1992b. Gatekeeping or not, sample selection in the roll call agenda matters. *American Journal of Political Science* 36(1):36–39.

Snyder, James M., and Tim Groseclose. 2000. Estimating party influence in congressional roll-call voting. *American Journal of Political Science* 44(2):193–211.

Stiglitz, Edward H., and Barry R. Weingast. 2010. Agenda control in Congress: Evidence from cutpoint estimates and ideal point uncertainty. *Legislative Studies Quarterly* 35(2):157–85.

Talbert, Jeffrey C., and Matthew Potoski. 2002. Setting the legislative agenda: The dimensional structure of bill cosponsoring and floor voting. *Journal of Politics* 64(03):864–91.

Welch, Susan, and Eric H. Carlson. 1973. The impact of party on voting behavior in a nonpartisan legislature. *American Political Science Review* 67(3):854–67.

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):367–79.