Studies in American Political Development, 30 (October 2016), 172–184.
doi:10.1017/S0898588X16000110
ISSN 0898-588X/16
© Cambridge University Press 2016

In Defense of DW-NOMINATE

Nolan McCarty, *Princeton University*

Several of the articles in this volume criticize the use of DW-NOMINATE in historical work in American politics and suggest alternative approaches to the use of roll call voting data. While many of criticisms are certainly valid, their practical implications are often overstated. Moreover, the suggested alternatives are either impractical for most historically oriented scholars and or do not adequately address the underlying problems. Almost all of the criticisms can be addressed by correct application of DW-NOMINATE results or those of other closely related measures.

1. INTRODUCTION

The contributors to this special issue all note that ideal point estimation and DW-NOMINATE have had an extraordinary impact on mainstream political science but that these techniques have failed to gain much traction in American political development (APD). Each article offers an explanation for this disconnect. The premise of Everson et al. is that because NOMINATE and other scaling techniques can be technically demanding, few historically oriented scholars have been willing to incur the start-up costs associated with using ideal point estimates and techniques in their research. Thus, the answer is to demystify ideal point estimation and make it more accessible to such scholars. Everson et al.'s excellent contribution makes a giant step forward toward these goals.

The other three articles provide quite a different answer as to why DW-NOMINATE is not used in historical work. Their argument is that DW-NOMINATE generally should *not* be used in historical research. Many reasons are offered. DW-NOMINATE's estimates of the ideal points and the underlying dimensions are hard to interpret. It is not clear whether DW-NOMINATE scores represent ideology, partisanship, or something else. Characterizations of DW-NOMINATE scores in one historical era may not be valid in another. DW-NOMINATE's assumptions about dynamics are too restrictive. Generalized measures of legislator preferences blur the differences across distinct policy areas. None of these claims are new, and very few are specific to historical research. Dealing with the concerns has been at the forefront of the methodological agenda for ideal point estimates for almost two decades. But the contributors to this special issue should be applauded for bringing them to the attention of the APD subfield.

My concerns, therefore, center primarily on the advice proffered as to what to do about these deficiencies. For Lee, the answer seems to be not to use ideal point estimation at all, or if one does, then the outcomes should be interpreted as a summary measure of partisanship. As I explain below, I do not agree with this advice. The articles by Bateman and Lapinski and by Caughey and Schickler also raise serious concerns about DW-NOMINATE but do not reject all ideal point estimation. Indeed, they generally advise using newer, more customized techniques to solve specific problems. On the surface, this seems like inherently sensible advice. Yet, I mostly disagree with it.

My counter-argument can be best expressed in terms of an analogy to men's fashion. Think of DW-NOMINATE as a mass-produced line of men's suits. The unmistakable message of Bateman-Lapinski and Caughey-Schickler is "never buy off the rack." If you do, the DW-NOMINATE suit just won't fit right. The crotch will be too long, the sleeves will touch the knuckles, and the shoulders will have divots. Moreover, wearing such an ill-fitting suit will show that you put very little thought into your purchase and maybe that you just really don't care how you look. Instead, one needs a bespoke outfit from Savile Row. With expert tailoring, the fit will be perfect. The signal of elegance and sophistication will be unmistakable. The problem with custom-made suits, however, is that they are beyond the means of many, and the insistence that one is a must will surely increase the popularity of business casual.

But, fortunately, there is a middle ground. The DW-NOMINATE line allows for many straightforward alterations that can improve the fit dramatically at a fraction of the cost. Although one may not be able to get real buttonholes on their jacket sleeves, the untrained eye may never know the difference. But what about those hard-to-fit sizes? Great news! DW-NOMINATE has some off-labels that might just do the trick.

I have nothing against bespoke suits. But "you must buy one" is not good general advice. Some people will not be able to, and owning one is not a necessity. Similarly, I have a deep appreciation of the difficult methodological work designed to overcome DW-NOMINATE's shortcomings. But implementing the alternatives may prove quite challenging for

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

historically oriented scholars, and doing so may not be entirely necessary. Tailored DW-NOMINATE may be the best value for the money.

In fleshing out this argument, I address four areas for which DW-NOMINATE received the most criticism in this issue. First, I take up the issue of interpretation and the relative roles of ideology, partisanship, and constituency interest in determining ideal points. I argue that the strongest claim that ideal points (at least in the Gilded Age) should be interpreted as partisanship is not supported by evidence showing that NOMINATE does a very good job of identifying persistent intraparty cleavages—a finding hard to reconcile with currently theories of partisanship. Second, I address the issue of dimensionality. Here I stress that much of the criticism of DW-NOMINATE centers on a misinterpretation of the second dimension and its misapplication in empirical work. Correct application of the second dimension can be important both for uncovering important legislative cleavages and for understanding policy-specific preferences. Third, I address questions about policy-specific ideal points and dimensions. While DW-NOMINATE uses almost all votes cast across all issues, its results, when used appropriately, can elucidate the politics of very specific political issues. On the contrary, the suggested alternatives to DW-NOMINATE may provide very little additional insight and have problematic statistical properties. Finally, I take up the issues of the dynamics of ideal points. Because DW-NOMINATE assumes that each ideal point evolves linearly over time, its ideal point estimates may be inappropriate for studying abrupt or nonlinear change in legislative preferences. While this feature is certainly a drawback, the alternatives proffered have several of their own. Moreover, an extension of DW-NOMINATE deals with this problem directly and appears to perform at least as well as the customized alternatives.

In what follows, I attempt to keep the technical discussion to a minimum, but some focus on the details of various ideal point models is unavoidable. Thus, readers who are less familiar with the methodological work in this area would be advised to review Everson et al.'s contribution.

2. INTERPRETATION

A central concern in this issue is the ongoing debate about exactly what ideal point estimates such as DW-NOMINATE are actually capturing. The debate revolves around the fact that ideal points are a low-dimensional representation of all the considerations that go into roll call voting. But in isolation from other forms of evidence, the ideal points do not reveal exactly which political, psychological, and strategic factors that underlie the coalitions we observe in Congress and the other legislative bodies to which these models have been applied.

Drawing on the arguments of Phillip Converse, Keith Poole and Howard Rosenthal explain ideal points and their low dimensionality in terms of belief constraint.[1] This is the notion that political elites share beliefs about which issue opinions go together. Support for tax cuts is correlated with support for deregulation, which is correlated with higher defense spending, because elites believe that those issues are all related. But constraint is really little more than a description of the phenomena, not an explanation. There is no logical underpinning for the issue configurations that define left and right. Why exactly should a pro-life position predict support for tax cuts? Moreover, these configurations have changed over time. Protectionism used to be the province of the conservative. Ideal point estimation and dimensional analysis can identify when such changes occur, but it has more limited value in explaining why.

Several pieces in this issue take exception with the Poole-Rosenthal interpretation of ideal points as reflecting legislator ideology. In her piece, Lee argues that the ideal point estimates and related measures of polarization of the Gilded Age are little more than reflections of partisanship triggered by the intense competition for control over patronage. The articles by Bateman and Lapinski and by Caughey and Schickler weigh in against the ideological interpretation in slightly less unequivocal terms. While conceding that DW-NOMINATE scores may reflect ideological concerns to a degree, the authors argue that the conflation of other factors undermine any direct relationship between the scores and coherent ideologies such as liberalism or progressivism.

It is hard to argue for the proposition that DW-NOMINATE scores are pure measures of legislator ideology. Clearly, partisanship, constituency interests, and regionalism will impact the estimates of legislator positions on these scales. It is equally clear that even if the scores were purged of these other factors, the resulting ideological scales would not pass a political theory seminar. But this is also true of the "ideologies" that constitute much of the public discourse, or at least the part that goes on outside of elite political magazines.

Nevertheless, I contend that DW-NOMINATE scores largely reflect an ideology-like substance. Let's call this substance ideo-lite. Ideo-lite shares two features with ideology. First, ideo-lite shares the ability to link choices across different issues together. Unlike full-strength, political-theory-seminar ideology, which requires that the linkage be born of logical deductions from first principles, the issue

1. See Philip E. Converse, "The Nature of Belief Systems in Mass Publics," in *Ideology and Discontent*, ed. David Apter (New York: Free Press, 1964); Keith T. Poole and Howard Rosenthal, *Ideology and Congress* (New Brunswick, NJ: Transaction Publishers, 2011).

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

linkages in ideo-lite may be largely politically and socially constructed. Second, ideo-lite generates consistency across legislative behavior over time, just as we would expect of the full-strength version. Finally, ideo-lite is not reducible to party or constituency. Like ideology, we would expect to see correlations due to the selection effects related to the processes of party affiliation and elections, but ideo-lite has explanatory power even when those factors are accounted for.

There is a substantial amount of evidence that DW-NOMINATE ideal points contain ideo-lite. First, DW-NOMINATE idea points clearly have significant explanatory power for votes across a wide variety of substantive issues. That DW-NOMINATE is also very successful in capturing intraparty divisions suggests that the issue linkages that it has uncovered do not simply reflect partisan agendas. Second, ideal point estimates are quite stable for politicians throughout their career. Of course, there are a few prominent examples of politicians whose positions did change, some of which are highlighted in the Bateman-Lapinski and Caughey-Schickler pieces. But for the most part, legislators' ideal points only move significantly if they switch parties (and of course party switching is quite rare).[2] Even a member whose constituency changes quite dramatically, either by elevation to the Senate or through major redistricting, rarely changes positions in a significant way. In a very careful study, Poole shows that the assumption that legislators maintain the same ideological position throughout their careers performs just as well statistically as the assumption that legislators are able to change positions in each biennial term.[3]

The second piece of evidence in favor of ideo-lite is that the behavior of legislators deviates in large and systematic ways from the preferences of their average or median constituent. This finding persists even when the mismeasurement of constituency interests or preferences is not a concern. For example, senators from the same state rarely vote identically. Most obviously, senators from the same state but different parties, such as Sherrod Brown and Rob Portman of Ohio, vote very differently, and the difference is picked up in their polarized ideal point estimates.

If the two senators are from the same party, they are, of course, more similar. Even here, however, there are differences. Consider California Democrats Diane Feinstein and Barbara Boxer. They not only represent the same state but also were first elected by exactly the same electorate on the same day in 1992. In the 113th Senate term, Boxer has a DW-NOMINATE score of −0.486, making her the fifteenth most liberal member of the U.S. Senate. Conversely, Dianne Feinstein's ideal point is just −0.351 making her the thirty-third most liberal. Moreover, there is nothing unusual about this California duo. Seven other states have pairs of senators from the same party whose DW-NOMINATE scores differ at least as much.[4]

House districts, being single-member, do not allow the same natural experiment that is possible for the Senate. It is possible, however, to compare the voting behavior of a member to his successor. The same-party replacements of House members can have ideal points that are very different from those of their predecessors. True, a relatively liberal Democrat is likely to be replaced by another liberal Democrat, but the variation in the scores of the same-party replacements is very large. It is about half as large as the total variation of positions within the party.[5] In other words, the ideal point of the outgoing incumbent is at best a crude predictor of the position of the new member even if they are in the same party.

The evidence that ideo-lite is not simply partisanship is also fairly well established. DW-NOMINATE has become widely applied in large part because it has considerable explanatory power beyond simple party indicators. It also does a very good job of explaining internal divisions within the parties. This feature is hard to reconcile with the notion that the DW-NOMINATE dimension reflects "partyness." Indeed, if that were the case, Bernie Sanders would be the pillar of Democratness while Rand Paul and Mike Lee would anchor Republicanness. But in reality, the reason why these senators have extreme DW-NOMINATE scores is that they vote quite often against their partisan colleagues.

Lee takes the strongest stance that the DW-NOMINATE scores are merely reflections of partisanship. She conceptualizes the parties of the Gilded Age as "long" distributive coalitions motivated exclusively by control over the government and patronage. She rejects the notion that the parties were differentiated ideologically.

To address her arguments, I re-analyze the patterns of roll call voting during the Gilded Age and find substantial evidence for ideo-lite. In particular, I show that there are very high levels of consistency in the intraparty cleavages across the era. A rightist

2. Nolan McCarty, Keith T. Poole, and Howard Rosenthal, "The Hunt for Party Discipline in Congress," *American Political Science Review* 95 (2001): 673–87.

3. Keith T. Poole, "Changing Minds? Not in Congress!" *Public Choice* 131, no. 3 (2007): 435–51. I do agree with Bateman-Lapinski and Caughey-Schickler that the cases of meaningful ideal point change can be historical and substantively important. But these changes don't negate my general argument about ideo-lite as intellectuals also occasionally change their theory-seminar ideologies.

4. These are the same-party pairs from KY, MA, MN, SC, TX, UT, and WV. There would be eight if we included the difference between "socialist" Bernie Sanders and Democrat Patrick Leahy in Vermont.

5. See Keith T. Poole and Thomas Romer, "'Ideology,' 'Shirking,' and Representation," *Public Choice* 77, no. 1 (1993): 185–96.

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

Republican in one term tends to be a rightist in the next. A leftist Democrat tends to continue to be a leftist. Moderates tend to stay in the middle.

Such a finding is difficult to square with Lee's argument about long distributive coalitions. To see why, it is important to delve into the underlying theory. The foundational premise of parties as long coalitions is that nonpartisan distributive politics will be unstable. The current losers should gang up on the current winners and win next period. And as this process repeats, we would observe a lot of different winning and losing coalitions. This understanding of distributive politics has led a number of scholars to conceive of political parties as long coalitions that hang together to obtain control of the government and then distribute the spoils internally.[6] But the long coalition model simply moves the instability of distributive politics into the party without clearly specifying how the spoils are divided internally. If internal party politics mimicked party-free distributive politics, we would expect to see lots of instability within parties. Even if parties could implement a stable coalitional agreement, it is not clear why it would lead to roll call voting that produces temporally stable ideal points. Given that members with more extreme ideal points typically participate in fewer winning coalitions, why would any politician want to join a long coalition that would place her at an extreme ideal point? Such a legislator and her constituents would obviously be persistently disadvantaged, making them vulnerable to distributive offers from the other party. Thus, the stability of the long coalition would be compromised. Thus, a finding that DW-NOMINATE can explain persistent intraparty cleavages is hard to reconcile with the idea that parties are simply distributive coalitions.

To document my claims about the stability of intraparty coalitions, I estimate static W-NOMINATE ideal points for each Congress of the Gilded Age.[7] I then correlate the ideal points for all members across consecutive terms to measure the overall stability of the ideal points. I then repeat the exercise for the legislators within each party. I do not use DW-NOMINATE scores for the obvious reason that the change in member's ideal points are assumed to be linear over time, a feature that would inflate the congress-to-congress correlations. While W-NOMINATE scores are not directly comparable over time, the scales should differ only by a linear transformation so that correlations will not be affected by the comparability issue. One issue that works against my findings about stability is that static W-NOMINATE scores will contain substantial measurement error related to the relatively small number of roll calls used to estimate them. This measurement error will tend to attenuate the correlations toward zero.

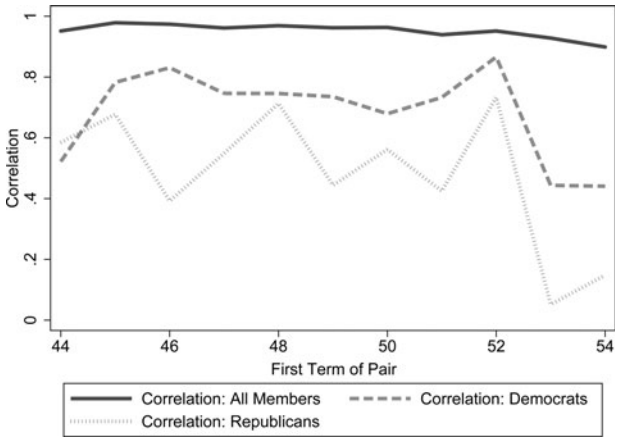


Fig. 1. Term over Term Ideal Point Correlation—House

Figure 1 presents my results for the Gilded Age House of Representatives. Each point represents the correlation of ideal points across a pair of congresses. The solid line shows the correlation for the entire House, which is consistently above 0.9. More interesting for my arguments are the within-party correlations. While the within-party correlations are obviously not as high as the overall ones, they are quite substantial. For Democrats, the correlation often exceeds 0.8. The Republican correlation is a little lower, around 0.5. Figure 2 tells a very similar story for the Senate. Overall correlations are always above 0.9, while the within-party correlations vary between 0.4 and 0.8.

It is important to remember that these figures represent conservative tests for intraparty consistency, given the probable measurement error in the static ideal point estimates. An alternative set of congress-by-congress ideal points less subject to measure error are those developed by Timothy Nokken and Keith Poole.[8] If the same analysis is carried out using those scores, the basic results are the same except that outliers such as the Republican Senate correlation from the 44th to 45th Congresses and the Republican House correlations from the 53rd through the 55th are much less pronounced.

6. For discussions of parties as "long" coalitions, see Thomas Schwartz, "Why Parties?" (Unpublished manuscript, Department of Political Science, University of California, Los Angeles, 1989); John H. Aldrich, *Why Parties? The Origin and Transformation of Party Politics in America* (New York: Cambridge University Press, 1995); Kathleen Bawn et al., "A Theory of Political Parties: Groups, Policy Demands and Nominations in American Politics," *Perspectives on Politics* 10, no. 03 (2012): 571–97.

7. These ideal points are estimated using W-NOMINATE for *R*. See Keith T. Poole et al., "Scaling Roll Call Votes with W-NOMINATE in R," *Journal of Statistical Software* 42, no. 14 (2011): 1–21.

8. Timothy P. Nokken and Keith T. Poole, "Congressional Party Defection in American History," *Legislative Studies Quarterly* 29, no. 4 (2004): 545–68. These scores are discussed in more detail in Section 4.

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

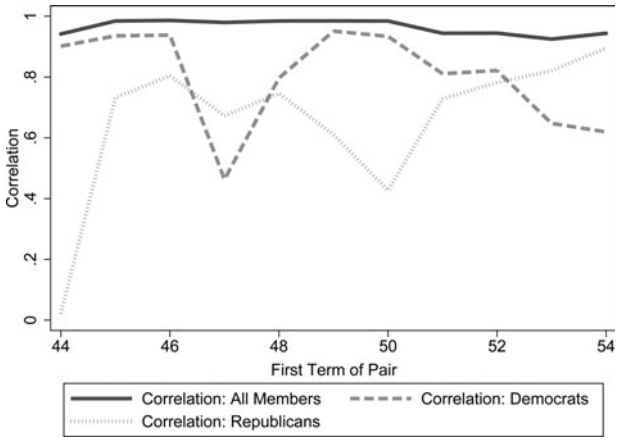


Fig. 2. Term over Term Ideal Point Correlation—Senate

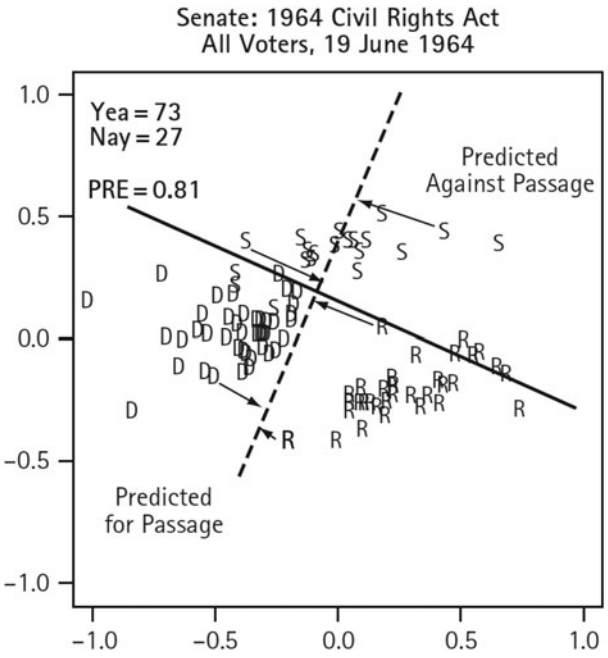


Fig. 3. Senate Passage Vote of Civil Rights Act

Lee concludes that "political parties are fully capable of waging ferocious warfare over spoils and office, even despite a relative lack of sharp party differences over national policy" (Abstract). That is undoubtedly true in principle, but the evidence I present suggests that during the Gilded Age, the parties received a substantial assist from ideo-lite.

3. DIMENSIONALITY

A key finding of the DW-NOMINATE research program is that conflicts within Congress can generally be represented by a single dimension. That dimension generally distinguishes between legislators' votes on issues related to economic policy and to the size and scope of the federal government. It has generally corresponded to colloquial, if not philosophical, definitions of the liberal–conservative divide. Analysis of the DW-NOMINATE model, however, has uncovered distinct time periods where a single ideological dimension was insufficient to explain important variation in roll call behavior. In such instances, a second dimension had substantial explanatory power.

During most of the episodes of "two-dimensional" politics, the second dimension was required to rationalize coalitions related to regional divisions centered around the issues of slavery, civil rights, and race. Consequently, the second dimension has often been called the "racial dimension." While such a labeling is convenient, it is technically incorrect and is the cause of considerable confusion, which carries over into both the Bateman-Lapinski and Caughey-Schickler pieces.

The second dimension is not a racial dimension in the sense that it is estimated to explain voting coalitions on racial issues. Rather, the second dimension is estimated as a policy dimension orthogonal to the first dimension that provides additional explanatory power on votes that do not fit the first dimension. These votes are often related to race. How a member votes on a racial issue is therefore a function of both his position on the first dimension and the second dimension. Analyses of the positioning on the second dimension such as Bateman and Lapinski's Figure 1 are, therefore, hard to interpret. To understand this issue, consider Figure 3, which represents Senate ideal points and voting on the 1964 Civil Rights Act. Clearly, the positions on the first dimension are inadequate to explain this vote, as they predict that many Southern Democrats would actually support the bill. But the second dimension alone does not do so well either. It would considerably overpredict the number of Republicans who voted for the bill. Instead, multidimensionality is reflected in the fact that the cutline for the bill deviates considerably from the vertical line one would expect for a first-dimension vote (yet it also is not horizontal, as would be expected by a purely second-dimension vote). But we can generate a single measure reflecting civil rights voting. First, we choose a line perpendicular to the cutline, such as the dotted line in the figure. Then we project the ideal points onto the perpendicular line. In other words, we move each ideal point to its closest point on the perpendicular line. The projection is illustrated by the arrows in Fig. 3. Note that the new positions do as good a job representing the vote as the two-dimensional configuration. Any senator with a projection below the intersection is predicted to support the measure and anyone with

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

a point above it is predicted to oppose it. Consequently, the projection represents the substantive civil rights dimension.

4. IDEAL POINT DYNAMICS

Both the articles by Bateman and Lapinski and by Caughey and Schickler note the limitations of DW-NOMINATE when it comes to capturing the dynamics of the positions taken by legislators. As each article discusses, the publicly released versions of DW-NOMINATE assume that legislators move linearly throughout their careers. This implies that the total movement of a legislator's ideal point is assumed to take place in equal increments over each term she serves. Poole and Rosenthal settled on this linear specification after extensive experimentation with more flexible polynomial trends for legislators. They found that overall legislators moved relatively little over the course of their careers and that little explanatory power was gained by fitting more complicated trends. But as the authors in this issue point out, there may be exceptions to Poole and Rosenthal's general findings, and these exceptions may be historically important. Moreover, cases of abrupt change may not have been well captured even by the higher order polynomials explored by Poole and Rosenthal. To this end, both sets of authors propose alternatives to DW-NOMINATE and demonstrate that there was significant movement of Southerners during the New Deal that was not well captured by DW-NOMINATE. I argue, however, that the methodological advice requires more caveats than those offered and that the evidence for the historical finding was already available in a publicly available version of DW-NOMINATE developed by Timothy Nokken and Keith Poole.[9]

Bateman and Lapinski suggest the following procedure. First, estimate static congress-by-congress ideal point estimates using either W-NOMINATE or the item response ideal point model developed by Joshua Clinton, Simon Jackman, and Douglas Rivers.[10] Second, use the method proposed by Tim Groseclose, Steven Levitt, and James Snyder, (hereafter GLS) to rescale the congress-by-congress results so that they are comparable over time.[11] This comparability is generated by the assumption that each legislator has a constant *expected* ideal point throughout her career. Like DW-NOMINATE, GLS also is based on the implicit assumption that the basic ideological space is constant.[12]

Bateman and Lapinski's claim that GLS estimates unrestricted movement in ideal points is based on the fact that the model produces each legislator's average ideal point and a residual for each term. These authors assume that average ideal point plus residual equals a member's session specific ideal point. Such an assumption is a very strong one, however. A nonzero residual from the GLS model, and therefore temporal variation in individual ideal points, can arise for a large number of reasons. It may indeed represent true change in the ideal point. But it may also arise from estimation error in the congress-by-congress ideal points. For example, assume that every member has the same true ideal point throughout her career. Let's call this ideal point x. In each session, due to limited sample size and the stochastic nature of roll call voting, we would estimate her ideal point as $\alpha_t + \beta_t(x + \varepsilon_t)$, where α_t and β_t are scale shift parameters estimated by GLS for session t, and ε_t is the measurement error in session t. Under these assumptions, the GLS model will correctly uncover x and ε_t. But because ε_t is measurement error, would we really want to say that $x + \varepsilon_t$ is the member's ideal point in session t? Second, the shift and stretch parameters will also be measured with error. Let $\hat{\alpha}_t$ and $\hat{\beta}_t$ be the estimates of these parameters in session t. Then the estimate of the session specific ideal point is $(\alpha_t - \hat{\alpha}_t) + \frac{\beta_t}{\hat{\beta}_t}(x + \varepsilon_t)$, which deviates from the true value whenever $\alpha_t \neq \hat{\alpha}_t$ or $\beta_t \neq \hat{\beta}_t$.

So, clearly the Bateman and Lapinski's interpretation of the GLS residual is capable of producing lots of false positive movement. But how well will it perform when there is substantial movement for a subset of members? If all members move a lot, the GLS will perform very poorly because the estimates of $\hat{\alpha}_t$ and $\hat{\beta}_t$ will be imprecise. Given this imprecision, there is no way of guaranteeing that GLS is recovering the true movement of the legislators. Thus, GLS can only be assumed to work if the set of legislators who move is small and that the remaining legislators are assumed to have constant positions. But I am unaware of any work that establishes how far we can deviate from the assumption that all members have constant preferences and still reliably pick up the ideal point movement of those who actually move.

Caughey and Schickler propose using the dynamic Bayesian ideal point model developed by Kevin Quinn and Andrew Martin.[13] This approach has some distinct advantages over the use of the GLS model. The most important is that it does not

9. Nokken and Poole, "Congressional Party Defection in American History."

10. Joshua Clinton, Simon Jackman, and Douglas Rivers, "The Statistical Analysis of Roll Call Data," *American Political Science Review* 98, no. 02 (2004): 355–70.

11. Tim Groseclose, Steven D. Levitt, and James M. Snyder, "Comparing Interest Group Scores Across Time and Chambers: Adjusted ADA Scores for the US Congress," *American Political Science Review* 93, no. 1 (1999): 33–50.

12. Thus, it is itself susceptible to the criticisms that Bateman and Lapinski levied in Section 2 of their article in this issue.

13. Andrew D. Martin and Kevin M. Quinn, "Dynamic Ideal Point estimation via Markov Chain Monte Carlo for the US Supreme Court, 1953–1999," *Political Analysis* 10, no. 2 (2002): 134–53.

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

assume that each legislator has a constant average ideal point. Instead, it makes the weaker assumption that a member's expected ideal point in period t is equal to her ideal point in period $t-1$. Formally, the model specifies that $x_t = x_{t-1} + \varepsilon_t$ where ε_t is a random shock with mean zero and variance σ^2. This statistical process, known as a *random walk*, has the advantage that it allows the expected ideal points of members to change over time.[14] A second advantage is that the Martin-Quinn model, unlike the GLS model, incorporates information in votes from time $t-1$ to estimate ideal points at time t. Thus, the dynamics are somewhat less likely to be driven by the estimation error of congress-by-congress estimates.

The Martin-Quinn model is not without drawbacks, however. The most important is that the estimated dynamics are determined by the variance of the random walk of ideal points, σ^2. If σ^2 were zero, then $x_t = x_{t-1}$, so that each legislator's ideal point is constant overtime. If σ were very large, x_t would be completely independent of x_{t-1} and the ideal points would not be comparable over time. Thus, correctly modeling legislator dynamics requires an accurate estimate of σ^2. Unfortunately, σ^2 cannot be estimated. Because low values of σ^2 act as a constraint on the statistical model, the likelihood function will always be maximized when $\sigma^2 = \infty$.[15] Thus, σ^2 has to be set at a fixed value a priori, and there is no way to determine whether it is set at its true value. The model, therefore, can be evaluated only by estimating several models with different values of σ^2 and showing that the results are robust. But invariably, higher values of σ^2 will be associated with greater movement, and it is impossible to determine exact values of σ^2 for which the observed movement comprises true ideal point changes instead of measurement error.[16]

The problem of how best to estimate dynamic models of ideal points is a thorny one, and the limitations of DW-NOMINATE in this regard are clear. But there other approaches that have recently been developed. A two-step procedure has been proposed by Timothy Nokken and Keith Poole.[17] In step 1, the full DW-NOMINATE model is estimated to generate estimates of the cutting lines for every vote. In step 2, ideal points are estimated congress by congress using the cutting line estimates from step 1. Thus, the ideal points from step 2 are comparable over time and can be used to test whether specific members change positions over time. One advantage of the Nokken-Poole estimates is that they are available for two dimensions. The second advantage to applied APD scholars is that, like DW-NOMINATE, the scores are available for download from http://voteview.com.

To assess the differences between these *prêt-à-porter* scores with customized alternatives, I replicate Caughey and Schickler's analysis of Southern senators during the New Deal using the Nokken-Poole estimates. To capture the dynamics of Southern positions on economic policy in the New Deal, Caughey and Schickler estimate a dynamic Bayesian ideal point model. To focus on economics, they only include votes related to economic policy and omit any related to race or international relations. For technical reasons, their model is limited to a single dimension despite the fact that the incremental classification gain for the second-dimension DW-NOMINATE in the Senate during the 1930s is greater than 4 percent. They defend the restriction to one dimension by noting that "even if voting on economic issues was occasionally two dimensional, the domain-specific ideal points can still be interpreted as weighted averages of MCs' support for liberal roll call positions, with weights determined by the correlation of voting patterns across bills" (sect. 3.2). The translation of this claim is that if economic voting were two dimensional, they would be estimating the projection into a single dimension. I will demonstrate, however, that focusing on a single dimension during this era is not costless.

I begin my replication with their analysis of the career of Theodore Bilbo. Figure 4 plots the Nokken-Poole score for Bilbo for the 74th through 79th Senates. The pattern that emerged follows that documented by Caughey and Schickler fairly closely. Consistent with their findings, Bilbo did not begin moving to the right on the first dimension until later in his

14. The difference between the Martin-Quinn model and GLS can be subtle in that each can produce any sequence of ideal points for a legislator. But the main difference concerns how each procedure adjusts ideal point estimates to ensure comparability over time. When the Martin-Quinn model estimates the scale for period t, it relies upon the estimates of x_{t-1}. GLS relies upon the overall average ideal points $\bar{x}$.

15. In more technical language, the parameter σ^2 is not *identified*.

16. Unfortunately, Caughey and Schickler do not report the value of σ^2 or describe the necessary robustness checks.

17. Nokken and Poole, "Congressional Party Defection in American History." Another newer alternative is Adam Bonica's dynamic extension of Keith Poole's nonparametric optimal classification (OC). See Adam Bonica, "The Punctuated Origins of Senate Polarization," *Legislative Studies Quarterly* 39, no. 1 (2014): 5–26; Keith T. Poole, "Non-parametric Unfolding of Binary Choice Data," *Political Analysis* 8 (2000): 211–37. The OC model estimates ordinal rankings of legislator ideal points instead of cardinal measures. Since rank orderings are independent of changes to the scale, OC need not worry about bridging across time. Movement can be assessed, for example, by noting that a member has shifted from the 85th percentile of conservatism to the 90th percentile. Since there is no issue of comparability over time, Bonica suggests estimating ideal points over fixed intervals of the agenda, such as every 200 votes. To minimize the likelihood of idiosyncratic shifts, Bonica suggests a smoothing process where a legislator's ideal point as on vote t is based on the votes from $t-200$ to $t-1$, while the ideal point at time $t+1$ is based on the votes from $t-199$ to t. Thus, Bonica's procedure produces a new ideal point ranking at each vote without the strong assumptions required of the GLS model (which are then violated if a legislator does in fact change position).

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

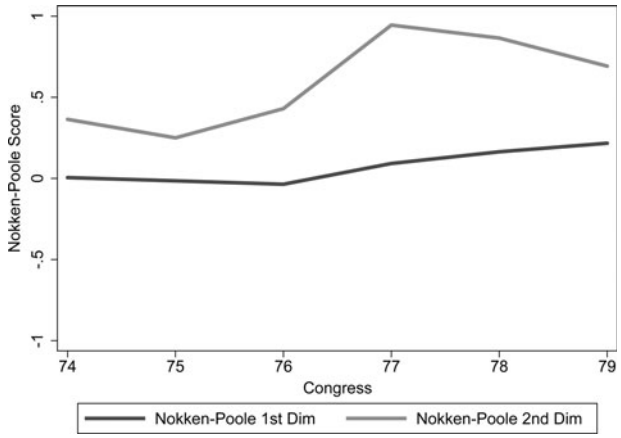


Fig. 4. Positioning of Theodore Bilbo on Nokken-Poole Scale

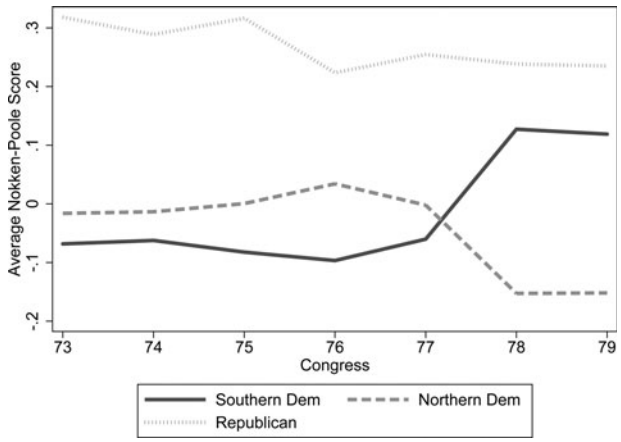


Fig. 5. Positioning of Senators on Nokken-Poole First Dimension

career. Nokken and Poole identify the shift during 77th Senate, while Caughey and Schickler suggest it was during the 78th. But Figure 4 shows that most of his movement was on the second dimension, where it shifted dramatically, also during the 77th Senate. This additional piece of information might be useful in unpacking the exact nature of his defection from the New Deal coalition.

Now I turn to the analysis of the positioning of the Southern Democratic bloc. In Figure 5, I plot the mean positions of Republicans, Northern Democrats, and Southern Democrats using the first-dimension Nokken-Poole scores. These results largely mirror that of Caughey and Schickler's Figure 6. The differences between the Northern and Southern Democrats were small until the 78th Congress (1943–44) when the Southerners moved sharply to the right and the Northerners moved sharply to the left. The only notable difference between the dynamic

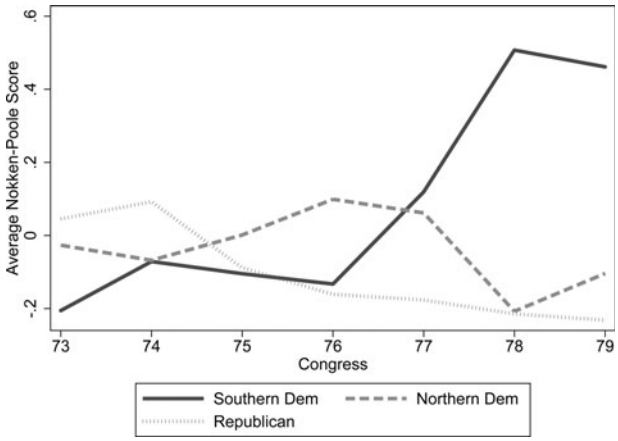


Fig. 6. Positioning of Senators on Nokken-Poole Second Dimension

Bayesian results and those of Nokken-Poole is that the latter show the Southern Democrats as slightly more liberal than the Northerners until the 78th Congress.[18]

Because they used a one-dimensional model, Caughey and Schickler may have missed a more substantial change in the behavior of the Southern senators. Figure 6 shows the positioning of the three groups of senators on the second dimension. The position of Southern senators on the second dimension changes dramatically between the 76th Senate (1939–40) and the 78th (1943–44). These results suggest that not only did Southern senators become more like Republicans, but they also became more regionally distinctive even before race relations and civil rights became an explicit part of the agenda.

When it comes to the first-dimensional estimates where the use of GLS, dynamic Bayesian models, and Nokken-Poole are all available alternatives, there is no conclusive evidence that any choice is necessarily the correct one. But for many applied APD scholars, the ready availability of Nokken-Poole results and the information about the second dimension should make using those scores an attractive option.

4.1. Policy-Specific Ideal Points

A major thrust of the Bateman and Lapinski piece is that DW-NOMINATE may not be appropriate for APD scholars who wish to study the trajectory of congressional voting across a single policy issue. The crux of the criticism lies in the fact that DW-NOMINATE uses (almost) all roll call votes in constructing its ideological measures. Therefore, the DW-NOMINATE positions are a weighted average of positions across a

18. The difference may be the result of the fact that Caughey and Schickler estimated their model only on explicitly economic votes.

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

number of political issues and therefore may be sensitive to the policy content of the agenda. While these criticisms have some validity (if not originality), there are many reasons to believe that the issue is not that important in practice and that Bateman and Lapinski's cure may be worse than the disease.

That the two-dimensional DW-NOMINATE models explains such a large proportion of the observed roll call voting suggests that if coalition patterns vary across policy issues, they do not do so significantly. These findings suggest that most policy issues project very nicely onto the DW-NOMINATE dimensions. And as discussed above, most issues except for regional cleavages (often but not always related to race relations and civil rights) map onto a single dimension. Thus, it is not surprising that numerous studies find that the DW-NOMINATE scores are the single largest predictor of policy-specific roll calls. This is not only because DW-NOMINATE's first dimension captures partisan differences, but also because DW-NOMINATE scores are generally very good at capturing intraparty disagreements.

Of course, this is not to suggest that issues-specific deviations from DW-NOMINATE are not interesting or important. Indeed, such deviations have been a central part of the literature spawned by DW-NOMINATE. A central feature of every DW-NOMINATE release has been a set of measures showing how well each vote fits the one- or two-dimensional model. Perhaps the most useful measure is the proportional reduction of error (PRE). The PRE computes how much the level of prediction error associated with random voting is reduced by the application of the model. The PRE can be assessed for a single vote or aggregated into a measure assessing the fit of a group of votes. Therefore, the aggregate PRE (APRE) is ideal for assessing the extent to which a policy area fits the DW-NOMINATE model. It can also be used to assess the relative importance of the first and second dimensions. Moreover, APREs for specific issues may be plotted over time to assess how the fit between DW-NOMINATE and issue-specific coalitions changes over time. As an example, consider the analysis I conducted with Poole and Rosenthal.[19] We were interested in how three broad categories of votes (civil rights, labor and minimum wage, and government management) map into the first and second DW-NOMINATE dimensions.[20] By plotting the APREs for the first dimension and the incremental increase associated with the second dimension, the figures reveal how mapping from these issue areas into the basic space has changed since the 1930s. Civil Rights projected primarily onto the second dimension until the mid- to late 1960s and, since then, has become increasingly tied to the first dimension. Labor votes during the Second New Deal were a mixture of the first and second dimension but have become increasingly tied to the first dimension since the 1940s. Government management votes have always been closely tied to the first dimension. Such analyses can be straightforwardly conducted for any issue codings that a scholar wishes to employ.

Perhaps one would like to go further than measures of fit to identify positions on an issue-specific scale. Again, off-the-rack DW-NOMINATE may be all an applied scholar needs. As I described in Section 3, any roll call vote can be projected from the two-dimensional DW-NOMINATE model into a single policy dimension using estimates of ideal points and the cutting lines, each of which are included in the DW-NOMINATE data release. Such a procedure can be done of a set of votes either by averaging the policy scores across a set of votes or by applying the average parameters of the cutting lines. These procedures can be accomplished with a minimal amount of technical know-how. It requires only computing a weighted average of each legislator's two-dimensional ideal point and averaging these scores across votes.[21]

Finally, the best procedure may be the easiest: multiple regression of a set of policy votes on DW-NOMINATE scores. By including additional explanatory variables for policy-specific constituency factors and ideological preferences, the researcher can provide substantive meaning to the variance in the policy votes that cannot be accounted for by DW-NOMINATE scores. Recent work by Atif Mian, Amir Sufi, and Francesco Trebbi provides a useful example.[22] In their study, they wished to understand the congressional voting patterns on several votes in 2008 on emergency measures aimed at the housing crisis. Thus, they estimated several regression models of roll calls that include the first-dimension DW-NOMINATE as a regressor. Focusing on Republican House members, they found that a larger DW-NOMINATE score is strongly associated with opposing the emergency measures.[23] But they also found that a number of other factors explain substantial variation in the voting behavior, most notably the mortgage default rates in the member's district (especially the default rates in Republican zip codes). Thus, their article shows exactly how one can use a broad measure like

19. Nolan McCarty, Keith Poole, and Howard Rosenthal, *Polarized American: The Dance of Ideology and Unequal Riches* (Cambridge, MA: MIT Press, 2006).

20. Ibid., Figures 2.19a–c.

21. For each roll call, DW-NOMINATE estimates of directed distances d_1 and d_2 and the dimensional weight w. If a legislator's two-dimension ideal point is (x_1, x_2), then her projected ideal point is $\frac{d_1 x_1 + w d_2 x_2}{\sqrt{d_1^2 + w^2 d_2^2}}$.

22. Atif Mian, Amir Sufi, and Francesco Trebbi, "The Political Economy of the US Mortgage Default Crisis," *American Economic Review* 100, no. 5 (2010): 1967–98.

23. These results further underscore that DW-NOMINATE scores are not simply measures of partisanship. DW-NOMINATE explains significant within-Republican party variation on a set of measures proposed by a Republican administration.

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

DW-NOMINATE as part of a study on a very specific policy intervention.

Although DW-NOMINATE (used correctly) addresses many of the issues raised by Bateman and Lapinski, it is not clear that the approach they advocate offers significant advantages. Their approach involves using subsets of votes defined by policy area to estimate static spatial models and then applying the GLS model to facilitate comparisons over time. The primary concern with this approach is the performance of static ideal point estimates based on small subsamples of votes. There are known precision problems associated with ideal point estimates based on a small number of votes.[24] So unless the policy subsamples are large, the policy-specific ideal points will be very noisy. But large policy samples may defeat the purpose of policy-specific ideal points, as they would necessarily contain a heterogeneous mix of issues. To the contrary, the projection procedure outlined above can produce an issue scale based on a single vote.[25] The application of the GLS model may also be problematic given the substantial measurement error in the static estimates. As I pointed out in Section 4, the ideal points produced by GLS are difficult to interpret in the presence of measurement error because we do not know whether the legislator has actually repositioned or whether the changes over time are due to poorly measured static preferences. Moreover, the shift and stretch parameters will be poorly estimated when there is substantial error in the static estimates. This compounds the problems associated with interpreting legislative movement.

Caughey and Schickler's analysis of voting during the Progressive Era provides a nice example of how projecting from the two-dimensional version of DW-NOMINATE can help account for policy-specific dimensions. In their article, Caughey and Schickler argue that the DW-NOMINATE model does a poor job of capturing the most salient ideological divides of the 1920s. To support this argument, the authors compare DW-NOMINATE scores with a set of ideal points estimated on votes selected by a prominent progressive group in the 1920s, the Conference for Progressive Political Action (CPPA). Their main finding is that DW-NOMINATE does a poor job of accounting for the votes on the CPPA agenda. While the first dimension of DW-NOMINATE correlates reasonably well with the CPPA scores of Republican senators, the correlation for Democrats is almost zero. The correlation of the second dimension for all senators is also essentially zero, although the within-party correlations are reasonably high. Given this evidence, Caughey and Schickler conclude that DW-NOMINATE measures "merely the divergence between the roll call patterns of the two parties on the main dimension of congressional conflict, not … ideological differences on a liberal–conservative scale" (sect. 2.2).

While the Caughey-Schickler analysis does raise some important caveats about the interpretation of DW-NOMINATE during the Progressive Era, their analysis overstates the case against DW-NOMINATE. The main concern with the Caughey-Schickler analysis is that it misapplies the two-dimensional DW-NOMINATE model by assessing the role of the second dimension in isolation from the first. This is exactly the issue I discussed in Section 3. The Progressive Era does have a distinctive second dimension. The two-dimensional DW-NOMINATE model correctly classifies 2–4 percent more roll call decisions over the period in question than the one-dimensional version. It is therefore important to correctly take both dimensions into account. When one does so, the DW-NOMINATE model does a remarkably good job accounting for the CPPA scores.

To illustrate this point, Table 1 replicates Caughey and Schickler, where I report the results in terms of regressions instead of correlations. The first three columns present the results comparing the first-dimension DW-NOMINATE with the CPPA scores for all senators and for each party separately. The overall R^2 and that for Republicans is fairly high. But for Democratic senators, the coefficient on DW-NOMINATE is not statistically different from zero and the R^2 is close to zero. So the results of the first-dimension DW-NOMINATE are anomalous for Democrats.

The next three columns show the results for the second dimension in isolation. In column 4 (all senators), the coefficient on DW-NOMINATE is not significant, and the R^2 is low. But the coefficients are significant, and the R^2 is substantial for both parties separately (columns 5 and 6). But as I explained above, columns 4–6 are not the correct way to assess the role of the second dimension. That dimension is not interpretable without controlling for the first dimension. To better capture the extent to which the two-dimensional DW-NOMINATE model predicts the CPPA scores, I regressed the CPPA score on both DW-NOMINATE dimensions. These models are reported in columns 6–9. The results alter the conclusions substantially. The two-dimensional model explains CPPA scores quite well. For all senators, the coefficients on both dimensions are highly significant, and the R^2 is 0.753. This finding implies that the projection of the two-dimensional DW-NOMINATE scores correlate with the CPPA scores at 0.87. This correlation is almost as high as the correlations we observed with DW-NOMINATE and interest group ratings in the

24. For a nontechnical discussion, see Nolan McCarty, "Measuring Legislative Preferences," in *Oxford Handbook of the American Congress*, ed. Frances Lee and Eric Schickler (New York: Oxford University Press, 2011), p. 82.

25. Of course, this feat is accomplished because the projection procedure leverages information from all of the other votes and the relationship between the vote in question and the basic DW-NOMINATE model.

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

Table 1. Projections of DW-NOMINATE to CPPA Scores

	D1-All	D1-D	D1-R	D2-All	D2-D	D2-R	Both-All	Both-D	Both-R
DW-NOM 1	−2.442	−0.664	−3.578				−3.035	−2.022	−2.775
	(0.157)	(0.712)	(0.332)				(0.147)	(0.624)	(0.348)
DW-NOM 2				−0.164	0.998	1.623	0.889	1.262	0.839
				(0.185)	(0.216)	(0.218)	(0.106)	(0.217)	(0.189)
N	145	65	78	145	65	78	145	65	78
R-Squared	0.629	0.0136	0.604	0.00546	0.253	0.421	0.753	0.361	0.687

Standard errors in parentheses

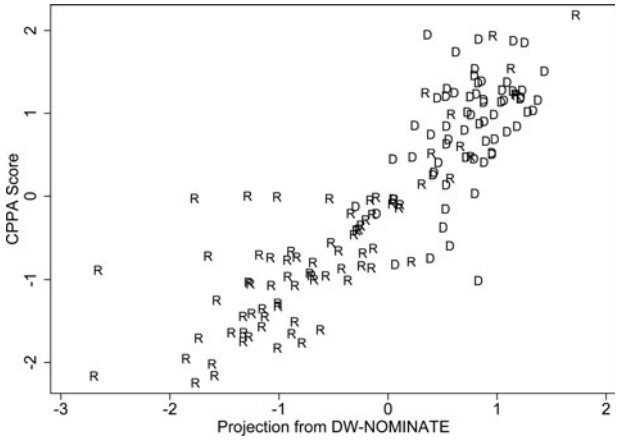


Fig. 7. Projection of DW-NOMINATE into CPPA Scores

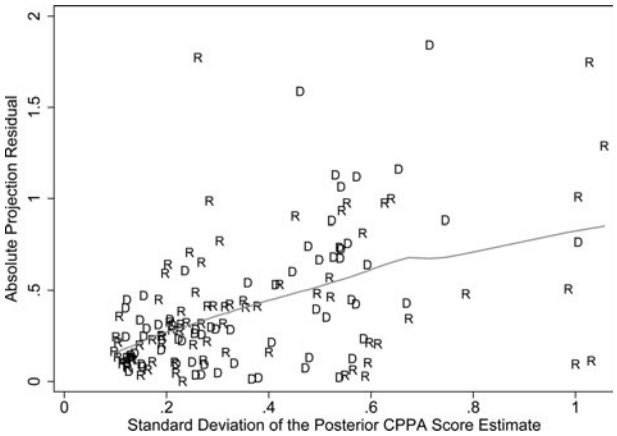


Fig. 8. Projection Residuals and Estimation Uncertainty

contemporary era. The results for the within-party models are weaker, but only slightly so. Even for the Democrats, both dimensions are highly significant. The R^2 of 0.361 implies that the projected DW-NOMINATE scores correlate with CPPA at 0.6. Figure 7 plots the overall projection against the CPPA scores. It is clear that the projection tracks the CPPA scores well and that there are no identifiable clusters of senators for which the projections perform poorly.

Instead, there is evidence that estimation error in the CPPA scores is likely behind the less-than-perfect correlations. Recall that the CPPA scores reflect only seventy-five votes over three Senate terms. That is not a lot of votes with which to precisely estimate the positions of more than 140 senators. Thus, the modest fit of the DW-NOMINATE model may be due to imprecise CPPA score estimates. To illustrate this possibility, Figure 8 plots the absolute value of the residual from the projection against a measure of uncertainty about the CPPA scores (the standard deviation of the posterior mean). Clearly, the residuals from the projection are largest when the estimation uncertainty of the CPPA scores is the highest. If I estimate the projection using the senators with below-median-estimation error, the projection correlates with the CPPA scores at the 0.924 level.

While the (correctly interpreted) DW-NOMINATE model does a good job of capturing the "progressive dimension," some puzzles remain. Why is the second dimension required to capture the divisions over progressivism? Why is the second dimension more important for explaining the divisions with the Democratic party? One possible answer lies in the restricted movement of ideal points assumed by DW-NOMINATE. The restriction of movement to linear trends may make it more difficult to capture emerging political divisions such as the one underlying progressivism. To assess this possibility, I compare the CPPA scores to static W-NOMINATE scores for each of the three Senate terms that are covered by the CPPA votes.[26] In using these static scores, the estimates of ideal points for each senator can change freely from term to term. Table 2 reports the projection of the static W-NOMINATE scores into CPPA scores. These results show that the projection of the static scores performs somewhat

26. These ideal points are estimated using W-NOMINATE for *R*. See Poole et al., "Scaling Roll Call Votes with W-NOMINATE in R.".

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

Table 2. Projections of Static W-NOMINATE into CPPA Scores

	66th-All	66th-D	66th-R	67th-All	67th-D	67th-R	68th-All	68th-D	68th-R
1st Dim	0.931	0.672	2.690	−1.306	0.0379	−1.360	−1.773	−1.059	−1.758
	(0.0888)	(0.271)	(0.533)	(0.0600)	(0.244)	(0.214)	(0.0943)	(0.306)	(0.322)
2nd Dim	−1.675	−1.281	−2.333	−0.819	−0.850	−0.859	−0.270	−0.476	−0.218
	(0.154)	(0.229)	(0.236)	(0.102)	(0.173)	(0.174)	(0.111)	(0.132)	(0.293)
N	100	49	51	104	39	65	100	43	55
R-Squared	0.724	0.406	0.673	0.845	0.405	0.813	0.794	0.507	0.616

Standard errors in parentheses

better than the projection of DW-NOMINATE, suggesting that the restriction of DW-NOMINATE to linear ideal point changes comes at some cost. But the differences are not huge, suggesting that the answer to the puzzles lies elsewhere.

The answer probably lies in the substance of 1920s politics. Progressivism was hardly the sole challenger to conservatism. The Democratic left wing was occupied both by progressives and agrarian populists. The results from DW-NOMINATE reveal that, at least in terms of explanatory power, the left–right spectrum was one that ran from the Southern populists to the Northeastern conservatives. The spectrum from progressives to conservatives therefore requires a second dimension. That the left wing of the Democratic party is split between populists and progressives while the right wing of the Republican party is fairly homogeneous suggests that the second dimension would be more important in explaining variation within the Democratic party.

Some evidence that the DW-NOMINATE first dimension is picking up a left–right dimension distinctive of progressivism can be gleaned by looking at the important votes that are not part of the CPPA agenda. If one looks at the votes that are best explained by the DW-NOMINATE model in the Progressive Era, many are not included on the list of the CPPA. For example, the best fitting votes for the 67th Senate were ones related to the ratification of the Treaty of Versailles. In the 69th, the DW-NOMINATE excels at predicting votes on the rates of the Fordney-McCumber Tariff. But votes on these issues were not scored by the CPPA. Only in the 69th Senate does the CPPA list the best fitting DW-NOMINATE roll calls when it included several votes on the Revenue Act of 1924 (the Mellon Tax bill). While the Versailles Treaty and the Fordney-McCumber bill may not have been central to progressivism, it is hard to say that these issues were not connected to the ideological divisions of the era. Perhaps one can object on conceptual grounds that they are not technically the issues that divided conservatives and liberals, but they clearly divided conservatives and nonconservatives.

The Caughey-Schickler piece does raise one important caveat to the conventional understanding of the DW-NOMINATE's second dimension. Poole and Rosenthal have generally labeled this dimension "race" and "civil rights" on the basis of its role in distinguishing Southern from non-Southern legislators. But as Caughey and Schickler's results show, the second dimension can be crucial in distinguishing these regional blocks even when racial issues are not explicitly on the agenda. Indeed, the second dimension also may be important for distinguishing between different types of economic ideologies.[27]

5. CONCLUSIONS

If the goal of this special issue was to introduce scholars to a robust discussion of the pros and cons of ideal point estimation and DW-NOMINATE, I think it has been largely successful. Ideal point estimation has been the topic of vigorous discussions among methodologists and legislative scholars for decades. So I hope that APD and other historically oriented scholars can benefit from the exposure to these well-rehearsed debates.

I also hope that the unsettled nature of these questions does not provide a deterrent to those who might otherwise find ideal point measures and methodologies useful in their own work. Indeed, the theme of my comments is that shortcomings and ambiguities related to DW-NOMINATE and ideal point estimation are not as significant as they are sometimes portrayed. A lot can be accomplished buying off the rack at the voteview.com warehouse. With a touch of tailoring, many significant methodological issues can be resolved.

But I also wish to make clear that I do not think DW-NOMINATE should be the last word in ideal point estimation. There are many significant, open methodological questions whose future resolution will further aid historical scholarship. One example is the research agenda that estimates more theoretically driven models of legislative behavior. Incorporating strategic voting and logrolling into ideal point models has been limited. Although the basic model can accommodate simple forms of strategic voting, richer models of strategic voting, such as those under incomplete information, are just now being considered. Similarly,

27. Recall from Section 4 that the second dimension may also explain regional differences on votes when the connection to race relations is not explicit.

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press

logrolling and vote trading are inconsistent with the assumptions underlying the basic model. How vote trading affects ideal point estimation and confounds the evaluation of dimensionality clearly needs further work, such as the use of scaling models to detect vote trading and other forms of legislator coordination.

While such technical advances will be welcome, even more work needs to be done in terms of interpreting the implications of ideal point estimates for our understanding of politics. In particular, we still know very little about the dynamics of how issues map (or not) into the major dimensions of conflict over time. Progress on this front will entail significantly more attention to methodological challenges of estimating the spatial local of policy outcomes in addition to ideal points. Success will invariably require bringing substantially more substantive information directly into the process of estimating ideal points and would therefore be an area where historically oriented scholars will make key contributions to the methodology.

https://doi.org/10.1017/S0898588X16000110 Published online by Cambridge University Press