Studies in American Political Development, 30 (October 2016), 147–171.
doi:10.1017/S0898588X16000080
ISSN 0898-588X/16
© Cambridge University Press 2016

Ideal Points and American Political Development: Beyond DW-NOMINATE

David A. Bateman, *Cornell University*
John Lapinski, *University of Pennsylvania*

This article aims to persuade historically oriented political scientists that ideal point techniques such as DW-NOMINATE can illuminate much about politics and lawmaking and be very useful to better understanding some of the key questions put forward by American political development (APD) scholars. We believe that there are many lines of inquiry of interest to APD scholars where ideal point measure could be useful, but which have been effectively foreclosed because of the assumptions undergirding DW-NOMINATE. In particular, we focus on three issues as particularly important: (1) the assumption of linear change; (2) the collapsing of distinct policy issue areas into a single "ideology" score; and (3) an agnosticism toward policy development, institutional context, and historical periodization. We go over these issues in detail and propose that many of these concerns can be addressed by taking seriously the proposition that policy substance, historical and political context, and the temporal dimension of political processes be integrated into the core of our measures and analyses. We also discuss a set of techniques for addressing these issues in order to answer specific questions of broad interest to both APD scholars and other Americanists.

1. INTRODUCTION

There is a long tradition of research on legislatures that seeks to find meaningful patterns in the recorded divisions or roll call votes. The most ubiquitous of these research projects in recent years has been the various NOMINATE measures produced by Keith Poole and Howard Rosenthal, which have fundamentally transformed the analysis of congressional politics. The DW-NOMINATE series in particular covers the entire history of Congress since 1789, generating information on nearly every non-unanimous vote and on the preferences of nearly every member of Congress. This analysis of literally millions of individual roll call votes by legislators is perhaps the most important empirical project in the study of American politics. By providing scores of legislator ideology comparable across time, Poole and Rosenthal have made the statistical analysis of roll call data central to the study of Congress, and they have greatly influenced scholars of the presidency, political parties, and American political history.

But scholars have often desired to leverage the information in legislative behavior to better understand political dynamics that range beyond the legislatures themselves.[1] By systematically recovering information about the structure of political conflict across different time periods and institutional settings, spatial maps of parliamentary voting such as those provided by NOMINATE can significantly advance our understanding of how legislatures interact with other political institutions and social forces to shape policy, politics, and political regimes. Given the centrality of Congress to American political life, its central role in shaping the state and economy, and the detailed historical information generated by Poole and Rosenthal, there seems to be a rich possibility for better integrating these measures into the study of American political development (APD).[2]

For the most part, however, this is a possibility that has largely been missed. Of eighty-eight articles published in *Studies in American Political Development* since 1995 that examine legislative dynamics in Congress, only ten used any variant of the NOMINATE

1. Duncan MacRae Jr., *Issues and Parties in Legislative Voting* (Chicago: University of Chicago Press, 1970), 1, 4.

2. The DW-NOMINATE and other ideal point scores are only one facet of the considerable resources compiled and hosted on the www.voteview.com website. Keith Poole and Howard Rosenthal have provided a central and easily accessible site for much of the information on roll call votes and members of Congress compiled by the Inter-university Consortium for Political and Social Research and have undertaken the not inconsiderable task of cleaning up, systematizing, and updating these data. Even without the different measures that they have generated, which are themselves essential for most students of Congress, the site would be an indispensable resource for students of congressional history.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

scores.[3] While this is more than the four articles that rely on Rice cohesion, likeness, or party unity scores, the low levels of use suggest that APD scholars have not found much value in these measures.[4] We argue that this is unfortunate, but also that it is not without good reason. Many of the assumptions undergirding the DW-NOMINATE project cut against motivating concerns that have long been at the core of APD as a subfield: the reduction of "ideology" to a coordinate along one or two dimensions, the projection of members into a trans-historical space organized around unchanging quantities of "liberalism" and "conservatism," and the measures' fundamental agnosticism to the substance of policy or the political or institutional context within which policy is being formulated. The DW-NOMINATE project, in this regard, perhaps exemplifies the trade-offs between complexity and parsimony in political science, providing researchers with a single measure of ideology that can ostensibly be applied across time but which relies on assumptions that many historically oriented scholars might find untenable.[5]

The argument of this article is that this does not need to be the case, and our hope is to contribute to the growing effort to bridge the divide between APD and other lines of scholarship in American politics.[6] In short, we aim to persuade historically oriented political scientists that ideal point techniques such as DW-NOMINATE can illuminate much about politics and lawmaking. But for such a project to succeed, we need measures that are sensitive to the methodological concerns of APD scholars that these impose too rigid a structure on the data, assume an unwarranted equivalency across historical periods and institutional contexts, and make overly broad inferences about their substantive meaning. We believe that there are many lines of inquiry of interest to APD scholars where ideal point measures could be useful, but that have been effectively foreclosed because of the assumptions undergirding DW-NOMINATE. We argue, however, that many of these concerns can be addressed by taking seriously the proposition that policy substance, historical and political context, and the temporal dimension of political processes should be integrated into the core of our measures and analyses.

Before we proceed, there are two points that we want to stress that will come up frequently in the pages that follow. First, the measures we will introduce, which we think will be appealing to APD scholars, have their own limitations. They will not address all of the issues that may concern historically oriented scholars. But we hope to make a persuasive case that, when used judiciously and in dialogue with other sources, they can inform and enrich our historical work. And second, the measures we introduce are by no means objectively better than those we critique. All statistical techniques provide highly structured, stylized representations of reality; but changing the structure that is imposed, as we propose, can certainly increase the complementarity between the measures used and specific research questions.

3. Four of these articles had Jeffrey A. Jenkins as first author, suggesting that the range of APD scholars who have found the scores to be useful is even smaller than the count of their use implies. Moreover, several used NOMINATE in only a very limited way, such as using the scores to identify a single member of Congress as liberal or conservative. To be clear, this is not a critique of the authors, who appropriately used the scores to advance their particular research questions. Nor does it mean that these measures are not being used in other sources, such as non-*Studies* articles and in books. But we suggest the lack of use in the subfield's premier journal is a reflection of the degree to which these scores have been found wanting by many APD scholars interested in Congress. See Jeffrey A. Jenkins, "Partisanship and Confederate Constitution-Making Reconsidered: A Response to Bensel," *Studies in American Political Development* 13 (1999): 279–87; Jeffrey A. Jenkins, "Partisanship and Contested Election Cases in the House of Representatives, 1789–2002," *Studies in American Political Development* 18 (2004): 112–35; Jeffrey A. Jenkins, "Partisanship and Contested Election Cases in the Senate, 1789–2002," *Studies in American Political Development* 19 (2005): 53–74; Jeffrey A. Jenkins and Timothy P. Nokken, "Legislative Shirking in the Pre-Twentieth Amendment Era: Presidential Influence, Party Power, and Lame-Duck Sessions of Congress, 1877–1933," *Studies in American Political Development* 22 (2008): 111–40; Robert C. Lieberman, "Weak State, Strong Policy: Paradoxes of Race Policy in the United States, Great Britain, and France," *Studies in American Political Development* 16 (2002): 138–61; M. Stephen Weatherford, "Presidential Leadership and Ideological Consistency: Were there 'Two Eisenshowers' in Economic Policy," *Studies in American Political Development* 16 (2002): 111–37; Wendy J. Schiller, "Building Careers and Courting Constituents: U.S. Senate Representation, 1889–1924," *Studies in American Political Development* 20 (2006): 185–97; David Karol, "Has Polling Enhanced Representation? Unearthing Evidence from the Literary Digest Issue Polls," *Studies in American Political Development* 21 (2007): 16–29; Daniel Carpenter and Gisela Sin, "Policy Tragedy and the Emergence of Regulation: The Food, Drug, and Cosmetic Act of 1938," *Studies in American Political Development* 21 (2007): 149–80; Richard M. Valelley, "The Reed Rules and Republican Party Building: A New Look," *Studies in American Political Development* 23 (2009): 115–42.

4. Two of the articles using these other roll call–based measures of voting behavior are also by Jenkins, "Partisanship and Contested Election Cases . . . House," and "Partisanship and Contested Election Cases . . . Senate." The remaining two are Scott C. James, "Building a Democratic Majority: The Progressive Party Mote and the Federal Trade Commission," *Studies in American Political Development* 9 (1995): 331–85; and Sean Farhang and Ira Katznelson, "The Southern Imposition: Congress and Labor in the New Deal and Fair Deal," *Studies in American Political Development* 19 (2005): 1–30.

5. Gregory J. Wawro and Ira Katznelson, "Designing Historical Social Scientific Inquiry: How Parameter Heterogeneity Can Bridge the Methodological Divide between Quantitative and Qualitative Approaches," *American Journal of Political Science* 58 (2014): 526–46.

6. See, for example, Devin Caughy and Eric Schickler, "Public Opinion, Organized Labor, and the Limits of New Deal Liberalism, 1936–1945," *Studies in American Political Development* 25 (2011): 162–89; Sean Farhang, *The Litigation State: Public Regulation and Private Lawsuits in the U.S.* (Princeton, NJ: Princeton University Press, 2010); Steven White, "The Heterogeneity of Southern White Distinctiveness," *American Politics Research* 45 (2014): 551–78; Wawro and Katznelson, "Designing Historical Social Scientific Inquiry."

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

We begin in Section 2 with a discussion of some of the more relevant limitations to integrating DW-NOMINATE and APD scholarship. We focus on three issues as particularly important: (1) the assumption of linear change; (2) the collapsing of distinct policy issue areas into a single "ideology" score; and (3) an agnosticism toward policy development, institutional context, and historical periodization. Our purpose, however, is not simply to provide an overview of new measures. Instead, we hope to persuade the reader that such measures can make an important contribution to different lines of inquiry of importance to APD scholarship. In Section 3 we discuss a set of techniques for addressing these issues in order to answer specific questions of broad interest to both APD scholars and other Americanists. We conclude in Section 4 by asking scholars working across different methodological orientations to consider the opportunities that exist to enhance the quality of our research by taking seriously the concerns of historical political science and using these to inform our statistical measures and our analyses.

2. DW-NOMINATE AND AMERICAN POLITICAL DEVELOPMENT

Given the centrality of the DW-NOMINATE enterprise to congressional scholarship and the study of lawmaking, it is striking that it has received so little attention from APD scholars whose work touches on Congress as a site for the construction of public policy. Part of the reason for this might simply be attributed to differences in methodological orientation, with APD scholars inclined more toward qualitative modes of analysis as opposed to quantitative ones. We suggest, however, that the lack of engagement also reflects reasonable concerns about the appropriateness of the measures' undergirding assumptions for addressing the types of questions APD scholars find most interesting.

In this section we outline some of these assumptions and highlight the sorts of questions that they foreclose. The basic mechanics of DW-NOMINATE and similar procedures has been gone over in detail elsewhere, and so we focus our discussion on three points that we believe are under-appreciated and consequential obstacles for integrating such scores into historical analyses.[7] First, there is no intrinsic metric to ideal points, and so for different institutions or the same institutions over time to be compared, certain constraints need to be imposed. The implications of this for historical scholarship are not always fully appreciated. Second, despite its ubiquity as a measure, the meaning of the dimensions estimated by DW-NOMINATE is unclear and unstable over time. As we will argue, this limits the degree to which the measures are useful for detailed historical work. Third, the scores are agnostic to other sources of information, and thereby impose an assumption that legislator voting patterns are invariable to historical context, to institutional change, and to policy development. This assumption in particular is unsatisfying to those who believe that history is not simply a repository of data, but that its study requires sustained attention to institutional and historical context.

2.1 Linear Change and Static Scores

There is little question that DW-NOMINATE has become central to congressional studies and to the historical turn in this field because of the groundbreaking work of Poole and Rosenthal in estimating cross-time and cross-chamber coordinates. One of the main reasons for its widespread adoption in much of Americanist political science is its seeming ability to claim that "Jesse Helms is more conservative than Robert Taft, Sr. even though they never served in the Senate together."[8] Nevertheless, there has been relatively little nontechnical discussion or debate about the assumptions and constraints needed to establish temporal or cross-institution comparability, nor has there been a full consideration of the type of research questions that these constraints effectively rule out.

The basic problem of bridging across institutions or across time is that the metric of the space into which ideal points and roll call parameters are projected is arbitrary: If two different sessions of Congress are estimated independently, the resulting scores will not be comparable, such that we could expect conservative members in the first set of estimates to located to the right of liberal members in the second. In order to ensure that ideal points are being projected into a stable space, a common reference point is required. Determining what the common reference point should be has generated discussion and sometime disagreement by scholars of ideal points. These can take the form of identical votes, such as on conference committee reports; expressed positions, such as the president's positions on congressional roll calls or senators' positions on Supreme Court decisions; or bridge actors, those persons who have served and voted in multiple institutions.[9] But

7. Keith Poole and Howard Rosenthal, *Congress: A Political-Economic History of Roll-Call Voting* (New York: Oxford University Press, 1997); Keith Poole, *Spatial Maps of Parliamentary Voting* (Cambridge, UK: Cambridge University Press, 2005).

8. Keith Poole and Howard Rosenthal, "D-Nominate after 10 Years: A Comparative Update to Congress: A Political-Economic History of Roll-Call Voting," *Legislative Studies Quarterly* 26 (2001), 8.

9. Michael Bailey uses explicit position taking by presidents and members of Congress on Supreme Court decisions to estimate a common scale across these institutions. Other important shared reference points are survey responses and campaign contributions. Michael A. Bailey, "Comparable Preference Estimates across Time

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

identical votes, where the same proposal is being voted against the same status quo, are few, and finding sufficient instances of clear position taking might be excessively time consuming, especially for scores being estimated across the entire sweep of congressional history. Accordingly, most efforts to establish comparability rely on so-called bridge actors, legislators who served in multiple Congresses or who moved from the House to the Senate.

This reliance, however, comes at a cost. Establishing a common scale through bridge actors requires an assumption of stability in the bridge legislator's preferences. As noted by Poole and Rosenthal, "if we allow a legislator's ideal point to vary freely from one Congress to another, we cannot pin the legislators down to a common space."[10] In DW-NOMINATE a members' mean ideal point is constrained to lie between [−1] and [1] on each of the two estimated dimensions, and a linear trend is estimated for those members who served in three or more Congresses. Each legislator's ideal point at any given moment is a function of their voting record over their entire career, and while the linear terms vary across members, any change is flattened out across the whole of a member's tenure.[11] This is a significant price of making NOMINATE comparable across time.

As an example, consider the transformation in the position of Southern Democrats over the twentieth century. Figure 1 traces the individual ideal point change in DW-NOMINATE (first dimension) for all Southern Democrats who served more than five terms between 1907 and 2009. Overlain are the location of the party means for the Republicans, the Southern Democrats, and the non-Southern Democrats. The shift in the Southern Democrats' position over the course of the twentieth century—represented in the aggregate by the party mean—is largely mirrored in the movements of its longer-serving representatives, despite a considerable amount of individual idiosyncrasy: Most individual Southern Democratic legislators became more conservative in the first half of the twentieth century, and the next generation became more liberal, as we would expect from the qualitative and biographical literature.[12]

When the perspective is an overview of the twentieth century, the constraint on individual preference change is not so important. The level of change in individual ideal points over a member's career is usually fairly small, and the loss in information is compensated by an ability to make cross-chamber or cross-time comparisons. With the assumption of linear change, Poole and Rosenthal were able to present some of the most important findings in contemporary Americanist political science, that the ideological distance between the parties has been starkly increasing since the 1970s, that polarization has occurred more in the House than the Senate, and that it is driven more by a rightward shift among Republicans than a leftward shift among Democrats.

For other sets of questions, however, these assumptions effectively close off lines of inquiry in which ideal points might otherwise be useful, alter the way in which particular periods and historical junctures are understood, and flatten legislators' ideological development across time. Consider, for instance, a question of central important to twentieth-century APD, the timing and causes of the Southern split within the Democratic Party. This is a question for which ideal point measures that track the development of members' preferences across time would be especially useful. And on first glance, DW-NOMINATE seems to provide exactly the sort of measure that would be needed.

In reality, however, the underlying assumptions of DW-NOMINATE make it an inappropriate measure for addressing this question. Precisely because information

and Institutions for the Court, Congress, and Presidency," *American Journal of Political Science* 51 (2007): 433–48; Boris Shor and Nolan McCarty, "The Ideological Mapping of State Legislatures," *American Political Science Review* 105 (2011): 530–51.

10. Poole and Rosenthal, "D-Nominate after 10 Years," 8.

11. An additional problem occurs when there are a large number of members whose voting records do not overlap, resulting in missing data in the agreement score matrix used for generating initial estimates of legislator ideal points. Poole developed a linear mapping technique to address this problem, used to create the Common Space scores that allow for comparisons between the House and Senate. Each session and chamber are estimated independently, and member coordinates are regressed against each other, generating predicted coordinates for the bridge members and session and/or chamber-specific linear transformations that can then be applied to those members who served in fewer than five sessions across chambers. Each of the bridge members is assigned their predicted coordinate from the regression, whereas nonbridge members are given the mean of their session-specific adjustments. A single Common Space score is estimated for the entirety of a legislator's tenure. The model as outlined by Poole is $X_0^k = \left[\Psi_k W_k' + J_n \mu_k'\right]_0 + E_0^k$, where X_0^k is an n by T matrix of legislators, Ψ_k is an n by 1 matrix of legislators coordinates on the k dimension, μ_k is a vector of constants of length T, and E_0^k is an n by T matrix of error terms. The transformation of ideal points for members who do not serve in five chambers is $\hat{\Psi}_{ik} = \frac{\sum_{t=1}^{T_i} \frac{X_{ikt} - \hat{\mu}_k}{\hat{w}_k}}{T_i}$, where T_i is the number of sessions in which legislator i served, X_{ikt} is the ideal point for legislator i in session t on dimension k as generated by the independent scaling, $\hat{\mu}_k$ is the estimated constant and $\hat{w}_k$ the estimated coefficient. Note that this is largely the same as the Groseclose, Levitt, and Snyder technique outlined in their article, except that it computes the mean across the session-specific transformations and imputes as a constant ideal point for all members, and that the regression coefficients are not estimated at the session level. Poole, *Spatial Maps*,

137–39; Tim Groseclose, Steven Levitt, and James Snyder Jr. "Comparing Interest Group Scores across Time and Chambers: Adjusted ADA Scores for the U.S. Congress," *American Political Science Review* 93 (1999): 33–50.

12. See, for example, James M. Glaser, *Race, Campaign Politics, and the Realignment in the South.* (New Haven, CT: Yale University Press, 1997).

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

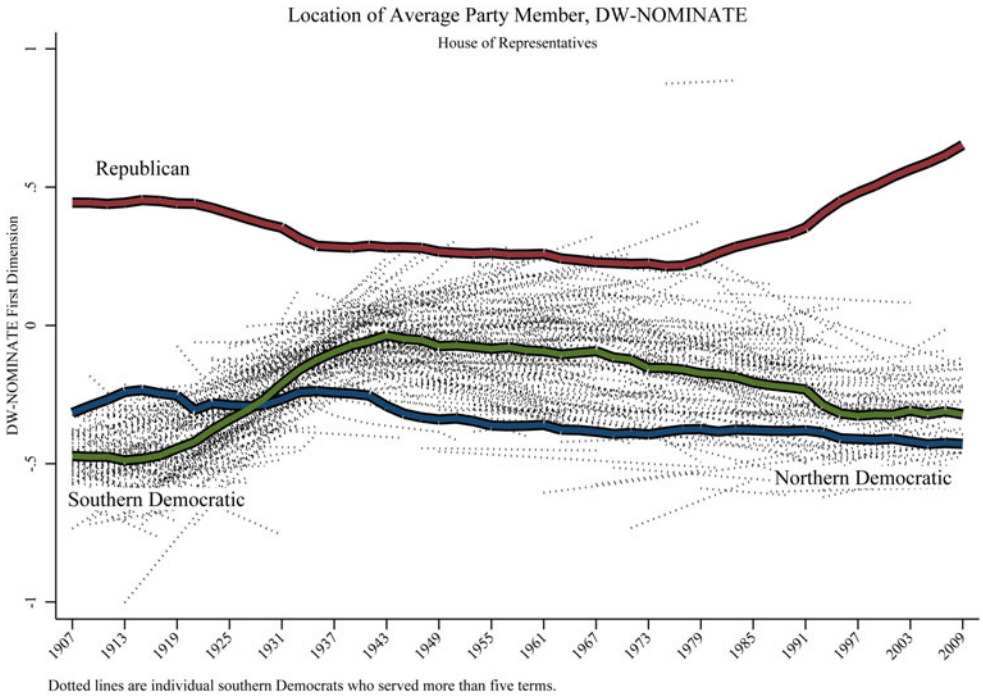


Fig. 1. Party Medians 1907–2009, DW-NOMINATE, First Dimension.

from later in a member's career is used to estimate the member's linear trend, we cannot use DW-NOMINATE to evaluate whether Southern Democrats started moving to the right in the 1920s, as is suggested by Figure 1, or whether this trend was a response to the New Deal. This limitation, however, is more general. Given the linear constraint, there can be no inflection points in legislator change. If a researcher is interested in how legislators may have changed in response to new circumstances, then the use of DW-NOMINATE will likely be inappropriate, as the measure will carry information from after the change into the period before. Over the entire course of American history, such details are probably not as relevant as capturing the essential trends, which DW-NOMINATE does very well. But if the researcher is interested in whether—and which—Southern Democrats began to moderate their positions after passage of the Voting Rights Act, or whether events such as Pearl Harbor, financial crises, or terrorist attacks resulted in distinct patterns of politics in Congress, it is an inappropriate measure and should not be used to probe such questions.

2.2 Ideology and the Substance of the DW-NOMINATE Dimensions

Perhaps a more fundamental obstacle to integrating DW-NOMINATE into APD scholarship stems from its reduction of "ideology" to a coordinate along two dimensions, as well as the corresponding labeling of these dimensions as stable quantities of "liberalism/conservatism" across American history. That is, DW-NOMINATE's ostensible ability to compare all members of Congress on shared dimensions of contemporary significance clashes with an attention to the substance of ideologies and their development over time that has always been a defining feature of the APD subfield.[13]

The spatial model underlying ideal point estimation assumes that policy alternatives can be placed on an ordered line, for example, that the options on defense spending range from zero to some large number, with each member having an ideal point corresponding to where they would prefer policy to be placed. Members are assumed to vote for proposals that move policy closer to their ideal and against those that move it further away.[14] The different

13. Poole and Rosenthal, "D-Nominate after 10 Years," 8; Rogers M. Smith, *Civic Ideals: Conflicting Visions of Citizenship in U.S. History* (New Haven, CT: Yale University Press, 1997); John Gerring, *Party Ideologies in America, 1828–1996* (Cambridge, UK: Cambridge University Press, 1998).

14. Member preferences are assumed to be single-peaked and symmetric: That is, as policy moves away from the point of highest preference—the ideal point—they will be worse off, and they will be indifferent between two options that are equally distant from their ideal. Members are presumed to vote sincerely, or if they do vote strategically, then they are assumed do so in a way that preserves the dimensional ordering. If a subset of a legislature is more likely to engage in strategic voting, while other members vote sincerely, then the placement of these members on the recovered dimension will likely be inaccurate. See Poole and Rosenthal, *Congress*, 17, 147; Howard Rosenthal and Erik Voeten, "Analyzing Roll Calls with Perfect Spatial Voting: France 1946–1958," *American Journal of Political Science* 48 (2004): 620–32; Arthur Spirling and

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

estimation techniques all require a sufficient number and distribution of votes to distinguish between members. That is, they require roll calls to be sufficiently dispersed along the ordered line so as to separate not just, say, hawks from doves but also finer gradations between members.

In reality, it is rare that a single proposed policy change—such as setting the defense budget—is associated with enough votes to cover the range of legislators' preferences. This poses a potential problem: One vote on a defense budget, or even a hundred votes in which the options fall in a narrow range of what the members might prefer, will not provide the needed information to distinguish between members. Fortunately, empirical analyses over the course of several decades have found that many issues share a common dimensionality, so that the line on which preferences on defense spending are arrayed might be largely the same as that on which preferences on health care spending are arrayed. While there might be an insufficient number of votes to distinguish between legislators on defense spending, if a range of issues share a common dimensionality, then the techniques will be able to reliably rank order members.

But given that the scores no longer refer to a specific policy proposal, what does the DW-NOMINATE coordinate actually mean? The answer is surprisingly unclear given the measure's ubiquity. Across their publications, Poole and Rosenthal refer to the two-dimensional space in a number of different and not entirely compatible ways. The first dimension is usually referred to as "ideology," with Poole and Rosenthal noting that while the continuum of positions is an abstraction, it is "convenient to use the word *ideology* as a shorthand code for these positions." This "continuum of ideological positions" can be thought of as ranging from "very liberal to moderate to very conservative" and is part of the "perceived reality" of contemporary America.[15] By affixing these labels, the positions of "very liberal" through "very conservative" are projected across American history through the dimensional structure—a member on the rightmost pole is a conservative, and a member on the leftmost is a liberal.[16]

The substance of the conservative and liberal positions is understood as being organized, "at the risk of some oversimplification," around "conflict over economic redistribution" or "the basic issue of the role of the government in the economy."[17] The second dimension is interpreted as capturing issue positions that cut across this primary ideological dimension. These issues are usually, but not always, associated with race: The second dimension is most salient "in periods when race issues are distinct from economic ones."[18] But the content of the second dimension has also been associated with bimetallism, internal improvements, or the management of public lands—all of which also involve questions about the basic role of the government in the economy and economic redistribution—as well as social issues such as abortion.[19] And the first dimension is also treated as capturing party loyalty "ranging from strong loyalty to one party… to weak loyalty to either party and to strong loyalty to the second, opposing party," although at other times party loyalty is said to be captured by the second dimension.[20] The confusing and residual nature of what the second dimension of NOMINATE actually captures is a central reason, in our opinion, why this dimension has been virtually ignored outside of Poole and Rosenthal's own work.

It is important to note that, labeling to the contrary, the conceptualization of ideology in DW-NOMINATE has no intrinsic or stable connection with policy positions. Poole and Rosenthal, like much APD work, draw on Philip Converse's definition of belief systems as "a configuration of ideas and attitudes in which the elements are bound together by some form of constraint or functional interdependence."[21] But where Converse maintains the emphasis on the

Iain McLean, "UK OC OK? Interpreting Optimal Classification Scores in the U.K. House of Commons," *Political Analysis* 15 (2007): 85–96.

15. Poole and Rosenthal, *Congress*, 4.

16. Since the foundational work in this area, the first dimension has also been interpreted as "partisanship" and the second dimension as "ideology." Technically, given their definition of ideology (discussed below), both the first and second dimension should be considered "ideology," but this term is usually reserved for the first dimension. Most discussions of DW-NOMINATE, including the work of Poole and Rosenthal and their coauthors, interpret the first dimension as an ideological dimension related to, but separate, from party. For example, Cheryl Schonhardt-Bailey notes that "the first dimension reflects a liberal-conservative divide over the role of the government in the economy" and that "it is this first dimension that Poole and Rosenthal summarize as 'ideology.'" McCarty, Poole, and Rosenthal describe the first dimension as "correspond[ing] to the popular conception of liberals versus conservatives." Cheryl Schonhardt-Bailey, *From the Corn Laws to Free Trade: Interests, Ideas, and Institutions in Historical Perspective* (Cambridge, MA: MIT Press, 2006), 359, note 6; Nolan McCarty, Keith Poole, and Howard Rosenthal, *Polarized America: The Dance of Ideology and Unequal Riches* (Cambridge, MA: MIT Press, 2006), 26.

17. Poole and Rosenthal, *Congress*, 6, 35. As they write on the www.voteview.com site, "The primary dimension is the basic issue of the role of the government in the economy, in modern terms liberal-moderate-conservative" (http://voteview.com/political_polarization2015.html).

18. Poole and Rosenthal, *Congress*, 51.

19. Ibid., 48; McCarty et al., *Polarized America*, 50.

20. Poole and Rosenthal, *Congress*, 46. The second dimension's capturing party loyalty is claimed to be partly responsible for the finding that "a slightly better accounting of roll call votes is gained by using two dimensions, even in periods when the race issue is largely inactive" (Poole and Rosenthal, *Congress*, 5–6).

21. Philip Converse, "The Nature of Belief Systems in Mass Publics," in *Ideology and Discontent*, ed. David Apter (New York: The Free Press, 1964), 207; John Gerring, "Ideology: A Definitional Analysis," *Political Research Quarterly* 50 (1997): 957–94, 980; Gerring, *Party Ideologies in America*; Sam DeCanio, *Democracy and*

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

"ideas and attitudes," Poole and Rosenthal define ideology as the stable continuum that enables a prediction of who will vote with whom—it is "fundamentally the knowledge of what-goes-with-what"—rather than a set of ideas or issue positions that are characterized by their functional interdependence.[22] In most contexts, when we say "member of Congress X is more conservative than member Y," we mean something equivalent to "the issue positions supported by member X are more aligned with the policy prescriptions of a relatively stable body of political thought, or of a relatively stable political movement, than those of member Y." That is, we assess a members' conservatism based on the substance of the positions that they take. Not so in DW-NOMINATE, and Poole and Rosenthal explicitly make clear that issue positions cannot be mapped on to this "ideological" structure.[23]

But the use of the labels "liberal" and "conservative"—the common meaning of which *do* entail issue positions—to characterize this structure creates a danger of misinterpretation, especially for periods in which the liberalism or conservativism might be less relevant axes of political alignment. As the content of the issues, ideologies, and partisanship separating members of Congress changes over time, the relationship between ideal points and identifiable liberal or conservative policy positions becomes tenuous. Representative Samuel A. Witherspoon (D-MS), first elected to the 62nd Congress, believed that states had the absolute right to "control all domestic questions and conditions, such as labor, education, domestic relations, preservation of order, good morals, encouragement of industry," and the only duty of the federal government was "to perpetuate an indissoluble union."[24] He was not more liberal than Dennis Kucinich or only slightly less liberal than Adam Clayton Powell Jr. in any sense in which the term is used in analyzing American politics, despite his comparable first dimensions DW-NOMINATE score.[25] Likewise, the claim that Helms is more conservative than Taft Sr. is not a reflection of the fact that Helms would likely support a free trade bill, but Taft, an opponent of free trade, would have been in opposition.[26] We cannot say Helms is more conservative if conservative implies a set of issue positions. All we can say is that assuming stability in the space into which the ideal points are projected, Helms is further from the center than Taft on the dimension that most closely resembles a liberal–conservative divide for the contemporary period.

We are not saying that Poole and Rosenthal were unconcerned with issue substance. They certainly were, treating the low-dimensional model as only a beginning point for a closer analysis of issue substance. By examining the angle of the cutline, the hyperplane that runs equidistant between the policy proposal and status quo and that theoretically should divide supporters from opponents, or a similar metric such as the proportional reduction in error that comes from the addition of a second dimension, Poole and Rosenthal across most of their published work give careful attention to how issues map on to the low-dimensional structure. It was on this basis that they affixed issue labels to the estimated dimensions. But the shorthand interpretation of the dimensions as issue positions can result in misleading interpretations of legislators' policy preferences if the specific content of the dimensions are not closely examined

the Origins of the American Regulatory State (New Haven, CT: Yale University Press, 2015), 35.

22. The idea of a constraint is of central importance to understanding ideology in public opinion research, but here too it remains closely tied to substantive issue positions. Converse, for example, defines "mass belief systems" in terms of a constraint, but the substance of the "specific belief elements" remains the key indicator of liberalism or conservatism. In recapitulating Converse's definition, John Zaller notes that "people who are liberal (or conservative) on one issue tend to be relatively liberal (or conservative) on a range of other issues." The conservation and liberal issue positions are affixed separately from the fact that they cluster together. In fact, the definition of ideology offered in *Congress* is perhaps best seen as an instance of an operationalization of a concept for a particular domain—public opinion research—becoming the definition of the concept. Keith Poole, "Changing Minds? Not in Congress!" *Public Choice* 131 (2007): 435–36; Poole and Rosenthal, *Congress*, 4. See also Hans Noel, "The Coalition Merchants: The Ideological Roots of the Civil Rights Realignment," *The Journal of Politics* 74(2012): 156–73; Converse, "Nature of Belief Systems," 208, 209; John Zaller, *The Nature and Origins of Mass Opinion* (Cambridge, UK: Cambridge University Press, 1992), 113.

23. "The simple ideological structure does not lead to a predictive model for specific issues. True, in the short term one can predict with accuracy. . .. But to obtain medium- and long-term forecasts, one would need to model how issues map onto the structure" (Poole and Rosenthal, *Congress*, 5). Although it is not always clear that it is "ideology" that is being measured, for short periods or for a single Congress it is not difficult to leverage other sources of information to discern what perhaps is being reflected in the first and second dimension scores. But to do this systematically across history is a much more difficult undertaking. More importantly, while labeling the dimensional structure "ideology," arrayed from "liberalism" to "conservatism," might make sense in some periods, it will be highly anachronistic and misleading in others.

24. "Address of Mr. Stephens, of Mississippi," *Samuel A. Witherspoon, Late a Representative from Mississippi, Memorial Addresses, delivered in the House of Representatives and the Senate of the United States* (Washington, DC: Joint Committee on Printing), 74.

25. Rather, the content of the Democratic Party's policy commitments changed dramatically over the twentieth century. This does not mean ideal point estimates and their cross-time comparisons are meaningless. Witherspoon was on the radical side of the Democratic Party, as it was constituted at the time and relative to the Republican opposition. "Address of Mr. Smith, of South Carolina," *Samuel A. Witherspoon, Late a Representative from Mississippi, Memorial Addresses, delivered in the House of Representatives and the Senate of the United States* (Washington, DC: Joint Committee on Printing),106; Gerring, *Party Ideologies in America*.

26. Indeed as Karol notes, opposition to trade liberalization was a position associated with conservatives in the 1950s, but had become a position associated with liberals by the 1980s. David Karol, *Party Position Change in American Politics: Coalition Management* (Cambridge, UK: Cambridge University Press, 2009), 44. Clarence E. Wunderlin, *Robert A. Taft: Ideas, Tradition, and Party in U.S. Foreign Policy* (Oxford: Rowman & Littlefield, 2005), 38.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

to see whether, at that particular historical juncture, the dimensions actually capture what they are claimed to be measuring.[27]

Consider the interpretation of the second dimension as capturing race, probably the most common label. Figure 2 shows the party medians on the DW-NOMINATE second dimension for the Senate and the House, from the end of Reconstruction to the present.[28] In the mid-twentieth century, Southern conservatism on civil rights likely informed the second dimension, and yet the scores can still be misleading. For instance, the measure bears little relationship to patterns of voting on civil rights in this period, other than the increasing isolation of the South. While the Republican Party in 1965 was probably more to the left on civil rights than it would be twenty years later, it is difficult to reconcile the historical record with the Republicans being to the left of Northern Democrats at this time, let alone twenty or even thirty years later.[29]

We are not taking issue with the decision of Poole and Rosenthal to estimate scores based on all votes or with their efforts to empirically assess what might be the content of these dimensions across time.[30] But the inconsistent association of the estimated dimensions with substantive issue positions limits the utility of these scores for specific questions that arise in the course of much APD work.

One of the main ways that APD has engaged with the study of Congress is by examining the dynamics surrounding the passage or defeat of specific bills, where we are often interested in identifying the pivotal members whose votes were needed for passage. For instance, an important area of study in APD has examined the construction of the "new American state" during the progressive era, with scholars debating which set of lawmakers propelled reform and which were well-situated within Congress to arbitrate between different reformist projects.[31] In these cases we are less interested in who the pivotal members are according to an ideological score estimated across all roll calls than in whether a given set of legislators was likely to be pivotal given their specific preferences on the issue in question at a specific time. In a similar vein, many APD scholars have traced the patterns of policymaking within well-defined issue areas. While DW-NOMINATE allows for a highly effective representation of polarization, it is less useful for scholars who want to know how politics has developed in a given policy area.

2.3 Policy Development, Historical Context, and Agnostic Ideal Points

The third issue we consider is DW-NOMINATE's agnosticism toward sources of information other than

27. Content is generally ascribed by seeing what issue areas seem to be more strongly associated with a given dimension at a given time. More precisely, the proportional reduction in error (PRE) achieved by a one- versus a two-dimensional model is calculated for each vote, and the issue topics for which the PRE increased by 0.2 from adding a second dimension are noted. The "substance" of the second dimension is whatever set of issues most commonly increased the PRE by the requisite threshold during periods in which the increase in the aggregate PRE was most substantial. The aggregate proportional reduction in error is frequently used in analyses of ideal point measures for assessing the improvement in fitting the data by different specifications of a model relative to some benchmark. The benchmark model used in these discussions is the minority vote. A unique PRE is calculated for each vote, and these are aggregated by summing across all votes. $APRE = \frac{\sum_{j=1}^{n} \left(Minority\ Vote - Classification\ Errors\right)_j}{\sum_{j=1}^{n} Minority\ Vote_j}$. Poole and Rosenthal, *Congress*, 30.

28. From 1865 until the election of 1896 (the 55th Congress), the second dimension is treated as capturing conflict over bimetallism and the currency. From the turn of the century to the 1940s, there is no consistent pattern, and it becomes a "civil rights" dimension only in the postwar period, and only really for the Senate. See Table 3.2 of *Congress* for the content of the second dimension in the House. Whereas civil rights does meet the threshold set by Poole and Rosenthal, it does so only in the 89th House. In the Senate, civil rights meets the threshold for most Congresses from the 81st on. More generally, for any estimated dimension, the appropriate interpretation will vary according to specific institutional features and the political context. Parliamentary legislatures, for example, consistently show voting organized along a government-opposition axis, in which the opposition votes en bloc against government proposals even when these move policies toward their preferred outcome. In these contexts, the second dimension can often be interpreted as capturing whatever left or right division is not covered by the government or opposition distinction. But not always, and in many contexts, the second dimension might be a geographical, linguistic, religious, or racial divide, and the left and right divide will not be directly captured at all. See Poole and Rosenthal, *Congress*, 48–51; Torun Dewan and Arthur Spirling, "Strategic Opposition and Government Cohesion in Westminster Democracies," *American Political Science Review* 105 (2011): 337–58; Simon Hix and Abdul Noury, "Government-Opposition of Left-Right? The Institutional Determinants of Voting in Legislatures" (paper presented at the annual meeting of the American Political Science Association, Chicago, 2013); Jean-François Godbout and Bjorn Hoyland, "Legislative Voting in the Canadian Parliament," *Canadian Journal of Political Science* 44 (2011): 367–88.

29. One possible response would be that these scores always need to be mapped into two-dimensional space. We agree, although we note that this is not usually done and applies equally to the first as to the second dimension. Nor does it result in scores that are more readily interpretable. For the relative timing of the non-Southern Democratic Party's move to the left of the Republicans on civil rights, see Brian D. Feinstein and Eric Schickler, "Platforms and Partners: The Civil Rights Realignment Reconsidered," *Studies in American Political Development* 22 (2008): 1–31; Jeffery A. Jenkins, Justin Peck, and Vesla M. Weaver, "Between Reconstructions: Congressional Action on Civil Rights, 1891–1940," *Studies in American Political Development* 24 (2010): 57–89.

30. It is Poole and Rosenthal, after all, who provide the many labels for the second dimension, precisely because they are sensitive to the fact that the issue content of this dimension, in particular, is unstable.

31. Elizabeth Sanders' *Roots of Reform*, for example, identifies a set of agrarian lawmakers who persistently advocated for progressive reforms in Congress, but who repeatedly were forced to make concessions to a pivotal bloc of lawmakers, generally from the Midwest, who were willing to accept the administrative discretion supported by northeastern conservatives.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

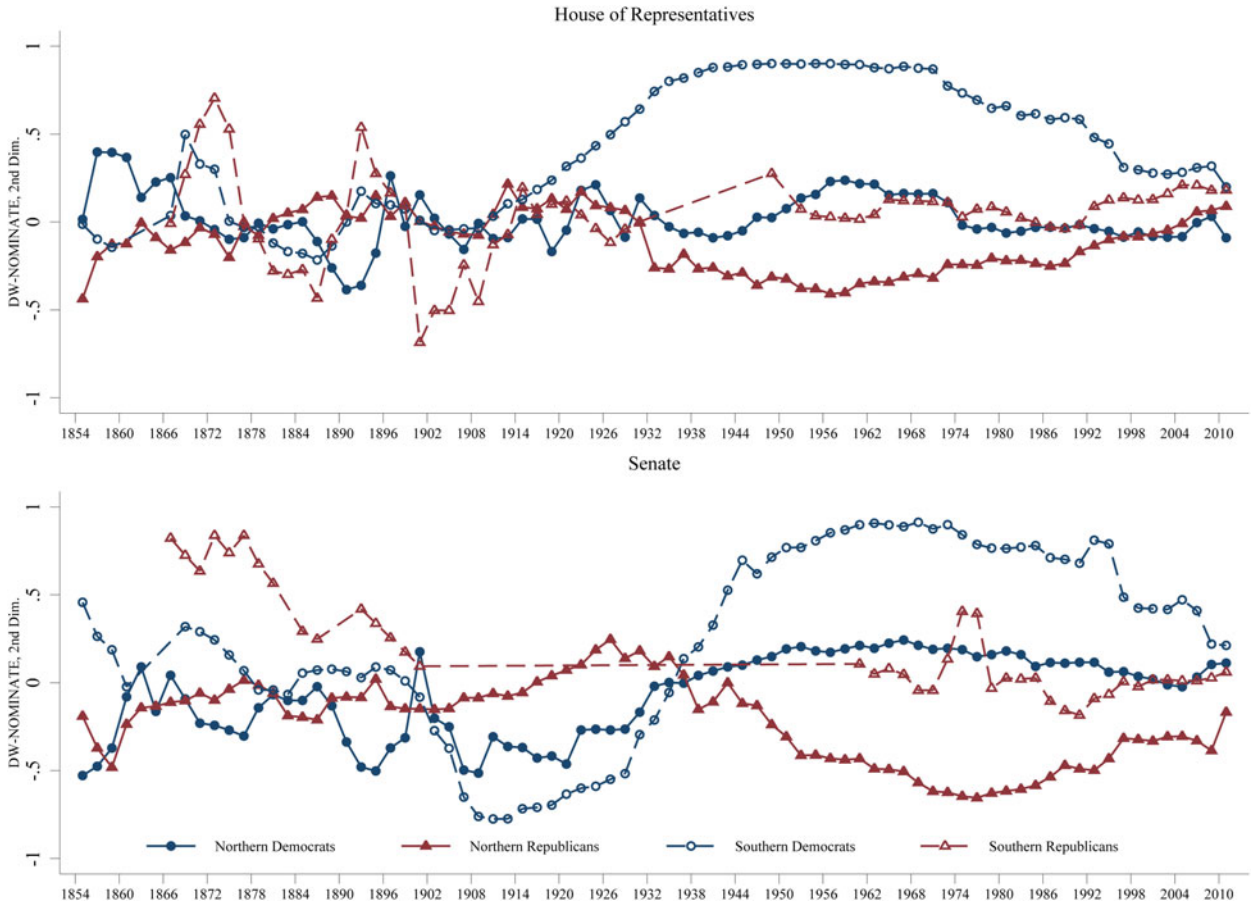


Fig. 2. Party and Regional Medians 1854–2010, DW-NOMINATE, Second Dimension.

roll calls. The implicit assumptions this agnosticism imposes make it less likely that the DW-NOMINATE results are comparable across historical periods and reliably array members according to their individual preferences. The only thing that is identified from the roll call matrix is an ordinal ranking of members along a set number of dimensions. To establish a measure of distance, ideal point procedures need to make additional assumptions, namely, that errors in how members vote—individual voting decisions that are incorrectly predicted from the rank ordering—are more likely when the midpoints between the status quo and the policy proposal are close to the members' preferred location, and that these errors are independently and identically distributed across members.[32] The promise of almost all ideal point techniques is that by leveraging these errors and by imposing some degree of constraint on legislators' movements, ideal points and roll calls can be arrayed on a common dimension or dimensions across time.

But the agnosticism of most techniques to the content of the agenda, to the policies at issue, or to the institutional or historical context in which the votes are being cast limits our ability to make good

32. The different techniques rely on different assumptions about the shape of members' utility curves and about the distribution of error. The differences are not consequential when the policy alternatives lie in the neighborhood of the legislator's ideal point, but they do differ when these alternatives are located far from the legislator's preferred location. As this is more common for extremists, the differences in the utility functions are most consequential at the extremes. The roll call parameters generated by NOMINATE are the midpoints for each dimension as well as the spread, the distance between the location of a yea vote versus a nay vote divided by 2. The status quo and proposal locations can be calculated from the midpoint and spread, but Poole and Rosenthal stress that these values (unlike the midpoint) are poorly estimated. Poole and Rosenthal, *Congress*, 235; David A. Armstrong, II, Ryan Bakker, Royce Carroll, Christopher Hare, Keith T. Poole, and Howard Rosenthal, *Analyzing Spatial Models of Choice and Judgment with R* (Boca Raton, FL: CRC Press, 2014), 223–24; Joshua Clinton, Simon Jackman, and Douglas Rivers, "The Statistical Analysis of Roll Call Data," *American Political Review* 98 (2004): 356; See Royce Carroll, Jeffrey B. Lewis, James Lo, Keith T. Poole, and Howard Rosenthal, "The Structure of Utility in Spatial Models of Voting," *American Journal of Political Science* 57 (2013): 1011; Keith Poole, "A Non-Parametric Unfolding of Binary Choice Data," *Political Analysis* 8 (2000): 211–37.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

on this promise. Ideal points reflect both the distribution of preferences and the issues that are brought to a vote, and changes in either one can result in changed estimates of legislator positions. Changes in agenda control, events, or policy development might result in a very different set of issues, with different cutlines and outcome locations being voted on across time.[33] But even if we hold members' preferences fixed, change in the substantive agenda can result in changes to the estimated ideal points. By simply adjusting the nature of the bills being voted upon, we can increase the amount of estimated polarization—the distance between party ideal points—even if preferences are held constant.

Similarly, systematic shifts in the location of policy and legislators' preferences will not be reflected in cross-time roll call parameters or ideal points. No ideal point technique can independently account for how the issue space might evolve over time, or how issues and legislators' positions on these might develop in tandem. If it were true that members had fixed or highly constrained preferences arrayed on a stable ideological dimension over the course of their careers, then identifying and controlling for change in the issue agenda would be relatively straightforward. Conversely, if it were true that the issue agenda did not move in any systematic direction over time, that members were voting on the same proposals and status quo points in 1968 as in 1933, it would be easy to determine who had become more conservative and by how much. But if *both* member preferences and the issue agenda are systematically moving over time—as seems likely for many issues—then we have no fixed location to enable a comparison. Such systematic shifts in policy preferences are a core feature of influential interpretations of APD.[34]

Again, this feature of most ideal point techniques can make such measures less useful for answering certain types of questions. Have the parties become more ideologically polarized on a given set of issues, in the sense that the distance between their preferred policies has increased; or have they become more partisan in their voting while still holding relatively stable and proximate policy positions? This is a question that is best addressed at the level of specific policy areas, but it is also one for which current ideal point techniques are likely to be inappropriate, as these do not anchor the policy space in such a way as to allow us to distinguish between polarization and partisanship. APD scholars tend to emphasize the degree to which process and sequencing matters, to examine the ways in which politics occurs in time. If issues and policies develop over time such that the number of votes that divide the parties internally declines over time—either by moving primary attention over these issues to an independent agency, such as occurred with the currency and a range of other once highly contested issues—then the mix of roll calls might come to be dominated by those issues that divide the parties from each other. In this case, our ability to compare, in a straightforward and easily interpretable manner, ideal points from across different periods will be significantly diminished.[35]

***

These are issues that in many ways fall along the methodological fault lines that define the APD subfield as a distinct endeavor. And they reflect particular decisions made by Poole and Rosenthal as to how to best analyze an enormous quantity of data. They were not the only decisions that could have been made, but they were by no means the wrong decisions. Our purpose in this article again is not to take issue these choices, which have been detailed, defended, and validated across decades of research. Rather, we want to consider how different choices, informed by APD scholarship, might allow for measures that are better suited for certain lines of historical inquiry. Indeed, we see this exercise in many ways as an extension of the work of Poole and Rosenthal. In the next section, we introduce new measures and consider how they might be enable us to address questions of central interest to APD.

33. Frances E. Lee, "Agreeing to Disagree: Agenda Control and Senate Partisanship, 1981–2004," *Legislative Studies Quarterly* 33 (2011): 199–222; Clifford Carrubba, Matthew Gabel, and Simon Hug, "Legislative Voting Behavior, Seen and Unseen: A Theory of Roll-call Vote Selection," *Legislative Studies Quarterly* 33 (2008): 543–72. The logic of the recovered ideal points being sensitive to the location of the cutlines is effectively the same as that laid out in Keith Krehbiel, "Party Discipline and Measures of Partisanship," *American Journal of Political Science* 44 (2000): 212–27. Although ideal points are an improvement over these other scores, in that they estimate the location of the cutlines, without integrating policy substance, there is no way to know whether the cutlines are comparable over time.

34. Realignment interpretations, for example, argue that extended stretches of American history are characterized by the relative ascendancy of a particular party, ideology, and set of issue priorities. Both the issues under consideration and the political implications of the roll calls being voted on are likely to change considerably across different periods in American history. Because the underlying roll call matrix does not have any information about the location of the policy being considered, but only observed decisions on discrete votes, changes in the process by which roll calls are generated or in the likelihood of voting error can lead to substantial changes in the estimated ideal points. And there is good reason to believe that with differences in agenda control, both across time and institutions, that there will be periods that differ starkly in the mix of roll calls that come to the floor. Jason Roberts and Steven Smith, "Procedural Contexts, Party Strategy, and Conditional Party Voting in the U.S. House of Representatives, 1971–2000," *American Journal of Political Science* 47 (2003): 305–17; Joshua D. Clinton and John Lapinski, "Laws and Roll Calls in the U.S. Congress, 1891–1994," *Legislative Studies Quarterly* 33 (2008): 511–41.

35. In a similar vein, most existing techniques treat all roll calls as occurring simultaneously, and within a legislative session, the ordering of roll calls can be rearranged without any change in the estimates.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

3. INTEGRATING IDEAL POINTS AND HISTORICAL ANALYSES

3.1 Nonlinear Movement in Legislator Positions

There are a variety of alternatives to the linearity constraint in DW-NOMINATE, each with their own strengths and weaknesses, most of which impose constraints on the movement of individual legislators.[36] An alternative way to proceed is to impose constraints not on individuals but on the movement of members in the aggregate. Tim Groseclose, Steven Levitt, and James Snyder (GLS) propose a linear mapping approach that adjusts interest group scores so that these can be compared across institutions and across time.[37] This approach takes scores generated for specific chambers—such as the yearly interest group ratings created by Americans for Democratic Action—then places these on a common scale by adjusting the initial scores to account for the differences in the scale relative to a given baseline, using the formula $\hat{y}_{it} = \frac{y_{it} - \alpha_t}{\beta_t}$, where $\hat{y}_{it}$ is the "adjusted" score for member i at time t, y_{it} is the initial nominal score, and α_t and β_t are the session-specific shift and stretch parameters.[38]

While the method was developed with interest scores in mind, it can also be used to place static ideal points—such as those generated in W-NOMINATE or IDEAL—on a comparable scale across time and institutions. We believe the payoff for using this technique on the individual level ideal points is potentially far greater. Using the GLS approach, we have placed on a common scale Congress-specific scores generated in W-NOMINATE and in IDEAL for the House and the Senate, from 1877 to 2011.[39]

In Section 2.1 we highlighted some questions where ideal points might be useful to APD scholars but for which DW-NOMINATE would be inappropriate. How does this measure address the concerns raised above? For one, it allows for the sort of inflection points in individual careers that we believe are often important. This recovers important nuances regarding individual change as well as allows for ideal points to be integrated into research questions where punctuated change is an important feature. While the broad arc of most legislators' careers might indeed be linear, there is little reason to believe that this change occurs equally across a legislator's career. Even if legislators do not do ideological backflips, a linear trend denies the possibility that they change at specific moments or in response to particular experiences or events.[40]

Consider the career of Jamie Whitten (D-MS), the second longest serving representative in American history. Whitten was elected in 1941, temporally adjacent to the birth of the conservative coalition between Southern Democrats and Republicans. But by the end of his career, he had moved gradually to the left, and by the 1980s was generally supportive of liberal policies and an opponent of the Reagan administration. As the *New York Times* remarked in an obituary otherwise devoted to detailing the government spending he brought to Mississippi,

> befitting his origins, Mr. Whitten ran as a conservative Democrat and opposed desegregation. . . . At first he did not support food stamps, a popular program among most House Democrats. But, mindful of the increasing number of black voters in his district, he joined the majority on that issue in 1975. . . . Gradually his ratings rose among organized labor and other cataloguers of liberal voting records, and when it came time for the Democratic caucus to elect a new chairman of the Appropriations Committee in 1978 he won, 157 to 88, in a secret ballot.[41]

36. The main exception to the practice of constraining individual members' movements are the one-Congress-at-a-time NOMINATE scores. These are generated by first estimating ideal points and cutlines in the constant model—with no linear change—and then by re-estimating Congress-specific ideal points while holding the roll call parameters fixed. These estimates are rarely used, although we believe they merit greater attention. The DW-NOMINATE framework itself allows for nonlinear scores: Legislators' scores across time are modeled as a polynomial function of time, so that legislator i's ideal point on dimension k at time t is $X_{ikt} = \gamma_{ik0} + \gamma_{ik1}T_{t1} + \gamma_{ik2}T_{t2} + \ldots + \gamma_{ikv}T_{tv}$, where v is the degree of the polynomial, γ are the coefficients of the polynomial, and the Ts are Legendre polynomials. For DW-NOMINATE, v is set at 1—estimating a linear score—because Poole and Rosenthal found that "essentially all movement is captured by simple linear movement," and that "the linear model in two dimensions was the best combination of explanatory power and number of parameters." The modeling decisions made by Poole and Rosenthal were carefully thought through and have been persuasively defended as justified relative to the potential gains of a parsimonious measure that can be compared across the entirety of American history. Poole, *Spatial Models*, 104–107; Poole and Rosenthal, *Congress*, 25, 236; Timothy Nokken and Keith Poole, "Congressional Party Defection in American History," *Legislative Studies Quarterly* 29 (2004): 545–68.

37. Groseclose et al., "Comparing Interest Group Scores."

38. Specifically, the method estimates a latent dimension, $y_{it} = \alpha_t + \beta_t x_i + \varepsilon_{it}$, with x_i being a mean-preference parameter—initially the mean score for a member over that member's entire career—and ε_{it} being an error term capturing individual change.

39. Given that these two measures are highly correlated in one dimension, the following discussion focuses on W-NOMINATE. The scores discussed in Sections 3.2 and 3.3, by contrast, were estimated with IDEAL. As with the Common Space scores, GLS-adjustments relies on bridge actors who served in both the House and the Senate, with the distinct advantage that rather than a static ideal point, members have scores that reflect the independently estimated sessions that are being adjusted.

40. Barry Burden, for example, demonstrates that individual events in legislators' lives can dramatically alter their preferences. Barry Burden, *Personal Roots of Representation* (Princeton, NJ: Princeton University Press, 2007).

41. David Binder, "Jamie Whitten, Who Served 53 Years in House, Dies at 85," *New York Times*, Sept. 10, 1995,

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

The two panels of Figure 3 shows Whitten's ideal point from his election in 1941 to his retirement more than fifty years later, with the location of the mean Southern Democrat shown for comparison. The basic narrative of his transformation from a conservative—against whom 109 Democrats and thirty-three Republicans voted against seating in 1965, in favor of Fannie Lou Hamer and the Mississippi Freedom Democratic Party—to a mainstream if not liberal Democrat—is evident in the DW-NOMINATE trend for his career (left panel of Figure 3). When we use the GLS technique to adjust static W-NOMINATE scores, a very different picture emerges, in which Whitten's progress maps on to the trajectory of his fellow Southerners.

Nor is Whitten exceptional in revealing the nuances of legislator's position change over time. Figure 4 compares the DW-NOMINATE to GLS-adjusted W-NOMINATE scores for nine long-serving members of the House of Representatives. Each of these members shows considerable change over the course of their career under the adjusted estimates. Even for those such as Carl Vinson (D-GA-8)—the first person to serve fifty years in the House—whose trend largely accords with the linear trend of DW-NOMINATE, there are important differences in the timing of their change: For Vinson, there is a clear move right near the beginning of his career, and then another during the New Deal and WWII. As we have seen, the arc of Representative Whitten's long career, which mirrored that of the Southern Democrats, will be misinterpreted using DW-NOMINATE, while the sharp move by Emanuel Cellar to the left (D-NY-10) during the New Deal is lost.

The greater flexibility in member movement aggregates up to important differences in how particular periods are represented. Across various periods, DW-NOMINATE and GLS-adjusted W-NOMINATE scores return considerably different representations of the structure of legislative voting, changing how we measure and thus how we explain polarization and party change over time. Consider the question we raised above, as to the timing of the Southern Democrats' break with the national party in Congress. The top panel of Figure 5 traces the mean location for the three most relevant factions in the House of Representatives, the Republican Party, the Northern Democrats, and the Southern Democrats, using both DW-NOMINATE and GLS-adjusted W-NOMINATE. The bottom panel traces the different measures' accounts of depolarization and repolarization over the course of the twentieth century.

The DW-NOMINATE scores show the Southern Democrats beginning to drift rightward during the 1920s, but we have already seen that these scores are problematic because information from later in Southern Democrats' careers will bleed into earlier period. With GLS-adjusted W-NOMINATE scores, by contrast, there is a slight drift toward the center before the New Deal, followed by a sharp move after the elections of 1942. These elections also seem to have occasioned a sharp move to the left by Northern Democrats, reflecting the defeat of many of its more conservative members. The GLS-adjusted W-NOMINATE scores map much more closely to the existing qualitative and quantitative analyses of the period, which finds the 78th Congress (1943–44) to have been the key inflection in time.[42]

This, of course, does not alone make it a measure better suited to this particular research question. Both the DW-NOMINATE and GLS-adjusted scores are abstracted representations of a much more complex reality, and we should not expect all the details of legislative conflict and party development will be reflected. But the fact that the scores and rank ordering of members do not include information from earlier or later Congresses makes it a methodologically more appropriate measure than DW-NOMINATE for certain questions. Indeed, for this particular question, as well as others where the researcher is concerned with identifying moments of inflection in individual or aggregate preferences, GLS-adjusted Congress-level scores will likely be better suited.

One advantage of GLS as a means of comparing ideal points across time is that, because it involves adjusting Congress-specific scores rather than estimating scores simultaneously across time, it does not change the rank ordering of members and therefore maintains the information from the chamber-specific estimates. The shift and stretch parameters estimated for a given Congress are applied to all members equally, and their ordinal rankings and relative distances are therefore unaffected. The left-hand panel of Figure 6 shows a scatterplot of unadjusted and adjusted W-NOMINATE scores for the 45th through 111th House of Representatives, as well as the 45-degree line. The relationship for each Congress is perfectly linear—the ordinal rankings and relative location of members have been maintained—but most have been visibly shifted toward the right and (to a lesser degree) stretched so that the adjusted scores are distributed in a wider or narrower range. If the chamber-specific estimate identifies you as the most conservative Republican or most conservative Democrat, you will not lose that distinction once the scores are adjusted. This is not true of DW-NOMINATE, which does not rescale W-NOMINATE but rather estimates ideal points using information

http://www.nytimes.com/1995/09/10/obituaries/jamie-whitten-who-served-53-years-in-house-dies-at-85.html.

42. Ira Katznelson and Quinn Mulroy, "Was the South Pivotal? Situated Partisanship and Policy Coalitions during the New Deal and Fair Deal," *Journal of Politics* 74 (2012): 604–20.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

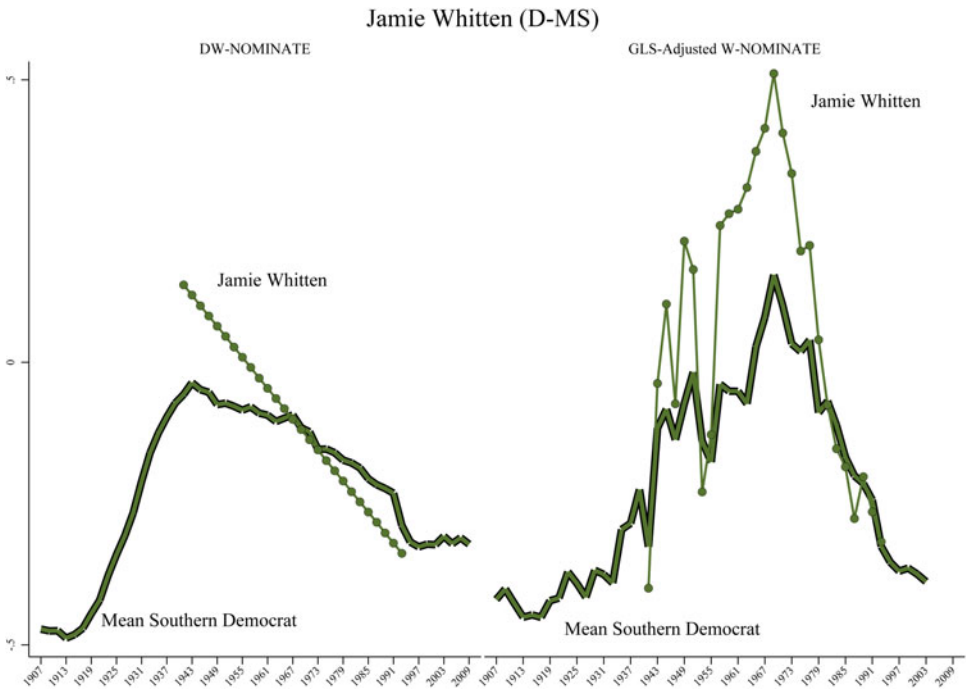


Fig. 3. Preference Development of Jamie Whitten (D-MS), DW-NOMINATE and GLS-Adjusted W-NOMINATE.

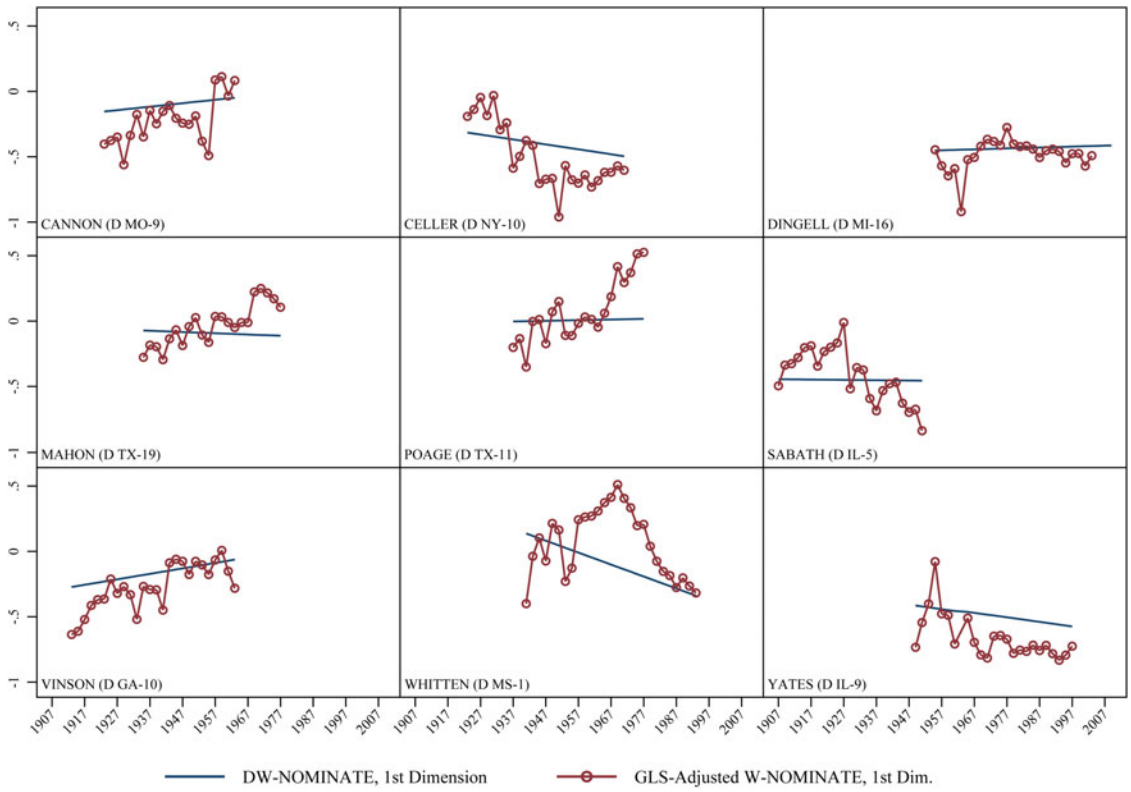


Fig. 4. Preference Development of Long-Serving Members, DW-NOMINATE and GLS-Adjusted W-NOMINATE.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

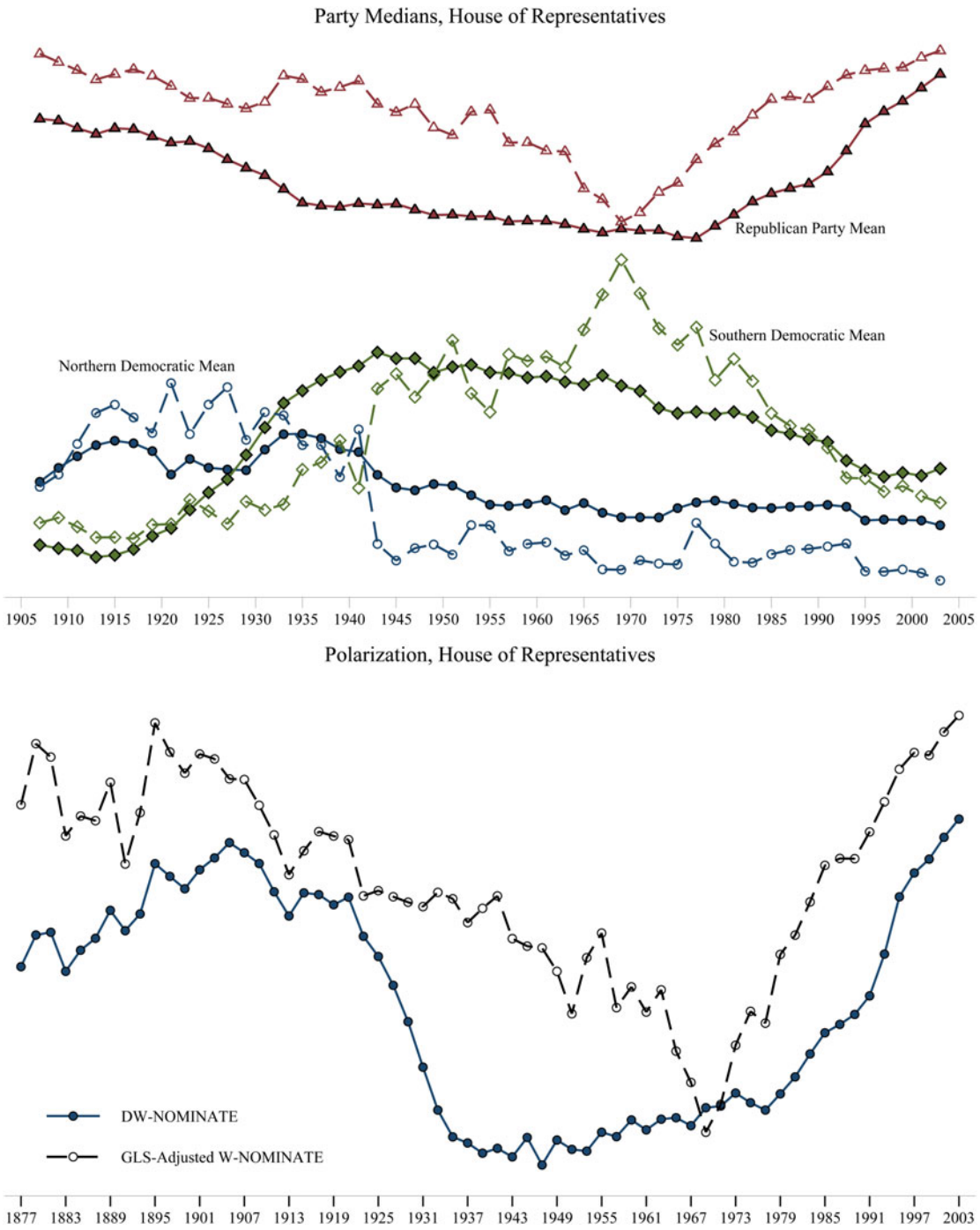


Fig. 5. Party Medians and Polarization in the House 1905–2005, DW-NOMINATE and GLS-Adjusted W-NOMINATE.

from the entirety of a legislator's career (right panel of Figure 6). As a result, there are considerable differences in the rank ordering and relative location of members between DW-NOMINATE and the Congress-specific W-NOMINATE scores. We think this is an extremely important point for APD scholars to understand.

The ordinal rankings and relative locations of members in any given Congress are not necessarily the most interesting or useful information, but they should also not be discarded too quickly. For one, cross-time comparability is not always essential for different research questions. We are often looking only at one Congress, or at how legislative dynamics

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

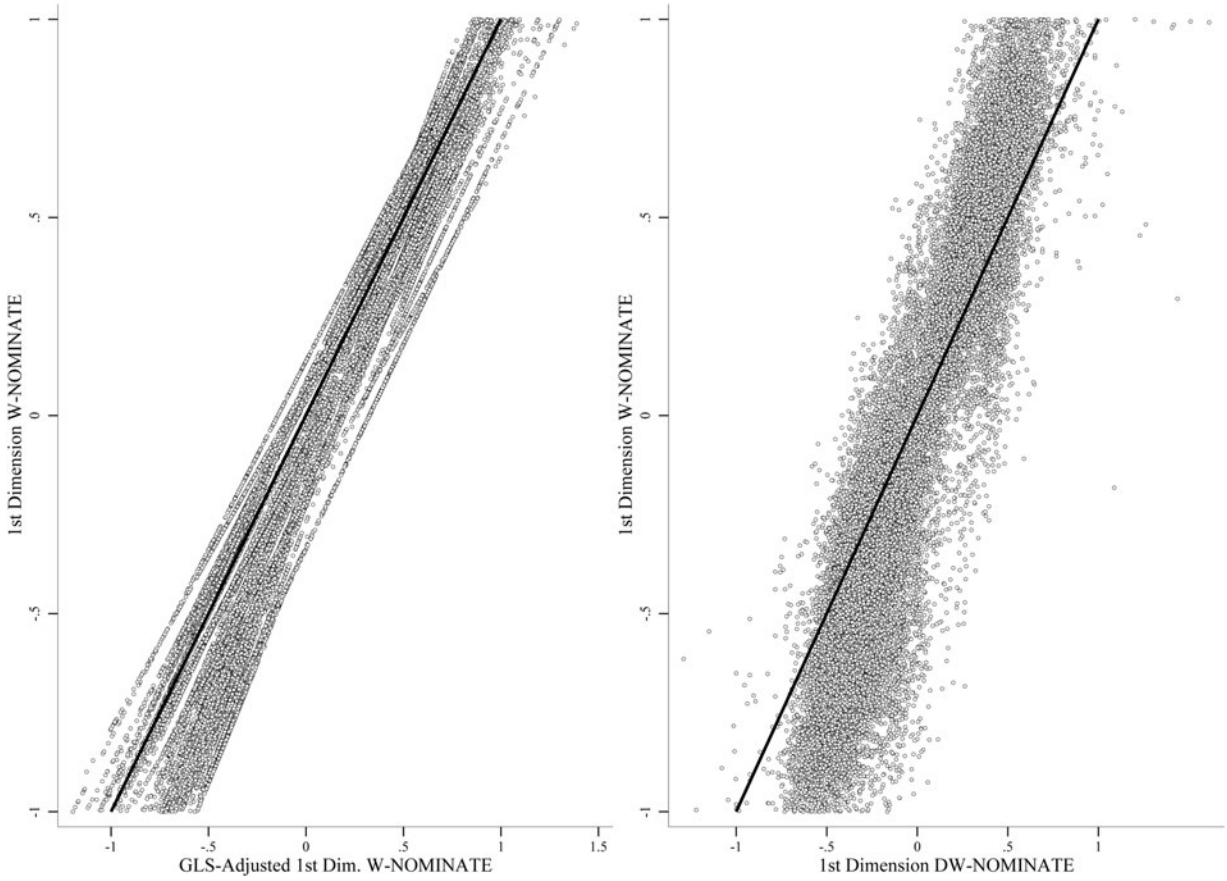


Fig. 6. Comparison of Congress-Level Scores, DW-NOMINATE and GLS-Adjusted W-NOMINATE.

changed across a few adjoining Congresses, in which a common scale is useful but so is maintaining the information specific to a given Congress. By maintaining the ordinal rankings of each legislative session, we limit the degree to which information from earlier or later years bleeds into our estimates for any given year.

To be clear, we are not claiming that GLS-adjusted ideal points are necessarily an improvement over DW-NOMINATE or other nonlinear measures. Rather, we are suggesting that in many cases the advantages to maintaining session-specific information and allowing maximum member flexibility on a common scale will outweigh the disadvantages of this approach relative to others. But the GLS method still imposes its own structure on the data, and the assumptions and limits of this method do present some disadvantages.

The most critical assumption for this technique is that for any subset of members who serve at the same time, without any turnover, the mean ideal point score will stay the same.[43] The shift and stretch parameters are estimated on the assumption that any individual-level movement is idiosyncratic: Any individual member can become more or less conservative or liberal, but across all members, there can be no *general* tendency for *individual* members to become more or less conservative or liberal. The mean of the chamber can move in a systematic fashion over time, but this will be largely the result of member replacement.

Caution is required when we believe that there might be systematic change in how members' individual preferences change over time. For example, if we expect that the Voting Rights Act and its subsequent reauthorization induced Southern Democratic members to become more liberal, this trend can be captured by GLS. But if a shift leftward among Southern Democrats were to occur at the same time as a liberal drift among all sitting members, Southern and non-Southern, then the

43. The assumption that the mean ideal point of members who continuously serve together without any turnover remains stable is an extension of work on judicial ideology. Lawrence Baum, "Measuring Policy Change in the U.S. Supreme Court," *American Political Science Review* 82 (1988): 905–12; Groseclose et al., "Comparing Interest Group Scores," 36.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

assumption would not hold and the estimated issue space would be biased.[44]

3.2 Ideology and Issue Substance

If we are to better approximate "ideology" through ideal point scores, more information needs to be integrated into the estimation procedures. And some scholars have begun to do just this, integrating ratings from external organizations to more precisely capture what is meant by liberalism or progressivism or projecting the written opinions of newspaper editors and columnists into the same space as members of Congress.[45] This is an important endeavor, and we are confident that the new generation of scores they are producing, drawing on a wider range of information as to the content of different ideologies, will be of considerable use for scholars of APD.

An alternative route is to sidestep the question of ideology altogether, not because it is unimportant but because it is not the only component of legislative politics, and because the equation of estimated ideal points with ideological beliefs is not always the most useful way of using these scores in our analyses. Rather than calculate additional dimensions relevant only at distinct moments and whose issue content is poorly defined, we disaggregate roll calls by coherent issue areas and use these to generate issue-specific ideal point scores.[46] The immediate advantages of doing this are that their interpretation is more straightforward and they allow for more nuanced accounts of legislator preferences across issues areas and of their role in policymaking.[47] As we shall discuss in the next section, they also facilitate the generation of nonagnostic scores that better pin down the policy space across time.

To generate the issue-specific scores, we rely on the coding scheme outlined in Katznelson and Lapinski.[48] The main advantage of issue-specific scores is that they have a relatively clear meaning: They are a spatial mapping of political conflict within a given issue area for a defined period of time. Figure 7 shows the location of the party medians on the tier 2 category of "civil rights" from 1877 to the present. Compared with the second-dimension scores in Figure 2, these have a relatively straightforward interpretation, with the positive pole being the more racially conservative or anti-egalitarian and the negative pole the more racially liberal or egalitarian position. The well-known switch between the Northern Democrats and the Republican Party is clearly visible, becoming evident in Congress during the New Deal and with 1964 being an inflection point in a trend already well underway. So is the much later movement of Southern Democrats toward the mainstream of their party.

44. That is, if sitting members on average drifted to the left during this period, then the recovered space would be presumed to be more conservative than it actually was. This is a fundamental problem with all scaling methods, none of which can fully account for changes in the underlying space. This is one of the central motivations for integrating information about the policy agenda into the ideal point estimates themselves. Another limitation to linear mapping methods such as GLS is that the different chambers or sessions being adjusted must be on the same underlying dimension. This is often relatively straightforward, but when it is not the case, the procedure will in effect be regressing two different sets of coordinates. In the case of GLS, the regression coefficient will accordingly be small. Recall that the regression coefficient β is used as the stretch parameter and the denominator in the adjustment formula $\hat{y}_{it} = \frac{y_{it} - \alpha_t}{\beta_t}$. As β decreases in size, $\hat{y}_{it}$ will increase, and the adjustment will lead to exploding estimates.

45. See Hans Noel, "Separating Ideology from Party in Roll Call Data: Why NOMINATE Doesn't Measure Ideology but Can Be Used to Measure Polarization," and Devin Caughy and Eric Schickler, "Structure and Change in Congressional Ideology: NOMINATE and Its Alternatives" (papers presented at the Congress & History Conference, University of Maryland, College Park, June 11–12, 2014).

46. As we have seen, the way in which issue substance is usually handled in an ideal point context is to treat the dimensions as recovering distinct positions across different issues. Although most early research on roll call voting in Congress had concluded that members "respond to many issues in terms of fairly broad evaluative dimensions," Poole and Rosenthal's work drastically reduced the number of dimensions from four or five to at most two, with the first dimension doing most of the work and the second at best a "second fiddle." Warren Miller and Donald Stokes, "Constituency Influence in Congress," *American Political Science Review* 57 (1963): 45–56, 47; Poole and Rosenthal, *Congress*, 54; Keith Poole and Howard Rosenthal, "Dimensional Simplification and Economic Theories of Legislative Behavior," *Economics and Politics* 6 (1994):163–71, 171.

47. The question of over what votes to aggregate—by issue area, by type of vote, by legislative period—is shared across all analyses of roll call data, including those that calculate party unity or Rice cohesion scores. See V. O. Key, *Southern Politics in State and Nation* (New York: Alfred A. Knopf, 1949); Duncan MacRae Jr., *Parliament, Parties, and Society in France, 1946–1958* (New York: St. Martin's Press, 1967); William O. Aydelotte, "Voting Patterns in the British House of Commons in the 1840s," *Comparative Studies in Society and History* 5 (1963): 134–63; Michael H. Crespin and David W. Rohde, "Dimensions, Issues, and Bills: Appropriations Voting on the House Floor," *The Journal of Politics* 72 (2010): 976–89; and John Lapinski, *The Substance of Representation: Congress, American Political Development, and Lawmaking* (Princeton, NJ: Princeton University Press, 2013).

48. The main advantage of this issue scheme in particular is that it is nested, with each roll call assigned a code corresponding to a set of fine-grained and deductively generated issue categories, which are then aggregated into bulkier categories corresponding to issue areas common to most sovereign states and national legislatures. The virtue of this arrangement for estimating ideal points is that it helps overcome the problem of insufficient data. There are unlikely to be sufficient roll calls on the question of "religion" in many Congresses to reliably estimate a set of issue-specific scores. But at the cost of some precision, we can treat the question of "religion" as bound up with other questions with which it is closely related, such as "loyalty and expression" or "privacy." We can then estimate scores for a bulkier category that theoretically encompasses the more fine-grained issue codes. Ira Katznelson and John S. Lapinski, "The Substance of Representation: Studying Policy Content and Legislative Behavior," in *The Macropolitics of Congress*, ed. E. Scott Adler and John S. Lapinski (Princeton, NJ: Princeton University Press, 2006), 96–126.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

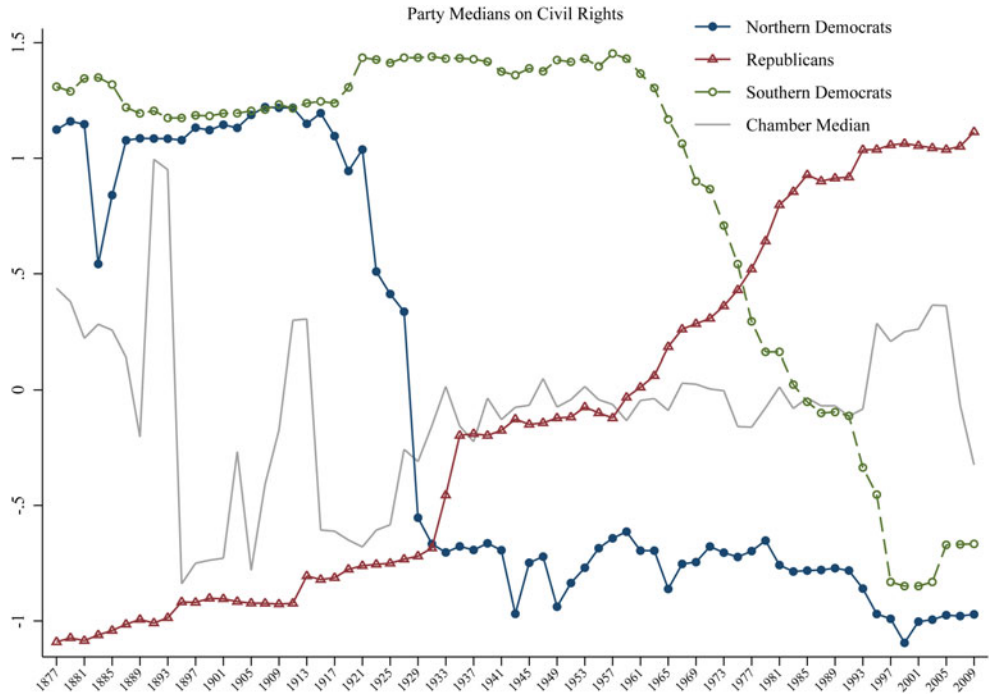


Fig. 7. Party and Regional Medians on Civil Rights.

The issue-specific scores will usually be highly correlated, both with aggregate scores and with other issues. This will be especially true if the issues were understood by legislators as being mutually implicated or where policy positions were structured by considerations common across the issues, such as partisanship.[49] But the degree to which they are different over time and for given legislators and blocs of legislators is potentially important. Table 1 shows the correlations between four issue areas, with the aggregate scores estimated using all roll calls.[50] There is a very high correlation between the aggregate scores and domestic policy as well as government organization. This is unsurprising: Most roll calls have tended to be in the area of domestic policy, while government organization includes votes on congressional organization that are likely to be highly partisan. But international relations and civil rights have much lower correlations, especially for particular historical moments (not shown), and these differences are even more pronounced when we look only at the correlations within each party bloc.The differences in member locations across issue areas can be especially important when we turn to more fine-grained analyses of legislators and their involvement in specific policy disputes. This is particularly important for much APD work, and we believe will help APD scholars better identify critical legislators and evaluate the structure conflict in a given issue area at a given time. Consider the case of Senator J. William Fulbright (D-AR), who was consistently among the more liberal and multilateralist members of the Democratic Party on issues of international affairs, but who was also a Southern Democrat opposed to most civil rights legislation and generally a conservative Democrat on domestic policy.[51] Using DW-NOMINATE, Fulbright is ranked as the fourth most liberal Democratic Senator, out of fifty-eight, for the 91st Congress. Using issue-specific scores, he is ranked thirteenth for international affairs and forty-second for domestic policy. Consider also Senator William Langer (R-ND), a staunch isolationist who voted against the United Nations Charter, but was also a populist leader of the state Non-Partisan League who supported maternity benefits for government employees and increased Social Security benefits, and was praised by *The Afro-American* as "essentially a liberal" whose vote could "always be counted on the side of civil rights legislation."[52] He ranks as the most liberal of fifty-two

49. In *Congress*, Poole and Rosenthal estimated ideal points for five distinct issue areas in the 95th House of Representatives, finding that the ideal points had high correlations across issue areas. Poole and Rosenthal, *Congress*, 55.

50. These are equivalent to the GLS-adjusted W-NOMINATE scores in Section 3.1.

51. See Richard L. Wilson, "Fulbright, J. William," in *American Political Leaders*, American Biographies (New York: Facts On File, 2002).

52. Eileen Boris, "'No Right to Layettes or Nursing Time': Maternity Leave and the Question of U.S. Exceptionalism," *Workers across the Americas: The Transnational Turn in Labor History*, ed. Leon Fink (Oxford: Oxford University Press, 2011), 171–93, 185;

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

Table 1. Correlations Between Aggregate and Issue Specific Scores

Correlations between Ideal Points across Issue Areas			
Issue Area	All Members	Democrats	Republicans
Civil Rights	0.133	0.490	−0.053
Organization and Scope	0.892	0.727	0.564
International Relations	0.626	0.500	0.317
Domestic Policy	0.964	0.912	0.839

Republican Senators for the 83rd Congress on domestic policy, but is among the most conservative Senators on international affairs. Such extreme variation across issue areas is not common, but it is also not rare.[53] And it is likely to be more important for the detailed historical work characteristic of APD, in which individual agency and policy entrepreneurship often plays a larger role in the causal narrative.

Variation in members' positions and rank ordering across issue areas also matters for assessing legislative dynamics, such as identifying the members who are likely to be critical on a given vote. Despite generally similar underlying structures across these different issue areas, the pivotal actors will likely be different, because of personal idiosyncrasies, investment in issue expertise, or strongly felt constituent interests. Aggregate estimates will likely mislead researchers as to the identity and location of the pivotal members, which matters not just for focused studies of issue areas but for analyses of more macro-oriented analyses of Congress; most votes in Congress, after all, concern a question of policy substance.

The left panels of Figure 8 report the proportion of times a given legislator was identified as the median member of the 89th Senate, for aggregate and issue-specific international relations ideal points.[54] The right panels of Figure 8 report the location of the median pivot—the average location of the median member across a sample of 500 observations from the posterior distribution—and the thirty surrounding members.[55] It is clear from even a casual inspection that the lists are significantly different, and only three Senators whose ideal points overlapped with the median pivot's confidence intervals for either the aggregate or issue-specific estimates appeared in both sets.

Issue-specific scores allow for more nuanced analyses of legislators' preferences and have a more straightforward interpretation than a general ideology score. But they also have their own limitations. For one, these scores share a broader limitation with all ideal point techniques in that they impose a theoretical model on a behavioral pattern that might not always be correct. That is, we assume voting is determined by spatial preferences over policy rather than partisanship or other non-policy-based rationales for voting decisions. It is certainly the case that at some times and for some members, this model will not usefully explain member voting patterns, in which case the generated spatial map will not array members according to their preferences. By more clearly specifying the votes over which these scores are generated, however, the researcher will be better able to evaluate what is producing the estimated scores. Another limitation is that by disaggregating roll calls, we are reducing the number that are included in the estimation, resulting in generally noisier estimates. And for some Congresses there will simply not be enough roll calls, meaning that if estimates are to be made, then roll calls from adjacent Congresses need to be pooled, a trade-off that flattens out some of the temporal information that we hope to retain. Finally, while the scores are more easily interpretable, they are still agnostic to policy development and institutional context.

These are real trade-offs, and in some cases, issue-specific scores will likely not be possible. But it is important to note that these are problems that can be moderated. For instance, if a researcher decides to pool roll calls within an issue area across adjacent Congresses, they can select the time period so that it does flatten out specific inflection points.[56] And, as we discuss next, issue-specific scores can also facilitate the estimation of scores

Eileen Boris, "Sen. William Langer," *The Afro-American*, Nov. 21, 1959.

53. For example, if each party caucus was divided evenly into three blocs—liberals, moderates, and conservatives—nearly 20 percent of House members would at some point be characterized as "liberal" and "conservative" across these two issue areas.

54. We follow Clinton and colleagues in identifying likely pivotal members by (1) sampling legislators' ideal points from the joint posterior distribution, (2) ranking the sampled ideal points, (3) identifying which member is in the pivotal position, and (4) repeating this a large number of times, reporting the proportion of times that a set of legislators are in the relevant position. See Clinton et al., "Statistical Analysis of Roll Call Data," 360. The dashed line is the mean proportion for those members who were at any point identified as the median member.

55. Note that the median pivot is more precisely estimated than the surrounding members.

56. In other work we have estimated two sets of pooled ideal points in the area of labor policy, but pooled only those Congresses before and after a suspected inflection point so as to not include information from after the inflection in the first set of estimates. David Bateman, Ira Katznelson, and John Lapinski, "*Southern Politics* Revisited: On V. O. Key's 'South in the House,'" *Studies in American Political Development* 29 (2015): 154–84.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

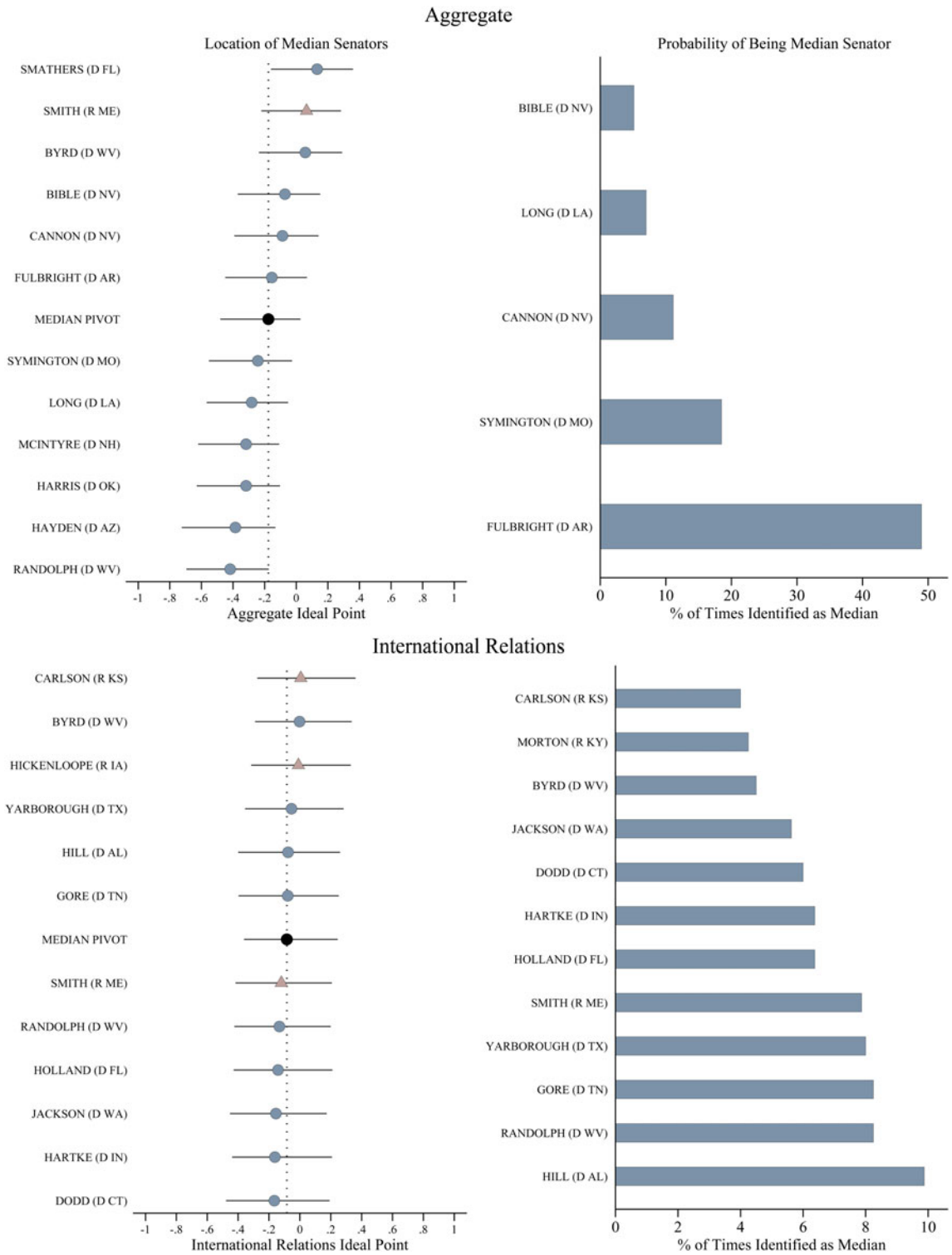


Fig. 8. Location and Probability of Median Voter in the 89th Senate, International Relations and Aggregate.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

that incorporates substantive knowledge about how a policy area has developed over time.[57]

3.3 Temporally Fine-Grained and Nonagnostic Ideal Points

DW-NOMINATE's agnosticism toward policy substance, policy development, and institutional and historical context potentially results in scores that are not genuinely comparable across distinct historical periods or long periods of time. The problem of issue change discussed in Section 2.3 for estimating reliable and comparable ideal points, however, is not entirely insurmountable. One way to proceed is by limiting our use of cross-time ideal points to discrete and relatively short periods, in which there might be good reason to believe that important contextual changes—such as the development of an American state with national reach and regular impact on the lives of its citizens, the re-emergence of a nationally competitive congressional environment, or the emergence of strong party control over the legislative agenda—are kept relatively constant.

The obvious disadvantage here is that we lose much of the interesting variation that accompanies having a distinct score for each legislator for each Congress. If we are examining only one session of Congress, a spatial map of legislators' preferences might be useful; but if we are interested in following how legislative dynamics developed over time during this session, a single set of scores provides relatively little information. One way around this is to follow Adam Bonica, who has estimated rank-ordered ideal points for each roll call. This allows for identifying inflection points in individuals' careers and in the relative locations of parties across much more restricted periods of time, including over the course of a single legislative session or an extended debate over an issue.[58]

We proceed in a similar vein. We first generate aggregate ideal point scores in IDEAL on a roll call matrix composed of all roll calls within a twenty-week window. Alternative specifications of time can be chosen to fit the research interests to the number of roll calls in a given period. We advance this window one week at a time, so that the first set of scores covers weeks 1–20; the second, weeks 2–21; the third, 3–22; and so on. Once distinct sets of ideal points are estimated for each interval, the scores are GLS-adjusted so as to be placed on the same scale. Each legislator, then, has a comparable ideal point for every week in which he or she was a sitting member and in which Congress was in session.[59]

This set of scores has the advantage of allowing scholars to closely assess what, if any, impact important events or shocks might have had on voting patterns in Congress. Consider, for example, the impact of American entry into the two world wars of the twentieth century. The top panel of Figure 9 shows the location of the median Southern and Northern Democratic Representative from 1928 to 1947, where each of the connected points corresponds to a single Congress. The bottom panel of Figure 9 shows the same information through 1948, but estimated on a weekly basis, with each of the connected points corresponding to the sectional median for the twenty-week window centered on that date. Like DW-NOMINATE, this includes information from the past and future; but this information is also limited to a distinct window of time, allowing the researcher to more closely examine whether shifts are associated with changes in the agenda or external events.

In both cases, many of the changes are associated with the start of a new Congress and the arrival and departure of members. But we are also able to identify clear inflection points or abrupt shifts associated with particular events and policy disputes. The attack on Pearl Harbor is accompanied by an immediate move by the Southern Democratic Representatives toward the Republicans, as the House voted on amendments to the Employment Stabilization Act, a bill to invest the Circuit Court of Appeals with original and exclusive jurisdiction over the detention of aliens, measures related to the Special Committee to Investigate Un-American Activities, and other bills on which the Southerners' preferences were increasingly aligned with the Republicans. And in the waning days of the 77th Congress, the positions were quickly reversed as Republicans and Northern Democrats joined against Southerners on an anti-poll-tax bill.[60] Neither of these abrupt changes in voting patterns is captured by the Congress-level estimates.

Figure 10 shows similar information corresponding to American entry into World War I. The top panel shows the location of the party and regional medians estimated at the Congress level, and the bottom panel shows the same thing but estimated on a weekly basis centered on an individual roll call.

57. Incorporating substantive information can both increase our understanding of policy change while mitigating to a certain extent the problem of sparse data. For example, we might have very strong priors about certain members, or we might have alternative but incomplete measures of policy liberalism. By integrating these into the estimates, we are able to use fewer roll calls more efficiently.

58. Adam Bonica, "Punctuated Origins of Senate Polarization," *Legislative Studies Quarterly* 39 (2014): 5–26.

59. For more information on this procedure, see Bateman et al., "*Southern Politics* Revisited."

60. The quick reversal of positions suggests that this particular twenty-week period might be on a different dimension than the others, making the use of GLS problematic. Where this occurs, it might make sense for the researcher to drop those periods that are off-dimension. But this technique also allows the researcher to note when such off-dimensional debates punctuate congressional voting, rather than reducing these to a single second-dimensional score estimated for the entire Congress.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

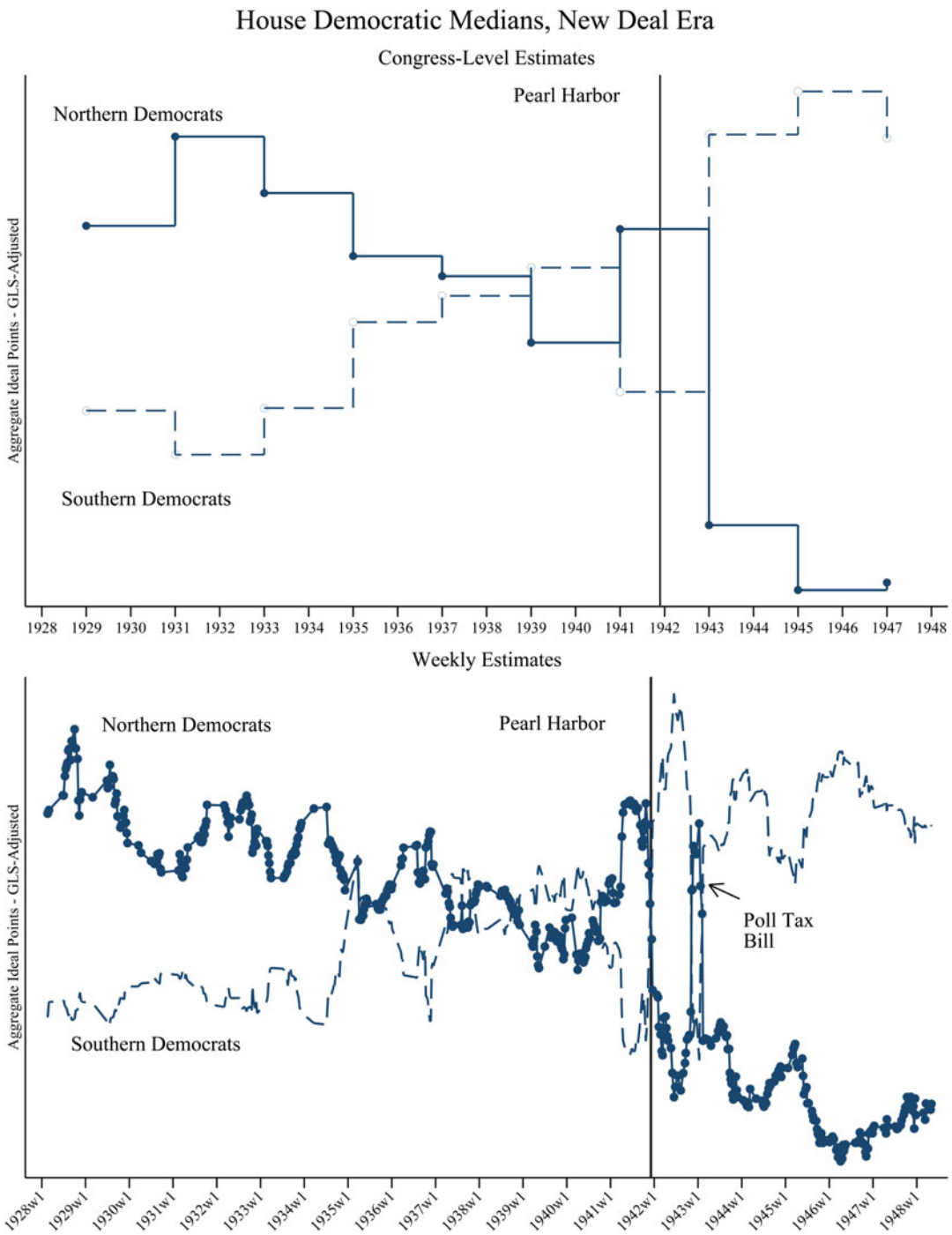


Fig. 9. House Democratic Medians in the New Deal Era, Before and After Pearl Harbor.

American entry into the war corresponded with a brief but important collapse of party lines, as the parties disagreed on conscription, on new revenue measures, and on whether to prohibit alcohol for newly mobilized soldiers. Party voting lines were quickly restored, but only after several weeks of off-dimensional political conflict. Again, this temporary outbreak of bipartisanship is obscured by Congress-level estimates. And while this might ultimately not have been important in the broader development of party positions, it is the sort of historically contextual change, with important consequences for the formation of national policy, to which APD scholars are likely to be attuned.

The trade-off to such temporally fine-grained measures is that changes in ideal points are more likely,

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

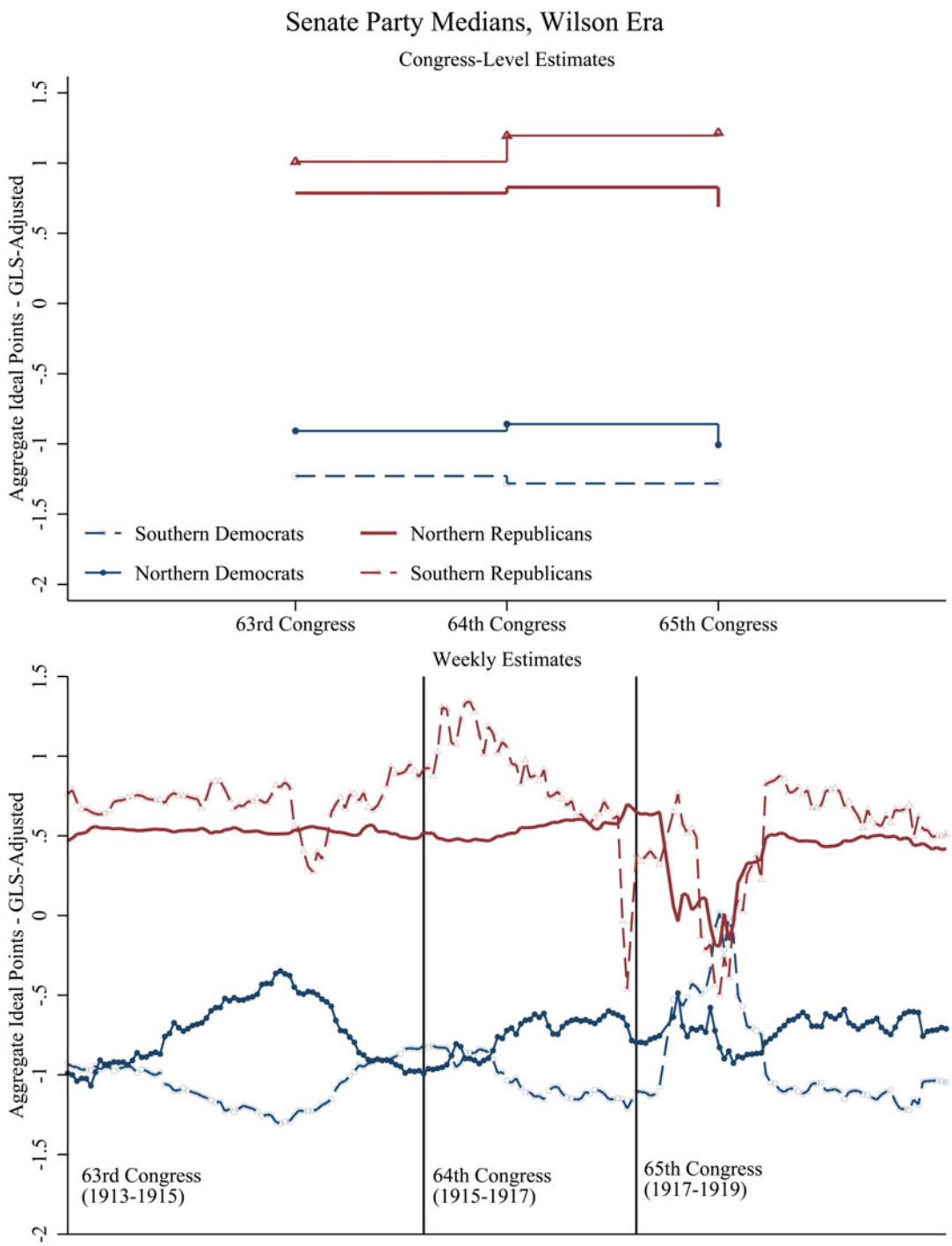


Fig. 10. Party Medians in the Senate during the Wilson Administration, Before and After Entry into WWI.

given the smaller number of votes, to reflect random fluctuations rather than substantive changes. As with all measures, caution in their use is required and scholars should avoid making much of small changes. Close inspection is required to assess whether the changes are random or reflect actual changes in voting behavior, however temporary. But by maximizing the information recovered from discrete periods, we are able to control for some of the changes in context that render cross-time comparisons difficult to interpret while generating scores that retain their utility.

Alternatively, we can generate scores that follow John Londregan's call for scholars to "think … more

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

carefully about the agenda on which legislators vote" and to begin integrating information about policy development directly into the estimation process itself.[61] An additional advantage of estimating issue-specific scores is that they allow the researcher to trace a single policy issue across time, integrating information not only about who members voted but about the policy proposals they voted upon.

Most of the attention given to DW-NOMINATE is paid to the legislator ideal points. But the roll call parameters—including the location of midpoint separating the status quo and the policy proposal being voted on—are also projected into the same space. These convey potentially useful information, and we can compare the location of the midpoints to our understanding of policy development.

The top panel of Figure 11 shows the location of the roll call midpoints—the first dimension DW-NOMINATE location halfway between the policy proposal and the status quo—for all votes on Social Security between 1935 and 2003. A set of final passage roll calls that moved policy in a more liberal direction are highlighted: the enactment of the program in 1935, the liberalizations and extensions of late 1940s through the 1960s, and the decision to reinstate the minimum benefits that would have been cut in the accompanying Budget Reconciliation Act of 1981.[62] If the policy space were stable, the liberalizations to the program should be accompanied by midpoints that moved to the liberal side of the space.[63] In fact, we see no coherent trend, and during the height of the program's liberalization, the midpoints remain relatively close to each other. In fact, the extensions of coverage and liberalization of benefits in 1949 and 1961 are to the right of the establishment of the program in 1935, and the 1961 extension was to the right of the 1949 Act. This strongly suggests that the policy space is unstable, that the policies being voted on are not arrayed on a stable dimension across relatively short periods of time in a way that would allow us to identify a more liberal or conservative policy.

But we do have substantive information about the policies under consideration, and for many votes we can say with some confidence that policy was moved leftward and, accordingly, that the midpoint on the final passage roll call should be to the left of the midpoint on an earlier roll call that set the status quo. This information can be directly integrated into the estimation procedure, by constraining the relative location of the midpoints for individual roll calls about which we have strong priors or by inferring from a legislator's support for liberalizing benefits at a given moment that they would have supported earlier liberalizations that these are building on, as well as the establishment of the program itself.[64] The result is a better anchored policy space, in which shifts in member ideal points more directly map on to shifts in policy preferences.

We first estimate a set of ideal points specific to Social Security roll calls, by merging the roll calls from different Congresses into one matrix with bridge members as the "glue."[65] This provides a baseline of agnostic estimates—generated only by looking at who votes with whom on Social Security votes—against which we can compare an "informed" set of estimates. We generate this second set of scores by using the *Congressional Quarterly* (*CQ*) *Almanac*'s Policy Tracker to examine the debates and votes on Social Security over time. Where voting yea on a roll call infers support for an earlier vote, we impute a yea vote for this earlier roll call for all members who voted yea on the later roll call.

The bottom panel of Figure 11 shows the location of the party means on Social Security, for both the agnostic and informed estimates. Both show a broadly similar story, of the Republican Party moving leftward in the 1940s through 1960s, and then gradually moving back to the right starting in the 1970s, and the Democratic Party trending leftward over time. What is different is the extent and rapidity of the Republicans' leftward shift, moving dramatically leftward in the 1930s and 1940s. While a significant portion of Republicans had opposed the creation of a contributory insurance fund in 1935, the 1948 Republican platform called for its extension and for an increase in benefits, Republican Senators proposed expanding the program to an additional 3.5 million workers, and in 1950 a large expansion of the program passed the House with Republicans committing to "support the legislation almost to a man."[66]

61. Londregan, "Estimating Legislators' Preferred Points," *Political Analysis* 8, no. 1 (1999): 36.

62. We only include votes on the Social Security program, rather than Medicare or other programs attached to the Social Security Act.

63. The midpoint is halfway between the existing status quo and the new proposal. If this proposal passes, as the five identified roll calls did, then the location of the policy proposal becomes the new status quo point. If a subsequent proposal would liberalize the program further, then the new roll call midpoint should be located to the left of the previous roll call midpoint.

64. In imputing votes based on inferred policy positions, we are following Bailey, "Comparable Preference Estimates." For our purposes, imputing votes and constraining the midpoints are effectively equivalent, and we rely on the former approach here.

65. Boris Shor, Christopher Berry, and Nolan McCarty, "A Bridge to Somewhere: Mapping State and Congressional Ideology on a Cross-Institutional Common Space," *Legislative Studies Quarterly* 35 (2010): 417–48.

66. Many Republicans did support moving the program more fully to a pay-as-you-go system funded through income taxes, but this was the extent of opposition, and the Senate passed the bill 81–2 and the House passed the conference report by 374–1. Even the lone Republican who voted nay explained that he felt obliged to vote against, as he had offered a motion to recommit as a tactical measure intended to save an amendment concerning when the federal government could withhold compensation funds from the states, and under House rules only those opposed to the bill could propose recommittal. "Social Security Act," in

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

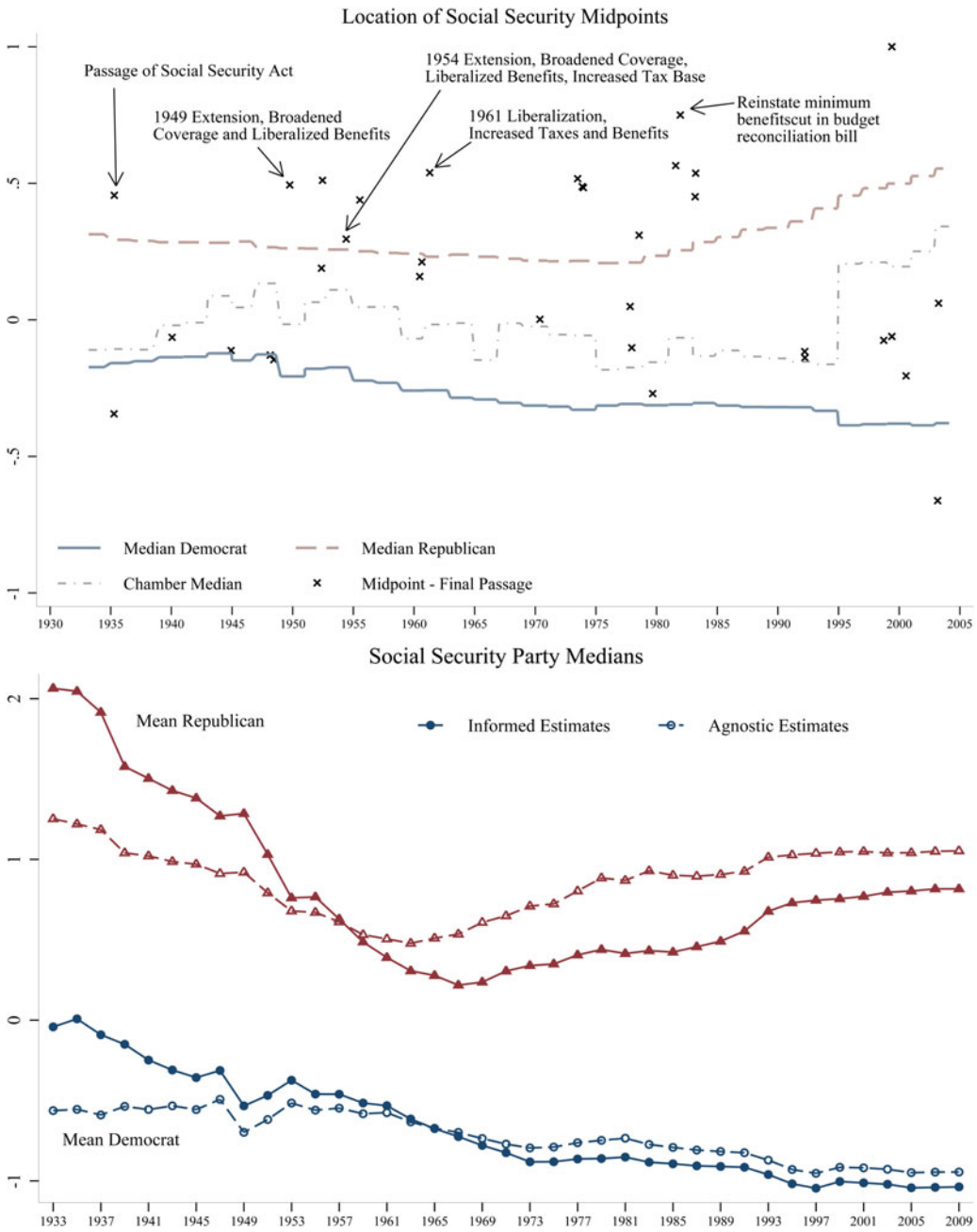


Fig. 11. Anchoring the Policy Space, Social Security.

Similar scores that anchor the policy space can be generated wherever policy can be arrayed on a coherent dimension, such as in civil rights, marginal tax rates, or the liberality of welfare state programs.[67]

CQ Almanac 1950, 6th ed., (Washington, DC: Congressional Quarterly, 1951), 165–77. http://library.cqpress.com/cqalmanac/cqal50-1377221.

67. Joshua Clinton and Adam Meirowitz also generate nonagnostic scores that integrate information about the roll call parameters. They note that given the underlying behavioral model the ordering of votes in a given session should matter, but that under standard techniques it does not. The status quo point changes with each roll call, but this fact is disregarded in most estimation techniques because we do not have an independent ability to locate the status quo or the proposal in a common space. They propose a model in which the estimated location of the winning position of the immediately preceding roll call is used as the status quo point for the next roll call to occur in the same issue area. See Joshua D. Clinton and Adam Meirowitz, "Testing Explanations of Strategic Voting: A Reexamination of the Compromise of 1790,"

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press

4. CONCLUSION

The measures produced by Poole and Rosenthal fundamentally changed how we have studied Congress, with important impacts on the study of American politics in general. Nonetheless, their pioneering effort has had little influence on the study of APD. This article's intent is not to insist upon the use of DW-NOMINATE or other ideal point techniques. For some questions, however, these tools can be immensely useful, and in these cases we believe they can be judiciously integrated into APD analyses to help us not only to better understand the legislative dynamics around particular issues at particular moments but also to make broader claims about political development. To achieve this requires measures that are sensitive to the methodological concerns of APD scholars, who rightly stress the role of historical and institutional context. We have introduced some possible alternatives in this article.

We believe that we are at a propitious moment in both the study of Congress and of APD. This truly is an opportunity for mutual gains from trade. For Congress scholars to make further advances in the statistical analysis of legislative behavior, we need to better integrate the substantive knowledge, historical methods, and theoretical frameworks motivating much APD scholarship. Congress scholars have much to gain from the accumulated historical knowledge, as well as the theories of institutional and policy development, that APD scholarship has been generating to help build measures that much better match the reality of what is actually going on in Congress. And scholars of APD, we suggest, will gain measures that are more accommodating to their methodological orientations and substantive interests. For these gains to be realized, however, requires the development of new techniques that do not treat such concerns as an afterthought, but that are designed from the outset to appeal across the persistent methodological divides in the study of American politics and APD. We hope this essay has provided some useful suggestions along these lines.

American Journal of Political Science 48 (2004): 675–89; Bailey, "Comparable Preference Estimates," 442, 439.

https://doi.org/10.1017/S0898588X16000080 Published online by Cambridge University Press