
David A. Bateman
Cornell University
Joshua D. Clinton
Vanderbilt University
John S. Lapinski
University of Pennsylvania

Abstract: The study of political conflict in legislatures is fundamental to understanding the nature of governance, but also difficult because of changes in membership and the issues addressed over time. Focusing on the enduring issue of civil rights in the United States since Reconstruction, we show that using current methods and measures to characterize elite ideological disagreements makes it hard to interpret or reconcile the conflicts with historical understandings because of their failure to adequately account for the policies being voted upon and the consequences of the iterative lawmaking process. Incorporating information about the policies being voted upon provides a starkly different portrait of elite conflict—not only are contemporary parties relatively less divided than is commonly thought, but the conflict occurs in a smaller, and more liberal, portion of the policy space. These findings have important implications for a broad range of work that uses elite actions to compare political conflict/polarization across time.

Replication Materials: The data, code, and any additional materials required to replicate all analyses in this article are available on the American Journal of Political Science Dataverse within the Harvard Dataverse Network, at: http://dx.doi.org/10.7910.

The study of political conflict is at the center of political science, and the study of political polarization in particular is among the most important avenues of research today, centrally important for studying lawmaking, representation, and the performance of American political institutions (Binder 1999, 2003; Fiorina, Abrams, and Pope 2005; Hetherington 2009; Krehbiel 1998; Lee 2008; Levendusky 2009; McCarty, Poole, and Rosenthal 2006). What polarization really means may be unclear (Lee 2015), but many have used measures based on the analysis of roll-call voting behavior to argue that the average policy preferences of the Democratic and Republican parties are now more extreme and divided than at any time following the Civil War (e.g., Hare and Poole 2014).

Claims regarding the historical levels of polarization have reached beyond the academy, and the characterization is often featured in the popular press. Beyond the study of polarization, similar assumptions regarding the ability of roll-call-based estimates to measure the policy preferences of elected officials have resulted in their pervasive use in the study of legislative politics and representation (e.g., Binder 1999; Binder and Maltzman 2002; Brady, Han, and Pope 2007; Chioi and Rothenberg 2003; Clinton 2006; Coleman 1999; Hetherington 2001; Johnson and Roberts 2004; Masket 2007; Schickler 2000; Theriault 2006).

We argue for caution in making such claims, and we demonstrate that it is difficult to reach any conclusion about the meaning of past and contemporary policy disagreements without accounting for how the policy content has changed across time. We do so by focusing on the political conflicts concerning African American civil rights since 1877. Perhaps no conflict in American politics is more important and long-standing than the conflict over black Americans’ legal status and civil rights, from the founding through the Civil War and Reconstruction and up to the present. Focusing on elite conflict in this issue not only probes the ability of political elites to deal with enduring issues, but it also provides a rich historical record against which competing characterizations of policy preferences can be compared.
Despite being more likely to cast party line votes in recent congresses, for example, elected officials have demonstrated that their preferences on civil rights issues are almost certainly more similar now than at any other time since the Civil War—members may disagree about the appropriate formula for preclearance under the Voting Rights Act, or even whether preclearance remains an important policy remedy, but few if any contest whether African Americans should have the right to vote.

We use the case of civil rights to argue for a revised interpretation of measures that are commonly used to characterize ideological disagreement and political conflicts in several respects. First, we show that current measures of political conflict provide erroneous conclusions about elite policy preferences regarding civil rights because they fail to account for the evolving policy agenda. For example, measures commonly used to characterize policy preferences over time implausibly suggest that contemporaneous Democrats would oppose federal policies designed to protect black Americans’ civil rights in the post-Reconstruction era and that there has been no change in the ideological scope of the policies being debated or the most preferred policies in over 100 years despite dramatic changes in American politics and society.

Second, accounting for the differences in the issues being voted upon by leveraging the relationship between linked policies such as the 1965 Voting Rights Act and its subsequent reauthorizations provides a much different characterization. While partisan-based divisions have increased in the post–World War II era, current inter-party differences are considerably diminished relative to historical levels. Moreover, the scope of political conflict on civil rights has narrowed considerably relative to the issues being debated in post-Reconstruction congresses; the domain of current political debates is far more limited and liberal.

Our findings have important consequences for how we study representation, lawmaking, and the performance of political institutions. For one, we do not believe that the problem we identify is limited to the area of civil rights, but will likely apply in other issue areas where there has been a substantial change in how issue positions map onto partisanship.\footnote{See the online supporting information for a list of possible issue areas.} If, as we argue, existing roll-call-based measures face difficulties when characterizing ideological disagreements over time, it is difficult to interpret the meaning of the many regressions that use such measures. While some have rightfully urged caution (see, e.g., McCarty 2011; Poole and Rosenthal 1989), there has not been a sustained investigation of the ability of roll calls to measure policy disagreement. Perhaps as a consequence, a great deal of scholarship fails to adequately grapple with the potential limitations. Our focused exposition and demonstration of how to account for policy context hopefully highlights critical limitations that are often too quickly assumed away.

### Measuring Political Conflict over Time

Given the importance of quantifying and comparing conflict over time, scholars have long used members’ votes in Congress to characterize changing patterns of political conflict. Roll-call-based measures were originally defined explicitly in partisan terms (e.g., Rice 1928), but modern measures are interpreted by many as capturing differences in policy preferences. For example, the large literature using roll-call-based estimates to measure policy disagreement in terms of “gridlock intervals” or polarization presumes that voting behavior reflects policy preferences and that the resulting estimates can be meaningfully compared over time.

Polarization can mean many things, but it is commonly defined as the difference between the two parties’ average ideal point, estimated using a statistical model applied to roll-call voting in Congress. Figure 1 provides a portrait of elite polarization in the United States that has been largely cemented as an empirical truth: The United States is now more ideologically polarized than ever, and the state of American politics is therefore, as one book title claims, Even Worse Than It Looks (Mann and Ornstein 2012).\footnote{The figure plots the difference between the median Republican and median Democrat. Common space scores assume ideal points are fixed; DW-NOMINATE allows ideal points to trend linearly.}

What exactly this trend means, however, is unclear. While the underlying data suggest that Democrats and Republicans are increasingly likely to vote in opposition to one another, scholars often interpret this increase in partisan conflict as also indicating a growing divergence in each party’s most preferred policies. Moreover, the broader interpretation equating ideal points and policy preferences forms the foundation of an enormous literature in political science and economics.

While members are certainly casting more votes along party lines in recent congresses than they have in the recent past, claims that members are therefore more ideologically divided now than they were following the Civil War seem implausible (a concern raised by Bailey 2007).
To examine the comparability of the recovered estimates over time, we consider how contemporary elites are predicted to vote on prominent past votes using these estimates. The ability to compare the policy preferences over time using ideal points—and therefore construct measures of polarization, gridlock intervals, or any other function of the estimated ideal points, as is done by many—relies critically on such over-time hypotheticals.\footnote{The claim that "Jesse Helms is more conservative than Robert Taft, Sr." (Poole and Rosenthal 2001, 8) implies that Helms is predicted to prefer more conservative policies. McCarty aptly notes that "some caution is obviously warranted in making too much of these comparisons. . . Being liberal in 1939 meant something different than liberal in 1959 or in 2009. So one has to interpret NOMINATE scores in different eras relative to the policy agendas and debates of each" (2011, 79). Determining the extent to which the seemingly common scale is problematic and the extent to which the changing agenda makes equating ideal points and policy preferences difficult over time is precisely our motivation.}

Considering two key votes from the late 19th and early 20th centuries for which we have strong priors as to their policy content—whether to limit federal authority to use the army to maintain peace at the polls in 1877, and whether to enact antilynching legislation in 1922—reveals support for this concern. Figure 2 plots the DW-NOMINATE estimates of every member who served in the House between 1877 and 2011, at the time they cast their first vote on a civil rights issue, relative to the cutting line that is estimated to divide supporters and opponents for each policy. The fact that both cutting lines are nearly vertical indicates that they were largely resolved by differences in the first dimension; ideal points to the left in the left-hand graph are predicted to vote to prohibit the presence of the army at the polls in 1877, and ideal points to the right of the cutting line in the right-hand graph are predicted to vote in favor of antilynching legislation.

Comparing the identities of the plotted ideal points relative to the estimated cutting lines in Figure 2 reveals immediate problems. Democrats sitting in Congress in 1877 and 1922 may not have supported the racially progressive positions associated with these two votes, but it is implausible to think that contemporary Democrats would share those views. It is difficult to imagine that Jack Flynt (D-GA), a signer of the Southern Manifesto, or former Confederate Vice President Alexander Stephens (D-GA)
would be more likely to support retaining the authority of the federal government to use the military to ensure black American males’ voting rights in the South than would John Lewis (D-GA) or Adam Clayton Powell (D-NY). The policy content of the first dimension has changed in ways that are not reflected in the DW-NOMINATE estimates.

The problems go beyond mere shifts of the policy space. That Democrats are always to the left of Republicans despite their changing first-dimension preferences over race suggests that DW-NOMINATE estimates reflect partisan rather than ideological differences—ideal points less than 0 are consistently Democratic, but they are not consistently liberal (Schickler, Pearson, and Feinstein 2010). Because the estimated dimensions are defined in terms of interparty versus intraparty voting coalitions rather than the issues involved, it is unclear how to interpret the policy content if the parties’ positions change over time. This is not a problem that is unique to the issue of civil rights, as it will potentially occur whenever the parties’ relative positions on an issue have changed substantively over time (Karol 2009).

Finally, the distribution of ideal points suggests that the scope of political conflict is unchanged. This stability seems hard to reconcile with the profoundly changing political, economic, and social circumstances in the United States over time and the extensive changes in policies and in preferences among legislators and society at large produced by the civil rights movement. In the early 20th century, many Southern Democrats supported the repeal of the 15th Amendment while at mid-century, Theodore Bilbo (D-MS) urged violence to prevent black voting (Newton 2010, 103–4). Fast-forward to 2006, when a Republican Senate voted 98–0 and a Republican House 390–33 to reauthorize the Voting Rights Act. The vote was not without partisan conflict or ideological disagreement, and conservative amendments to revise the act’s triggering formula and to strike the extension of bilingual ballots were defeated 96–318 and 185–238. Contemporary policy differences are important, but any reasonable measure should show that they pale in comparison to earlier differences.

There are at least two reasons why it is difficult to infer the most preferred policy positions from existing ideal point estimates. First, while a behavioral model of spatial voting can be used to motivate a statistical model of roll-call voting, nothing ensures that the estimates generated by the latter reflect the interpretations suggested by the former. Second, even if we are willing to assume voting is ideologically based, ignoring the changing content of the congressional agenda produces estimates that fail to account for the dramatic changes in policy over time and that are hard to square with historical understandings.

When analyzing elite behavior, most models assume that members vote for the policy alternative that is closest to their most preferred policy, with some
idiosyncratic voting error (e.g., Clinton, Jackman, and Rivers 2004; Heckman and Snyder 1997; Poole and Rosenthal 1989). Estimators differ in what they assume about the distance function members use to evaluate the competing options, but common behavioral models assume that members vote in favor (1) or against (0) a vote depending on

$$\Pr(y_{it} = 1) = \Pr(u(x_i - \theta_{y(t)}) - u(x_i - \theta_{n(t)}) > \epsilon_{it}),$$

where $y_{it}$ is the vote of legislator $i$ on vote $t$, $x_i$ is the most preferred policy of legislator $i$ (ideal point), $\theta_{y(t)}$ is the location associated with the success of vote $t$, $\theta_{n(t)}$ is the location associated with the failure of vote $t$, and $\epsilon_{it}$ is an idiosyncratic error that is assumed to be independent across votes and legislators. Because the only observable parameters in Equation (1) are the votes $y_{it}$ cast by legislator $i$ on vote $t$, interpreting the meaning of the recovered estimates is difficult.

As with any latent variable model, the meaning of the recovered parameters is not defined, and the estimates can be rationalized by any behavioral model that might produce persistent voting coalitions. While the stability in individual voting behavior captured by $x_i$ is often thought to reflect policy preferences (because of the behavioral model used to justify the statistical model), many observationally equivalent behavioral models are possible.

The fact that many behavioral models can provide observationally equivalent rationalizations of the observed voting patterns is revealed by considering the reduced-form representation of Equation (1) provided by Clinton, Jackman, and Rivers (2004):

$$\Pr(y_{it} = 1) = \Pr(\alpha_i + \beta_i x_i > \epsilon_{it}).$$

Because only $y_{it}$ is observed, many interpretations can be given to the estimated parameters. While it may be impossible to statistically identify many of the alternative behavioral models that could be proposed, the lack of statistical identification does not preclude their truth. As a result, while the pattern of divergent ideal points in Figure 1 is usually interpreted as increasing ideological divergence, nothing ensures that this is the correct interpretation.

Similar difficulties arise when interpreting the meaning of the recovered dimensions. Akin to an eigenvalue-eigenvector decomposition in exploratory factor analysis, the dimensionality of the political conflict recovered by NOMINATE models is determined by iteratively fitting higher-dimensional models to account for the residual variation from the lower-dimensional model. As such, the recovered dimensions are based on conflict rather than content; differences in voting coalitions—not differences in the issues being voted upon—distinguish the estimated dimensions. As Poole and Rosenthal (1997, 46) note: "The first dimension divides the two major political parties. The dimension can be thought of as ranging from strong loyalty to one party to weak loyalty to either party and to strong loyalty to the second, opposing party. The second dimension differentiates the members by region within each party."

Even so, the meaning of the dimensions is often interpreted in ideological terms—the first dimension reflects preferences in the liberal-conservative dimension, and the second dimension reflects cross-cutting issues such as those related to race and region—but nothing in the statistical model ensures that either dimension has any relationship to policy outcomes in a stable ideological space. In fact, as with any exploratory factor analysis, the recovered dimensions could be arbitrarily rotated to produce new (correlated or uncorrelated) dimensions that account equally well for the observed behavior. As a result, if parties' positions change—as happened with civil rights (Carmines and Stimson 1989)—it becomes difficult to interpret what estimates based on stable party-based voting coalitions imply about the underlying ideological conflict. The desire to make comparisons of policy positions over time exacerbates these ambiguities; the ability to make comparative statements depends on the claim that an ideal point of 1 in 1980 represents the same set of policy preferences as an ideal point of 1 in 1880 and that the meaning of a one-unit difference in ideal points is unchanged. While statistical identification is often achieved by imposing constraints on the ideal point parameters—assuming that legislator $i$'s ideal point varies parametrically over time (e.g., constant, linear, polynomial)—it is unclear whether they can be sensibly interpreted as reflecting ideological content over time, as is commonly assumed.

Some have explored how the vote-related estimates vary over time (e.g., Clinton 2012), but the studies are limited by issue specificity and temporal reach.\(^5\)

\(^4\)The concerns raised about using roll calls to detect party influence (Krehbiel 2003) can be raised against interpreting ideal point estimates as measuring policy preferences.

\(^5\)An entire literature explores the critical issue of how to compare measures of elite behavior over time (Bailey 2007; Groseclose, Levitt, and Snyder 1999; Martin and Quinn 2002; Poole and Rosenthal 1997). As Figure 1 suggests, most results are not sensitive to the choice of over-time constraint.

\(^6\)Some work attempts to take account of information about the votes, usually within a single congress. Krehbiel and Rivers (1988) consider a sequence of votes on amendments to the minimum wage in the Senate. Clinton and Meirowitz (2004) analyze several dozen
it is common to assume that the statistical identification of roll-call parameters over time also guarantees that it is sensible to give the estimates a consistent ideological interpretation, the sensibility of this assumed interpretation is unknown.

**Elite Conflict over Civil Rights, 1877–2011**

We demonstrate the difficulty of interpreting ideal point estimates as reflecting policy differences over time by focusing on political conflict over black Americans’ civil rights. Focusing on this issue is justifiable for several reasons. For one, it is a vitally important issue in American politics, and it has been since the founding. We also have strong priors about the direction of policy change and policy preferences over time. Policy differences between racial liberals and conservatives may remain large, but any reasonable measure should show that these differences have declined since the early 20th century. Moreover, the scope of the policy debate has also undeniably shifted leftward over time; many racially conservative policies that were once popular and nationally supported are no longer widely acceptable.

To be clear, we do not mean that potent differences on racial politics no longer exist, and we do not deny that partisan differences on race have increased over the last few decades. Instead, we question the prevalent interpretation of ideal points as reflecting policy preferences; measures that suggest Tom Delay (R-TX) is as conservative on race as was Theodore Bilbo, for example, discount the magnitude of white supremacy, the commitment of political elites to maintaining it, and the tremendous costs that it imposed.

We argue that roll-call-based estimates face difficulties in measuring changing policy preferences because the policy content being voted upon is not accounted for. Even if we assume that the behavioral model is a true representation of elected officials’ behavior, common statistical roll-call voting models ignore the policy content being voted upon; no information is used when estimating the location of policy proposals in the policy space \( \theta_y(t) \) and \( \theta_y(m) \), and these parameters are only identified votes pertaining to a supposed logroll. Clinton (2012) considers the relationship between amendment votes on the minimum wage in a congress. Pope and Treier (2011) consider votes related to the Great Compromise in the Constitutional Convention. And Jeong, Miller, and Sened (2009) impose constraints in order to estimate the location of bill proposals in the same space as legislator preferences, allowing them to calculate the uncovered set on civil rights policy in the 1960s and 1970s.

We account for policy content in two ways. To define the meaning of the recovered dimensions in terms of policy content rather than political conflict, we restrict our analysis to those votes that directly involve civil rights.

We use the statistical model of Clinton, Jackman, and Rivers (2004) to fit a single one-dimension model using all civil rights votes cast in the House. For comparability, we assume that members’ ideal points are constant across time, but our results do not depend on this strategy, and the supporting information shows similar results for alternative approaches. By assuming ideal points are fixed over time, the changes in polarization we estimate are due to replacement rather than conversion or adaptation.

To be clear, we do not consider the possibility that the assumed behavioral model is false and voting behavior is driven by factors besides policy preferences. Instead, we investigate what is possible if we are willing to use the same statistical models and assumptions about the equivalence between ideal points and policy preferences by incorporating more information about the policies being voted upon. In particular, can accounting for the relationship between the policies being voted upon produce patterns in the resulting estimates that are more consistent with historical understanding and expectations regarding how policy preferences on civil rights have likely changed over time? If so, accounting for policy content may increase our confidence in the ability to interpret differences in ideal point estimates as reflecting policy preferences.

Figure 3 traces the median ideal point and bootstrapped 95% confidence intervals in the House of Representatives on votes pertaining to African American civil rights for three groups: Republicans, Southern Democrats, and all Democrats. Unlike the estimates of

---

7As a result, Poole and Rosenthal (1997) explicitly warn against using these parameters, and Clinton, Jackman, and Rivers (2004) estimate a reduced form that estimates cutting planes rather than location parameters.

8In doing so, we follow Poole’s suggestion to subset roll calls by issue area to better understand changes in how a given issue maps onto the structure of voting across time (2005, 185). Because the existence of a vote on an issue depends on the willingness of elected officials to consider the issue (Lee 2008, 2009), using the entire roll-call record to make inferences about civil rights preferences possibly conflates policy debates because voting coalitions on the issue break along both party and regional divisions.

9Civil rights roll calls were identified by Katznelson and Lapinski (2006).

10In each case, we assume the estimated ideal points have mean 0 and variance 1 to define the scale.
A HOUSE DIVIDED?

Figure 3 Party Medians on House Civil Rights Votes Using a Civil Rights Dimension, 1877–2009

Party Medians on Civil Rights, House

While the rank ordering of the implied policy preferences held by these groups of elites in Figure 3 is far more consistent with historical understandings than the ordering evident in Figure 2, problems remain. Not only does the magnitude of current party differences on civil rights issues implausibly equal those in the immediate post-Reconstruction period, but the most preferred policies of contemporary Republicans are supposedly equivalent to those of Southern Democrats immediately following the Civil War.

These problematic implications suggest that constraining ideal points over time alone cannot adequately ground temporal comparisons given the evolution of civil rights policy. To do better, we use the sequential and cumulative nature of civil rights policy change in the United States to incorporate information about the relationship between the policies being voted upon. Between 1957 and 2006, for example, a sequence of votes on federal protection of African American voting rights was taken in Congress that can be used to ground the relative location of the policies being voted upon (for a similar approach, see Bailey 2007).

Figure 4 depicts the hypothetical ordering of one such sequence of policy debates concerning voting rights. Because the ordering is based on the qualitative comparison of the content being voted upon, what matters is the relative ordering; we have no information about the actual distances involved (hence, the lack of a scale for the y-axis). Put differently, wherever the pre-1957 status quo...
might have been in the policy space, the 1957 civil rights bill proposed to move policy leftward, albeit by a modest amount. When that bill passed, it became the status quo for the next vote that proposed to move policy even further to the left. As such, when the 1960 Civil Rights Act was proposed, the location of the midpoint on final passage for the 1960 bill should be to the left of the midpoint for the 1957 legislation absent an exogenous shift to the status quo. Upon passage, the 1960 act, in conjunction with the 1957 act, became the new status quo.

Careful consideration of the policy content reveals what can and cannot be assumed about the relationship between policies. For example, the 1965 Voting Rights Act unambiguously moved policy considerably to the left, and so we can infer that the midpoint for the vote to pass the 1965 act should be to the left of the midpoint on final passage of the 1960 act. In 1970, however, although Section 4 of the Voting Rights Act was set to expire, the status quo was not a reversion to the pre-1965 act because many of its provisions had been made permanent. As a result, the status quo of no legislative action represented a conservative move, but a move that would remain to the left of the 1960 status quo. Because the 1970 proposal again moved policy leftward, the midpoint on final passage should be to the left of the 1965 midpoint. The same was again true of the 1975 reauthorization and extensions, which also made the nationwide ban on tests permanent so that the status quo of no legislative action would be to the left of where it had been in 1975 when the expiring sections were next up for reauthorization in 1982. The 1982 amendments and reauthorization largely maintained the 1975 bill but made also made a key section conditional on lowering the burden of proof to require showing discriminatory results rather than intent. As such, we can infer that the midpoint on final passage should be to the left of where it had been 7 years earlier.\(^{11}\) Because the 2006
reauthorization did not significantly expand the scope of the act, and from a policy perspective the status quo was generally where it had been in 1982, the midpoint should not change dramatically from its location of 25 years earlier.

Although we cannot make precise predictions about the locations of the expected midpoints, the narrative sketched above predicts they should drift in a liberal manner over time as the status quo is incrementally amended and adopted under changing circumstances. This is not what we find. Figure 5 plots the midpoints of the key votes presented in Figure 3, overlain on the House median, and it reveals that the estimated midpoints are moving in entirely the wrong direction. The roll-call-based estimates implausibly suggest that racial policies in the United States were more racially conservative in 2006 than they were in 1957.

Why are the midpoints estimated as becoming more conservative over time even though the status quos and policy proposals are known to be moving to the left? The discrepancy results from the fact that as the votes became more consensual over time, due to increasingly large majorities of Democrats and Republicans voting for reauthorization, those legislators voting in opposition tended to be estimated as more extreme. To account for the differences in the voting coalitions over time, the statistical model responds by shifting the estimated midpoints in a conservative direction rather than shifting the overall policy space to the left. Rather than assume that most members voting in 2006 have more racially liberal policy preferences than those voting in 1965 and that the policy space has shifted to the left as a result of prior lawmaking activity, the model instead assumes that the underlying issue space is stable—neither drifting nor stretching and contracting over time. As a result, accounting for the changing voting coalitions implies that the midpoints must become increasingly conservative to account for the increasingly consensual voting behavior.

Northeastern University
The implications of this are very hard to reconcile with what we know about the nature of policy change on this issue. It is unlikely that most of the Republican senators who voted with Strom Thurmond in 1990 against the Civil Rights Act of that year—which would have required employers whose practices had been shown to have a disparate impact on the basis of race, color, religion, sex, or national origin to demonstrate that the practices were justified by business necessity—would have also voted with Thurmond against the Civil Rights Acts of the 1950s and 1960s. And yet, because there were no proposals to repeal these earlier acts, such members receive estimates placing them on par with the racial conservatives of the 1950s.

Without accounting for the policies being voted upon, the statistical model is unable to account for instances in which both the issues being considered and the issue preferences of members are changing over time. Moreover, in such instances, the interpretation of the resulting estimates cannot be sensibly compared.

**Accounting for Policy Content**

To account for changing policy content requires incorporating more information into the estimation to help anchor policy change over time (Bailey 2013). Several approaches have been used to do so. Clinton and Meirowitz (2004), for example, propose imposing constraints on the location of the bills being voted upon by equating successful policies with subsequent status quos. This requires strong assumptions about policy equivalence that are hard to sustain on an issue such as civil rights, where the surrounding political and social context is changing so rapidly. Other researchers attempt to model policy change using information such as cosponsorship (Peress 2013) or survey responses (Richman 2011), but such data do not always exist. Rather than imposing assumptions about policy locations, constraints could alternatively be imposed on the estimated midpoints. Bailey (2007, 440), for example, reasons that if a Supreme Court justice voted in one case to allow the execution of those under the age of 16, then it can be logically inferred that this justice would allow the execution of those over 16 in a separate case, even if he or she did not actually serve on the Court that considered the second case; Bailey uses this information to define informative priors for selected cutpoints’ locations.

A third approach, and the approach we take, is to impute votes for particular legislators based on historical understandings and the sequence of policy change. For example, if a member of Congress voted in favor of the 1965 Voting Rights Act, we can infer—all else being equal—that the member would have also likely voted in favor of the 1957 Civil Rights Act because it was a much more modest federal action; if a member voted to reauthorize a much more liberal version of the Voting Rights Act in 1982 or in 2006, then we can infer that he or she would have also likely voted in favor of the original measure. Imputing votes incorporates the information that is implied by the ordering of policy cutpoints without hard-wiring a formal constraint into the statistical model.

We utilize imputation in the text because it requires somewhat weaker assumptions due to the assumption of probabilistic voting.12 In particular, whereas parameter constraints are necessarily informative and impactful, incorporating information using imputed votes involves weaker assumptions because the imputations could be treated as voting errors by the statistical model if they are sufficiently anomalous. As a consequence, using imputed votes rather than cutpoint constraints more easily allows for the possibility of imperfect (or incorrect) information about the policy content being voted upon because of the presence of assumed voting errors in the statistical model. Moreover, it is possible to vary the amount of information being assumed by imputing only a random subset of the implied votes (see Appendix B in the supporting information).13

The logic of imputing votes follows from the spatial voting model and the assumption that if a legislator supports a bill that moves policy to the left from an already left-leaning status quo—such as strengthening the federal role in elections established by the Voting Rights Act—then such behavior is consistent with a policy preference that would have also supported an earlier bill to move policy leftward to the current status quo. Examining the substance of the issues under debate reveals whether proposed policies were intended to move policy to the left or right relative to a status quo set by prior identifiable legislative action.

To do so, we take a limited approach and only impute votes on final passage based on prior final passage votes involving the same fundamental issue. Thus, support for the Voting Rights Act implies support for the Civil Rights Act of 1957, as both acts were fundamentally concerned with voting, but it does not necessarily imply support for

12The supporting information reports similar conclusions from constraining midpoints rather than imputing votes and by relaxing the number of imputed votes by imputing votes to each selected roll call with only a .25 probability.

13As an anonymous referee noted, imputing votes is likely to bias the results toward a finding of more polarization, as it would lead to reduced error in voting, which in turn would produce ever more extreme estimates. This would bias the results in the contrary direction, so it does not invalidate our results. However, it does suggest caution against an overzealous application of this method.
the Civil Rights Act of 1964, which was more concerned with prohibiting discrimination in commerce.\textsuperscript{14} We also use only more radical policy proposals to infer votes on earlier ones. Opposition to the Voting Rights Act of 1965, for example, does not necessarily imply opposition to the Civil Rights Act of 1957, but opposition to the Civil Rights Act of 1957 does imply opposition to the Voting Rights Act of 1965. These linkages can involve policies spanning long time periods—supporting the 1965 Voting Rights Act, for example, implies supporting the Lodge Force Bill of 1890.

To be clear, the results employ the same statistical models used in existing work and assume that the estimated ideal points reflect policy preferences despite the theoretical ambiguities involved. What distinguishes the analyses of this section from the prior section is the fact that the roll-call matrix that we are analyzing has been augmented to contain imputed votes based on the changing policy content and a spatial voting model.

Estimates generated using the augmented matrix of roll calls produces dramatically different conclusions about the nature of elite conflict over time.\textsuperscript{15} Figure 6 graphs the median ideal points of Democrats (upper left), Republicans (upper right), and Southern Democrats (lower left) and the distance between the median Democrat and median Republican (lower right), with associated 95\% confidence intervals.

As Figure 6 reveals, accounting for policy content has relatively little effect among Democrats. The only noticeable impact is that the trend using imputed votes (dashed line) shifts the estimated ideal points of Democrats before the 1930s in a more racially conservative direction relative to the estimates that ignore policy content (solid line). Changes in the average ideal points of Democrats (upper left) and Southern Democrats (lower left) reflect the changing composition of the Democratic Party over time: Whereas the racial conservatism of the average Democrat shifted dramatically during the New Deal to a more liberal position following World War II, the average Southern Democrat remained racially conservative until the early 1970s.

The most dramatic changes occur among Republicans graphed in the upper right of Figure 6. Without accounting for policy content, the solid line suggests that the average Republican in 1877 was more racially liberal than they would ever be and that they were also (implausibly) more racially liberal than the average Democrat in the 21st century. Moreover, while the position of the average Republican is estimated to drift more conservatively over time, it is hard to believe that the average Republican in 2009 is as racially conservative as the average Southern Democrat in 1877, as the estimates suggest.

Incorporating policy content changes these characterizations considerably. As the dashed line in the upper-right plot of Figure 6 reveals, accounting for policy content sensibly suggests not only that the median Republican in 1877 was not more racially liberal than the median Democrat in 2009, but also that the median Republican of 2009 was not as racially conservative as the median Southern Democrat in the 19th century.

Accounting for the content of votes also reveals a liberal shift in the policy space over time—the conflict between contemporary Democrats and Republicans occurs entirely within what was the liberal half of the policy space following Reconstruction. Whereas the party medians ranged roughly from −0.5 to 0.75 in 1877, by 2009 they ranged only from −0.5 to 0.0 and they occupy only 40\% of the earlier space; this shift has important implications for the estimated distance between party medians (lower right of Figure 6).\textsuperscript{16} Ignoring policy content (solid line) suggests that contemporary differences in party preferences are equivalent to the policy differences of the 19th century, but including policy-specific information (dashed line) reveals that contemporary differences are considerably less than they were in the post-Reconstruction period, even with the estimated increase since the 1970s.

The additional information has improved our confidence in the interpretation of ideal points as reflecting changes in policy preferences, but questions can still be raised about the temporal comparability given the patterns revealed in Figure 6. For example, how confident can we really be in our comparisons of 1990 to 1890, or of the scores in the early 20th century? Such questions highlight an inherent limitation in using behavioral measures to characterize unknown policy preferences. The characterizations depend not only on the willingness of elites to take action, but also upon our ability to interpret and relate the meaning of those actions. Unfortunately, knowing how broadly comparisons can be made is not something that is statistically discernible. Instead, judgment is required to determine the extent to which comparisons can reasonably be made. While this is unfortunate from the perspective of providing ready-made,

\textsuperscript{14}The supporting information provides a detailed discussion of the constraints. Out of 4,166 instances where a member voted on multiple constrained roll calls, the votes cast were contrary to imputations in only 358 instances. Much of this reflects preference change among long-sitting members.

\textsuperscript{15}Appendix Table A1 in the supporting information contains model fit information.

\textsuperscript{16}The estimates are set to the same scale by measuring polarization as a fraction of the distance in party medians in 1877.
Figure 6  Estimates of House Party Medians Using Civil Rights Votes, 1877–2007

Note: Gray lines in bottom right indicate where more than 10% of members had votes imputed.
all-purpose estimates, the fact that the temporal comparability is contestable and unclear has the benefit of forcing a more careful consideration of the extent to which available measures are sensibly interpreted as reflecting policy change and the potential limits of so doing.

When there are few votes, or if the votes are difficult to relate to past or future legislative activity, it is hard to account for policy content. For example, because the New Deal Democrats purposely kept issues involving race off the congressional agenda as much as possible to prevent their legislative coalition from fracturing (Katznelson and Mulroy 2012), we have relatively little information with which to estimate or adjust elite policy preferences for this period. The same is true of the first decades of the 20th century when neither party showed much interest in civil rights. The fact that the two time series are often similar is attributable to the endogenous agenda being analyzed and the consequential inability to relate votes across time.

Figure 7 reveals why the results diverge most in the more recent periods by highlighting the relationship between the roll-call agenda and the patterns in Figure 6. The figure graphs the average percentage of a member’s total votes in each group that we impute. For instance, Duncan Hunter (R-CA) cast 72 votes on civil rights bills or amendments during his 28 years in Congress (1981–2009), and an additional 22 votes (23%) can be imputed based on his voting record. Because votes with the clearest connections to one another occurred in the post–civil rights period, the impact of our approach is most noticeable when estimating the behavior of recent members.

The Case of Social Security in the House

This pattern is not specific to civil rights issues. There are many issues where party positions have changed over time, as has occurred to varying degrees on civil rights, free trade, and abortion, but where the partisan character of voting has not (Karol 2009). Even when parties have not swapped positions, accounting for policy
content can affect our characterization. Votes on Social Security, for example, reveal a similar but less dramatic effect of the importance of accounting for the relationship between the policies being voted upon. Figure 8 traces the party medians and the distance between these from 1935 to 2006 in the U.S. House, using estimates generated with and without taking policy content into account. Both sets of estimates show a depolarization in the 1940s and 1950s, followed by a steady increase after the 1960s. However, the relative level of contemporary differences is overstated if policy content is not taken into account.

Incorporating policy content produces a pattern that appears to fit well with the historical trajectory of Social Security. Although many Republicans initially opposed the program, and a majority of the party supported an effort to strike the pension system from the bill, final passage in 1935 was characterized by bipartisan support, with an equal number of Democrats and Republicans in opposition. Within only a few years, even this opposition had collapsed: A substantial expansion of the pension program passed 361–2 in 1939, the 1948 Republican platform called for a further increase in benefits, and in 1950 a large expansion of the program passed with only three nay votes (“Social Security Act” 1951). By the late 1970s, the program was under financial stress and votes were more contentious; the average vote margin declined from a peak of 91% in the 1960s to 62% between 1980 and 2004. While policy differences between the parties have increased, it seems unlikely that the Republican majority elected in 1994 would have voted to strike the pension program entirely, as a majority of their co-partisans did in 1935. This is what is suggested by estimates that fail to account for policy content.

**Conclusion**

Studying the correlates and consequences of political conflict in the legislative arena touches on questions that are

---

17 Changes in the early 1990s focused “mainly on ‘fine-tuning’ the program” (Kollmann 1996, i).
fundamental to characterizing and evaluating democratic institutions and performance. Existing studies into such questions nearly always use analyses of roll-call voting behavior to characterize policy preferences, but the reasonability of equating ideal points and policy preferences is often assumed rather than interrogated. Given the critical importance of policy preferences in both theoretical and empirical work, the lack of any systematic examination into the ability of commonly deployed measures to characterize policy change is unfortunate and consequential, especially when we consider the number of studies relying on such characterizations.

A virtue of extant measures is that they allow scholars to fit statistical models consistent with behavioral voting models to observable actions without engaging in the difficult and time-consuming task of determining what is being voted upon and why. We show that this virtue can have negative consequences; remaining agnostic as to the content of the congressional agenda and the meaning of the estimated parameters has important consequences that are often not fully appreciated: Because the estimated policy space(s) are defined primarily by stable voting coalitions, rather than policy content, it is difficult to interpret the recovered dimensions and ideal points in terms of policy preferences over time.

Examining conflict over civil rights in the United States clearly reveals the shortcomings of using roll-call-based estimates to characterize policy preferences over time. Commonly used estimates profoundly and dramatically mischaracterize the magnitude of policy change in consequential ways. While the parties in recent decades are more likely to vote against one another on race-related issues, and the two parties do indeed differ in their most preferred policies, it is implausible that the differences separating racial liberals from conservatives are as great today as they were in the late 19th century, or that there has not been a substantial leftward shift in the policy space since the 1960s. Indeed, the aspiration of the civil rights movement was to secure durable changes in public policy that would positively transform American life (Valelly 2005). As Georgia representative John Lewis noted in urging passage of the 2006 Voting Rights Act reauthorization, “Yes, we’ve made some progress; we have come a distance,” but “the sad truth is, discrimination still exists.”

We might still be a house divided, but the divisions are smaller than they were.

Methods and measures based on roll-call votes that ignore the changing congressional agenda not only fail to detect this progress but also suggest a misleading ideological equivalence between contemporary conservatives and those who opposed any federal action to protect black civil rights. Beyond diminishing the accomplishments of the civil rights movement, this also obscures our understanding of contemporary politics.

Focusing on the case of civil rights provides a clear illustration of the difficulty in using existing methods and estimates to characterize changing policy preferences, but the concerns we identify are not limited to the issue of civil rights. Using roll-call-based measures to characterize policy preferences over time will be exceptionally difficult wherever there has been a systematic shift in member preferences or the policy space, possibly wherever a policy creates a constituency large or powerful enough to make fundamental changes to the program an unpopular position.

Our analyses provide an important corrective in terms of both the characterization of the level of political conflict related to civil rights and the way in which the content of policy is accounted for in estimators of elite behavior, but our solution is necessarily a partial remedy. Even though our characterization better reflects the historical understanding of how policy and preferences have changed over time, it does not account for the fact that members themselves are deciding which votes to take. The investigation also follows conventional usage in assuming that the ideal points reflect policy preferences, despite the fact that many other behavioral models could also explain the observed voting patterns and that nothing ensures the ideal points reflect sincere policy preferences. There are inherent limits in our ability to characterize individual preferences on the basis of observed behavior. But by incorporating more information, at least some of these obstacles can be mitigated.

References


Brady, David, Hahrie Han, and Jeremy Pope. 2007. “Primary Elections and Candidate Ideology: Out of Step with the
Primary Electorate?" Legislative Studies Quarterly 32(1): 79–106.


---

**Supporting Information**

Additional Supporting Information may be found in the online version of this article at the publisher’s website:

- **Appendix A:** Measuring Polarization
- **Appendix B:** Constrained Midpoints and Probabilistic Imputation
- **Appendix C:** Race and the “Second Dimension”?
- **Appendix D:** Policy-Induced Voting Imputations
- **Appendix E:** Predicted Probability of Imputed Ideal Point Estimates
- **Appendix F:** Civil Rights and the Senate
- **Appendix G:** Possible Issue Areas for Future Exploration
- **Appendix H:** Members With Contrary Imputations