Agreement Scores, Ideal Points, and Legislative Polarization

Betsy Sinclair  
*University of Chicago*

Jennifer Victor  
*University of Pittsburgh*

Seth Masket  
*University of Denver*

Gregory Koger  
*University of Miami*

Correspondence should be directed to betsy@uchicago.edu

Abstract

A number of scholars have described a steady increase in ideological polarization between Democrats and Republicans in the U.S. Congress. This research often relies upon ideal point estimation methods that assume legislators are atomistic individuals registering a sincere, independent preference on each roll call vote. Yet scholars who employ social network techniques, such as estimating each legislator’s centrality score, typically assume that legislators’ voting decisions are a consequence of social influences. This paper examines the distinctions between these two approaches. We introduce the agreement score, which measures the extent to which pairs of legislators vote the same way. We demonstrate the relationship between both agreement scores and centrality scores, a common measurement of influence in social network analysis, as well as between centrality scores and ideal point estimates. We then focus the ways in which agreement scores can be used to illuminate theories of legislative polarization for members of the U.S. House from 1967 to 2008. Agreement scores are not only more conducive to the assumptions typically made in social network analysis but also the empirical patterns of these scores are consistent with major explanations of legislative partisanship.
Introduction

For over a century, scholars have sought to understand legislative behavior by analyzing roll call votes. We understand legislation better by studying who votes for it; we understand legislators better by analyzing what they vote for and who else voted the same way. Measures of legislators’ similarity underlie our research on political parties, presidential influence, committees, and entrepreneurship. This article describes a classic measure of similarity known as the agreement score and uses social network analysis measures to illustrate the sources of political polarization from 1967 to 2008.

Currently, the most prevalent method for describing the preferences of legislators and the (dis)similarity between pairs or groups of legislators is some form of item-response theory (IRT) model, especially the family of NOMINATE scores developed by Keith Poole and Howard Rosenthal (Poole 1998, 2005; Poole and Rosenthal 1985, 1997; Poole, Rosenthal, and McCarty 1997, 2006). This approach offered a marked improvement over previous measures of legislators’ ideology (e.g., interest group ratings) and has contributed immensely to the study of the U.S. Congress. This approach is based on a set of strong assumptions that posit legislators as autonomous, fully-informed actors who vote for the option “closest” to their ideal policy. Yet some types of legislative behavior are best explained using alternative assumptions, e.g., that parties are influential because legislators are not fully informed and the perceived content of legislation can be framed by legislative leaders (Lee 2009, 46; Sinclair 2002). In particular, research that stresses the importance of informative signaling (e.g., Kingdon 1989) or social ties between legislators (e.g., Fowler 2006a, 2006b; Porter et al. 2007; Victor and Ringe 2009) is more compatible with measures of legislator affinity that are not based on fully informed, autonomous actors.
As a complementary alternative to IRT measures, we describe a classic measure of legislator similarity: agreement scores. We combine these scores with the methods and measures of social network analysis to illustrate how they can be used to trace and explain the partisan polarization of Congress. This network approach offers new evidence in support of extant claims about the sources of legislative polarization, including chronological, cultural, and institutional influences. Agreement scores provide a simple and intuitive measure of legislative cohesion while avoiding the strong, problematic assumptions underlying IRT-based measures. The next section provides historical context for agreement scores, then we describe how the scores are calculated, how they can be combined with social network analysis (SNA) techniques and contrast with ideal point estimation techniques, and then finally we apply our approach to polarization.

**Studying Congressional Votes**

For over a century, scholars have analyzed legislative voting to gain substantive insights into a complex process. The earliest studies by A. Lawrence Lowell (1902) and Stuart Rice (1925) focused on the cohesion within legislative parties and the differences between parties. In the 1940s, Congressional Quarterly began summarizing the vote scores for individual members of Congress based on their party loyalty on “party votes,” *i.e.*, votes on which most Republicans voted against most Democrats. Over time, Congressional Quarterly would also tabulate legislators’ support for the President and for the “conservative coalition” of Republicans and Southern Democrats. Each of these statistics was based on legislators’ support or opposition to an exogenously identified group but, as Rice explains, the underlying idea is that we can generally assume that the more legislators vote together, the more similar they are.
In an effort to develop a more nuanced understanding of Congressional parties, factions, and social relationships between legislators, David Truman (1959) pioneered a new approach based on the “agreement” between the voting patterns of every pair of legislators. Once he had calculated the agreement score for each pair of legislators, he identified clusters of legislators who voted together during the 81st Congress. This was very laborious, of course, in the age of punch cards and reserving time on the university’s “computer.” Not surprisingly, then, several subsequent analyses focused on party cohesion across issues, controlling for constituency (e.g. Mayhew 1966; Sinclair 1976, 1977). Others, however, sought to verify and extend Truman’s interest in the relationships between legislators, including the role of signaling and cue taking (Buchanan 1963; Kingdon 1989[1973]; Matthews and Stimson 1975).

The study of Congressional voting was revolutionized by the NOMINATE scaling methods developed by Keith Poole and Howard Rosenthal (1985, 1997). Poole and Rosenthal sought to identify the ideological alignment of legislators by treating Congressional votes as expressions of their preferences (with errors) given the assumptions of the spatial voting model (Poole 1999). That is, the estimation technique assumes that legislators have ideal points on each policy dimension (in practice, one or two dimensions) and vote to minimize the distance between their preferred policy and their ideal points. In the NOMINATE framework, legislators who vote together often do so because they have similar ideal points. Poole and Rosenthal have generated and shared a variety of NOMINATE-based scores, providing a convenient measure of individual and collective ideology over the scope of Congressional history, e.g. demonstrating the emergence of partisan polarization in Congress (McCarty, Poole, and Rosenthal 2006; Theriault 2008).
The NOMINATE revolution largely supplanted the study of cue-taking and social influence on Congressional voting, instead popularizing the premises of the spatial model and making it easier for scholars to test models based on this approach. This includes the assumption that legislators derive utility from the outcome of votes, rather than receiving electoral praise or blame for the positions they take per se. Second, legislators presumably vote for the option “closest” to their ideal points, even if both options are very distant. Thus, a legislator who is a socialist would vote for a bill that raises the minimum wage from $.01/hour to $.02/hour, even if the legislator considers the bill insultingly inadequate. Finally, legislators presumably treat each vote like a single-shot game, whereas legislators who expect that their decisions may influence the future options and outcomes of a game may vote differently (Penn 2009). The assumptions of the spatial model, especially the opaque origins of legislators’ preferences, make it difficult to discern the influence of constituents and parties (Krehbiel 1993, 1999; Sinclair 2002).

Some scholars have criticized the NOMINATE algorithms for failing to estimate the full range of possible coalitions, reducing the dimensionality of estimated bill and legislator positions (Londregan 1999). Instead, Londregan recommends exogenously estimating the nature of the legislative agenda. In this vein, Clinton, Jackman, and Rivers (2004) advocate a Bayesian estimation technique that allows scholars to incorporate additional information about the voting process (e.g. party whipping or agenda-setting) and estimate the uncertainty around ideal points. Both these critiques share NOMINATE’s underlying framework: rather than questioning the enterprise, the authors propose improved techniques to apply the spatial model to voting.

These NOMINATE/Item-Response Theory (IRT) approaches represent the best means available for scholars to achieve a very specific task – estimating the ideological predilections of legislators using observable public information. The information derived from it can be applied
to a great many avenues of legislative study, including legislative polarization and dyadic representation.

However, the focus on IRT models has led scholars to largely ignore another important aspect of legislative behavior: the degree to which legislators influence one another. At the most basic level, we know that logrolling occurs with great frequency, resulting in members voting against their own preferences in exchange for the promise by another legislator to vote their way on some future bill (e.g., Calvert and Fenno 1994). More importantly, we have numerous examples of members being influenced quite strongly on votes by things other than their own ideological predispositions, constituencies, and party affiliation. For example, Young (1966) offered evidence that members in the early American Congress were strongly influenced by their roommates in their boarding houses (although see Bogue and Marlaire 1975). Matthews and Stimson (1975) and Kingdon (1973) found evidence of members of Congress following voting cues not only of legislative leaders and committee chairs, but also of friends and respected colleagues. More recently, Masket (2008) found members of the California Assembly following the votes of those sitting near them on the chamber floor. We know from other research that legislators may be influenced by cosponsorship cues (Koger 2003; Fowler 2006) or by fellow members in legislative caucuses (Victor and Ringe 2009).

For scholars interested in the formation of legislators’ preferences, including signaling, party deliberations, and other forms of social interaction, it is necessary to use measures that are compatible with their research framework. For such research, we advocate the use of agreement scores as pioneered by David Truman. These scores require minimal assumptions and are amenable to studying the relations between legislators and groups of legislators.

**Agreement Scores**
We use legislative roll call votes to calculate agreement scores for each dyad of legislators. We relax the assumption that legislators cast independent votes on each roll call bill based upon their respective distance between the yea and nay locations and their ideal point. Instead, we consider the roll call votes from each legislative session to be the result of a set of ideologies, strategic interactions, and cue-taking behavior. This leads us to focus upon the correlation of voting behavior between each legislator-pair and is referred to as an *agreement score*. Agreement scores focus attention on the relationships between legislators as the quantity of interest.

We conceive of the legislator relationships as an M-by-M matrix, where M is the number of legislators in a Congress. The values in the matrices we produce help describe relationships between all pairs of legislators by calculating the rate that any two legislators voted the same way, given that they both cast a vote. More specifically, we use the roll call votes to define an implicit network between legislators by counting, for each individual $i$ and $j$, the number of instances where $i$ and $j$ have agreed upon a particular bill. This produces an adjacency matrix, $A$, where each entry in the matrix describes the rate of agreement between $i$ and $j$. That is, each entry is the number of times each pair of legislators $i$ and $j$ have voted similarly on all roll call votes divided by all possible eligible opportunities to do so. Each entry, $a_{ij}$, can be thought to represent behavioral proximity between legislators $i$ and $j$. We refer to the entries in the adjacency matrix as agreement scores. This representation is an improvement on existing literature in that it captures the instances where $i$ and $j$ behave similarly while also providing a measurement for all other interactions of $i$ with other legislators $k$ and similarly for all other interactions of $j$ with other legislators $k$. By locating the behavior of these two legislators $i$ and $j$ within the context of their behavior in the legislature as a whole, we are better able to identify
patterns of legislative prowess, for example. Each individual legislator will have M-1 agreement scores, where each score represents agreement with a particular other legislator. The overall pattern of agreement scores reveals some characteristics about that legislator. A legislator with a high average agreement score, for example, might be a legislator like mid-20th century House Speaker Sam Rayburn, who was able to vote similarly both with members of his own party as well as with members of the other party. Agreement scores illustrate the relationship between a particular legislator and the rest of the chamber. Agreement scores make minimal assumptions about legislative behavior. In particular, they require no assumptions about independence and sincerity. The primary disadvantage of agreement scores, however, is that the agreement score is calculated by treating all bills equally, regardless of whether the bills proposed are extreme deviations from the status quo.

**Ideal Points, Agreement Scores and Centrality Scores**

We next explore the differences between agreement scores and ideological scaling measures like NOMINATE. We illustrate using a classic example of five Senators (called Nunn, Helms, Gore, Kerry and Dole) who vote across four issues (called B-2, Cambodia, Tower, and MLK).\(^1\) Legislators’ votes on these issues are described in an incidence matrix, such as in Table 1. Here, a “yea” vote is indicated with a 1 and a “nay” vote is indicated with a 0. Each

---

\(^{1}\) This example is borrowed from Poole and Rosenthal 1997. NOMINATE procedures are typically intended for large-n examples, while the Bayesian IRT methods are more appropriate for smaller legislatures, so for the purposes of our example, we estimate ideal points with IRT for these legislators (Clinton, Jackman and Rivers 2004). We do so only to demonstrate that characterizations of legislators under the assumptions of spatial voting may yield different results than characterizations of legislators under the assumptions of relational voting, and we would encourage readers to have caution when estimating ideal points on a small number of votes for a small number of legislators.
legislator’s ideal point is assumed to be close to the “yea” outcome if he voted “yea” and close to the “nay” outcome if he voted nay. That is, both legislators and roll call bills have a spatial location.

Table 1

For purpose of example, we focus on a one-dimensional policy space. Then the goal of this estimation is to produce an ideal point \( x_i \) for each legislator \( i \). Let \( N \) represent the number of legislators in each session. We observe \( M \) different roll call votes for each individual \( i \in \{1, \ldots, N\} \). Each bill \( m \in \{1, \ldots, M\} \) has a yea location \( \varsigma_m \) and nay location \( \gamma_m \) in a one-dimensional policy space. Legislators are assumed to have quadratic utility functions over the policy space, so that

\[
U_i(\varsigma_m) = -||x_i - \varsigma_m||^2 + \eta_{im} \quad \text{and} \quad U_i(\gamma_m) = -||x_i - \gamma_m||^2 + \nu_{im}
\]

where \( x_i \in \mathbb{R} \) is the ideal point of legislator \( i \) and \( \eta_{im} \) and \( \nu_{im} \) are errors which are normally distributed and independent across both legislators and roll calls. We assume that each legislator maximizes her utility for each bill, taking into consideration only her own ideal point and the yea and nay locations of the bill. Then, legislator \( i \) votes yea (so that vote \( y_{im} = 1 \)) if \( U_i(\varsigma_m) > U_i(\gamma_m) \) and nay otherwise (so that \( y_{im} = 0 \)). This framework is referred to as the spatial voting model.

Identification of the ideal points of legislators in this model requires estimation of the cut point between the yea and nay bill locations as well as an ideal point for each legislator. This is typically accomplished either via Bayesian MCMC algorithms, W-NOMINATE, or factor analysis (Clinton, Jackman and Rivers 2004; Poole and Rosenthal 1997; Heckman and Snyder 1997).

Using the five Senators’ votes over four bills in Table 1, we can calculate ideal points of this very small legislature. Table 2 presents these scores. The scores for our hypothetical
congress show that these legislators have ideologies ranging from -0.12 (Helms) to 0.21 (Kerry). The NOMINATE procedure produces a logical, spatial, one-dimensional score for each legislator that is comparable across legislators. These points could easily be arranged on a one-dimensional spatial model. The appeal of such scores is obvious and it is easy to see why so many scholars (the authors of this paper not excluded) have relied upon NOMINATE scores as proxies for the ideology of legislators for some time. However, in order to use these scores one must assume that there is no consequential social interaction between members, so if these senators met as parties (which they did) or if, say, the other senators asked Nunn his opinion on the B-2 bomber and this interaction biased their information or independently influenced their votes, then one cannot use these scores without reservation.

Table 2

An alternative strategy is to estimate agreement scores. When two legislators vote similarly on a particular bill, they have expressed a level of agreement between them that is both a function of ideological similarity as well as their personal interactions. That is, we retain the assumptions that each legislator has quadratic utility over the policy space but adjust the components of the utility function to incorporate interactions between legislators so that \( U_i (\zeta_m) = -||x_i - \zeta_m||^2 + \sum \eta_{ijm} \) and \( U_i (\gamma_m) = -||x_i - \gamma_m||^2 + \sum \nu_{ijm} \forall j \neq i \). An individual’s utility for each bill is now determined by not only their respective distance from the yea and nay positions of the bill and their ideal point but also incorporates the behavior of the other legislators \( j \) into the error terms \( \eta_{ijm} \) and \( \nu_{ijm} \). This is problematic for identification using the standard methods for estimating ideal points. Thus, instead of estimating each of these parameters, we calculate the

2 Note that these values are not robust to a 180-degree switch (that is, there is no substantive significance associated with which values are positive or negative, only the ordering is relevant here).
correlation between any two legislators’ voting behavior. The primary disadvantage of agreement scores is that they treat each vote equally, making estimation of the bill cut point location impossible. Not every bill should carry equal importance yet agreement scores do not account for the specifics of each distinct bill.

To demonstrate agreement scores, consider the above example of five senators and four votes and refer to the series of tables below. First imagine the senators in a typical incidence matrix where rows are legislators and columns are votes (Table 1). We can then calculate the number of votes on which both $i$ and $j$ agree given that they both voted by summing their votes and dividing by the number of opportunities they had to vote together. Table 3 shows the matrix of agreement scores for our simulated actor network.

Table 3

Where ideal point estimation tells us that the “correct” ordering of these legislators from left-to-right is Helms, Dole, Nunn, Gore, Kerry, and produces an ideal point for each legislator, the agreement procedure does not produce a single ideal point for each individual. Rather, we have produced a matrix that describes the extent to which legislators agree with other legislators. Agreement scores quantify voting patterns between legislators within a legislative session.

Agreement scores have an additional advantage. Since all items on the agenda are treated equally, it is possible to compare agreement scores across sessions. Ideal point estimation has typically required that either some legislators or other institutions, such as interest groups, have

---

3 The reader will notice that these scores are not particularly representative of the ideal points of these actual senators. Of course, the votes on which they are based are, for the purposes of this example, pure fiction. However, this also points to a limitation of the NOMINATE procedure – it tends to break down in small legislatures that cast few votes. Estimated ideal points in these conditions are unstable and unreliable.
fixed ideologies over time. This enables these legislators or institutions to serve as bridges, which results in the type of ideal points produced in DW-NOMINATE scores. This process forces ideal points to the same scale, such as with Common Space Scores (Poole 1998, 2005). Yet, if we focus our substantive question on polarization within the legislature over time, then we cannot evaluate the extent to which the chamber has shifted if we require the same legislators to have fixed ideal points. With agreement scores, it is possible to directly compare legislative agreement over time. By quantifying legislative behavior in this way, it is now possible to evaluate if legislators, for example, are less likely to agree with members of a different party in the 110th Congress than they were in the 90th Congress.

While ideal points are commonly used in the political science literature, agreement scores are not. To a large extent this is attributable to the fact that agreement scores do not distinguish between bills. Yet a growing body of political science literature has been working to document the ways in which legislators’ social interactions impact their voting decisions. These scholars have typically relied upon centrality scores, a common technique in social network analysis. We next explain the relationship between centrality scores and ideal point estimation. By examining this relationship, the advantages presented by agreement scores for scholars wanting to incorporate social network analysis methods into studies of legislative behavior become clear.

Centrality is a concept derived from graph theory that seeks to describe the most “important” node in a graph. Each individual in the graph is represented by a node; connections between individuals—here, voting in concert with another individual—is represented by an undirected link. An individual's importance, then, is calculated as a function of these links in the
Centrality measurements have been used to ascertain legislative influence when it is not possible to directly observe network connections but instead to rely upon implicit influences, such as co-voting or co-sponsoring (Fowler 2005; Porter 2005; Fowler 2006). Legislators with similar centrality scores tend to also have other social ties, such as sharing leadership positions or holding similar previous offices (Fowler 2005).

If we were to directly compare NOMINATE scores to their closest parallel in the social networks literature, we would employ a definition of centrality defined as Bonacich power centrality (Bonacich 1987), similar to alpha centrality as defined by Bonacich and Lloyd (2001). Bonacich power centrality is a particularly appealing definition for legislative vote data because of three key characteristics. First, it assigns each individual within the network some small initial importance. Without the assumption that each individual begins with an initial importance, it is possible that an individual who had no one with whom they voted similarly then they would be assigned zero importance. The second feature of this definition of centrality that is appealing is that it assigns centrality values to each individual based on the principle that connections, even indirect connections, to more highly-connected individuals should carry more weight than connections to lower-connected individuals. In terms of evaluating legislative connectivity, the intuition is that a legislator is assigned a high value if he or she votes similarly with a large number of other legislators, and that this legislator’s value should be made even larger still, if the individuals he or she votes with are also highly connected. Thus the most

---

4 For a detailed introduction to network analysis, see Wasserman and Faust 1994.

5 This notion of centrality is very similar to that of eigenvector centrality, with the difference being that each individual is given some initial exogenous value in the network regardless of the links that form between individuals.
central legislator could be thought of as being representative of the chamber--they are the legislator with whom the most legislators agree, and with whom similarly central legislators agree.\(^6\) Finally, we appreciate the feature that this measurement of centrality can be normalized across legislative sessions, as the amount of influence is fixed for each legislative session.

More formally, let \(N\) represent the number of legislators within a particular session and let \(A\) be an \(n \times n\) adjacency matrix, where each value \(a_{ij}\) describes the intensity of the relationship between individuals \(i\) and \(j\). For any given pair of legislators \(i\) and \(j\), the value \(a_{ij}\) is the number of times that legislator \(i\) has cast the same vote as legislator \(j\) in that particular legislative session.

Given the adjacency (or agreement) matrix, \(A\), presented above, the vector of centrality measures \(c(\alpha, \beta)\) is defined as:

\[
c(\alpha, \beta) = \alpha(\mathbf{I} - \beta \mathbf{A})^{-1} \mathbf{1}
\]

for \(i = 1, \ldots, N\) legislators; \(\mathbf{I}\) is an \(n \times n\) identity matrix; \(\alpha\) is a scalar; \(\beta\) is an \(n \times 1\) vector of initial individual weights; and \(\mathbf{1}\) is an \(n \times 1\) vector of ones (Bonacich 1987; Bonacich and Lloyd 2001).

The scalar \(\alpha\) is chosen such that \(\sum_{i=1}^{N} c_i(\alpha, \beta)^2 = N\). The scale factor \(\beta\) is chosen to be “appropriately small.”\(^7\) Here, \(\beta = \frac{1}{(1 + \lambda_1)}\), where \(\lambda_1\) is the largest eigenvalue of \(\mathbf{A}\).

---

\(^6\) There are a large number of centrality measurements, and the choice of the particular centrality measurement is dictated by the substantive problem, as described by Freeman (1977). Other methods to compare networks include options to compare the Gini coefficients for different degree distributions, as in Goyal, van der Leij and Morago-Gonzalez (2006).

\(^7\) Here “appropriately small” is defined to be less than the inverse of the largest eigenvalue of the adjacency matrix, \(\mathbf{A}\). See Bonacich (1987) for further details.
Table 4 shows the power centrality scores of our example data. Here we show that Nunn is the most central (most agreeable) in the network, while Helms and Kerry are the least central (least agreeable). Using this approach, we have made no assumptions about legislators’ orientation toward one another or the independence of their voting. However, unlike the agreement scores, it becomes more difficult to discern the polarization of the network, legislators’ relationships to one another, and their relative positions in the network. If we compare the estimates in Table 2 to those in Table 4, one ordering of legislators in terms of ideological extremism (remembering that the scale is irrelevant) is Kerry-Helms-Dole-Nunn-Gore. Ordering the centrality scores generates Helms-Kerry-Gore-Dole-Nunn – that is, centrality captures something other than ideological distribution and is unlikely to be an appropriate tool to identify the presence or absence of coalition-building legislators like Sam Rayburn. Centrality, like agreement scores, has the disadvantage of treating all bills equally. Yet agreement scores are capable of capturing the essentials of polarization trends. While centrality scores are becoming frequent tools to look at legislative networks, it is important for scholars to think carefully about the underlying models of legislative behavior that their quantitative tools rely upon, as well as the primary motivation or question for quantifying patterns in roll call votes. For investigations into polarization or other questions designed to identify legislators like Sam Rayburn, we believe agreement scores are well-suited to illustrate these patterns.

Table 4

We graph our simulated actor network to demonstrate these relationships. Figure 1 is a network graph of our sample network. The lines (or edges) between actors (or nodes) are thicker for actors that have higher agreement scores. The size of each node indicates each actor’s
centrality score. The placement of nodes on the graph represents actors’ relative closeness to one another. This graph shows the strong relationships between Nunn and Dole, or Nunn and Gore, for example, and the weak relationship between Kerry and Dole. It also emphasizes the central position of Nunn in the network. Centrality analyses in social networks is often used to illustrate the role of an individual in the context of a broader set of relationships. Centrality analysis has a clear parallel in political science research – that of ideal point estimation – although the distinction is that in social network research, the focus is on understanding the ways in which one legislator’s behavior relates to the others. In more traditional political science research, the focus instead has been on isolating a legislator’s behavior to understand her particular ideological preferences.

Figure 1 Goes Here

Another clear difference between the ideal point estimation approach and the agreement score approach is that the former is an estimate of a latent trait while the latter is a description of existing data. Social scientists are primarily interested in testing falsifiable hypotheses and are therefore often interested in knowing the standard error or confidence interval around any quantitative metric. Agreement scores are not estimates of anything; they are simply summary descriptions of some quantity of data. However, they provide a very clear, and methodologically simple, measure of legislative behavior that enables us to study the relationships between individual members. Understanding what drives the reports of legislative polarization over time is a difficult problem, and one that has been primarily addressed with ideal point estimation techniques. In the next section, we illustrate the advantage of agreement scores and provide some leverage on the fundamental root causes of legislative polarization.

Application: Political Polarization
There is a general consensus that legislative partisanship has increased in the U.S. Congress over the past few decades. This phenomenon is seen as having multiple possible causes that fall into three general themes: mass secular partisan polarization, institutional conditions, and party cultural differences. In this section, we demonstrate how the agreement scores (and centrality) method can be used to evaluate each of these three causes. Our analysis shows some support for each of the extant hypotheses about the sources of polarization. The application demonstrates the ease with which these methods can be applied to illustrate the ways in which agreement scores and centrality scores follow consistent patterns with existing theories of polarization. Moreover, the extant evidence for the various theories about the sources of polarization relies on spatial analysis, such as data provided by NOMINATE, to support its claims. Here, we find evidence for the sources of polarization using a far simpler approach—agreement scores.

The first main cause of congressional polarization may be described as mass partisanship. This term encompasses the totality of partisan forces outside the legislature, whether induced by activists, party elites, or the voters themselves. That is, regardless of what legislators or legislative leaders may desire, voters and others may simply be sending more ideologically coherent party caucuses to Washington than they used to. This is a result of several trends, including a steady, secular movement by southern white voters away from the Democratic Party and toward the Republican Party since the 1960s. Generally speaking, more conservative voters have increasingly identified with the Republican Party while more liberal voters are increasingly calling themselves Democrats (Brewer and Stonecash 2007; McCarty, Poole and Rosenthal 2008). The existing evidence for this explanation is primarily based on NOMINATE-style examination of roll call data.
If increasing mass partisanship is, in fact, what is inducing congressional polarization, we should see evidence of it captured in the agreement scores of members of Congress. That is, we should be able to detect a slow but steady drift by the party caucuses away from each other as voters increasingly sort themselves into ideologically homogenous districts (Bishop and Cushing 2008). More specifically, we should observe the following trends in our data:

1.1 First, the mean agreement among same-party members should be increasing steadily over time. This is consistent with greater intraparty cohesion.

1.2 Second, the mean agreement among opposite-party pairs should decline as those parties become ideologically distinct from each other.

The second aspect of congressional polarization on which we focus concerns institutional forces. More specifically, it is the result of majority party leaders compelling their rank and file members to vote the party line based on the party’s need to prevail on roll calls. When the majority holds only a narrow seat advantage over the minority, this pressure will be strong, as even a few defections can result in significant policy losses for the majority. Conversely, when the majority party is very large, this pressure will be relaxed, as more members will be free to vote against the majority without endangering the majority’s policy victories. This argument is well advanced by Lebo, McGlynn, and Koger (2007) and Patty (2008).

If the size of the majority party is related to legislative polarization, we should see this manifested in several ways in congressional agreement scores:

2.1 First, as the size of the majority party increases, mean agreement among same-party pairs should decline.

2.2 Second, mean agreement among cross-party pairs should increase as the size of the majority party increases. Larger majorities allow for increased inter-partisan voting.

The final hypothesis is that legislative polarization is a function of party culture. One party may simply be more disposed than the other toward running the chamber in a partisan
fashion. As mentioned previously, some literature finds that the Republican Party is the more hierarchical and more ideologically unified of the two, suggesting that the GOP would run the House in a more partisan fashion (see Freeman 1986; Skinner Masket and Dulio forthcoming).

If, in fact, differences in party culture could explain changes in legislative polarization in the House, we would expect the following findings from a study of agreement scores:

3.1 First, the mean agreement among intra-party pairs should be greater when Republicans are in the majority.

3.2 Second, mean agreement among cross-party pairs should be lower when Republicans are in power.

**Results: Congressional Voting, 1967-2008**

We use agreement score matrices for the 90th through 110th Congresses (1967-2008) to look for evidence of each of the explanations of the sources of congressional polarization based on our expectations.

First, we look for evidence of mass polarization over time by examining mean agreement across time among members of Congress. Figure 2 shows the mean agreement for same and opposite party members during the time period of analysis. Here we observe that the slope for same party agreement is (slightly) positive and statistically significant (slope: 0.008, standard error: 0.0007). The slope of mean agreement for opposite partisans is (slightly) negative and statistically significant (slope: -0.0093, standard error: 0.0018). These findings are consistent with our expectations. Same party members have, over time, have tended to agree more frequently with one another, while opposite party members, over time, have tended to agree with one another less often. This is consistent with notions of a secular polarization driven by voters; as congressional districts sort themselves out and become more ideologically homogeneous, they
are sending more coherent party caucuses to Washington. The members of Congress produced by these districts vote more with members of their party and less with members of the other party because those are the sorts of representatives voters are increasingly selecting.

Second, we examine the institutional hypothesis – the notion that legislative polarization is a function of majority size. We do so in Figure 3, which compares the seat advantage held by the majority party across time with the mean agreement rate for pairs of legislators. The horizontal axis represents the percent (ranging from 0 to 20) of the seat advantage held by the majority party over the chamber (that is, the percent greater than 50 percent, so that a majority party that controls 51 percent of the seats is represented by .01 on the horizontal axis), and the vertical axis represents the mean agreement for pairs of majority party legislators. Time is represented in the graph by the size of the character—larger characters indicate more recent congresses. Figure 3 shows that as the size of the seat advantage held by the majority party increases, the rate of agreement among majority party pairs declines (slope: -0.9073, standard error: 0.2697). This is consistent with our expectation. Likewise, Figure 4 shows the same thing for minority party members. As the majority party seat advantage increases, the mean agreement rate among minority party pairs decreases (slope: -0.6757, standard error: 0.2039). The negative slopes in these figures emphasize this point, and are statistically significant at the 0.05 level.

We also examine the extent to which opposite party pairs agree, with respect to majority seat advantage. In Figure 5 we observe that the mean agreement among opposite party pairs increases as the majority party gains in size. The slope of the line is slightly positive, although
not statistically significant (slope: 0.3905, standard error: 0.3703). This is weaker support for the institutional hypothesis.

Figure 5 Goes Here

Finally, we examine the relationship between party culture and legislative polarization expectation by studying same-party and cross-party voting patterns over time. Recall that the expectation was that same-party pairs should have a higher agreement score when Republicans are in power. Figures 3 and 4 offered some support for this prediction, although such findings were also consistent with the hypothesis that more narrow majorities would produce greater discipline. (Republicans have tended in recent decades to lead the House of Representative with smaller majorities than Democrats have.)

Figure 6 shows a density plot of the distribution of agreement scores in the House over time. The y-axis shows the density of the agreement rates. The x-axis shows the standardized agreement rates from 0-1. Same-party pairs are solid lines while cross-party pairs are dotted lines. Finally, Democratic and Republican majorities are indicated with color (Red=Republicans, Blue=Democrats), and darker lines indicate more recent congresses. In this graph we observe that same-party pairs have greater agreement rates than cross-party pairs (solid lines are further right than dotted lines). Also, as expected, Republican majorities tend to see greater agreement rates among same-party pairs than do Democratic majorities (the red solid lines are further right than almost all the blue solid lines). The only exception to this occurs in the 110th congress, where a Democratic majority behaved more like a Republican majority (this is demonstrated by the blue peak on the rightmost part of the graph). We also observe that during Republican-controlled congresses, opposite-party agreement occurs less often than in
Democratic controlled congresses (dotted red lines tend to be further left than dotted blue lines). These findings are all consistent with our expectations.

Figure 6 Goes Here

Conclusions

The above analysis provides support for a number of existing theories about sources of legislative polarization in the U.S. Congress. The parties are becoming more internally homogenous and agreeing with each other less over time as the voters themselves behave in a more partisan fashion. Parties tend to vote with greater discipline when they are near numerical parity in a chamber, as cohesion can prevent embarrassing losses for the majority party or provide occasional surprising wins for the minority. And finally, Republicans tend to vote with greater internal cohesion than Democrats do. While our data do not allow us to conduct particularly penetrating tests into the relative importance of these three sources of polarization, it should be noted that we find modest evidence that institutional causes trump cultural ones. Democrats almost always enjoy larger House majorities than the Republicans do, and in the one case in our dataset in which Democrats held a smaller, Republican-sized majority, they behaved roughly as the Republicans did.

While our findings are consistent with some prior research, what we have shown is an ability to reach the same sorts of conclusions as others have without resorting to complex mathematical computer programs that make unrealistic assumptions about the knowledge and independence of legislators. The method we have outlined thus offers the advantages of parsimony, intuitiveness, and accuracy without sacrificing much in the way of explanatory power.
What is more, the analysis presented above represents just a small sample of what can be done with agreement scores to examine legislative behavior. It would be possible, for example, to examine the factors that determine agreement, including party, region, cohort, seating assignments, office locations, race, religion, caucus membership, and other factors. Researchers can also use this approach to study the determinants of legislative power. For example, are highly central legislators invariably party leaders and ranking committee members, or do they also include other members who are atypically moderate, charismatic, or something else? Similarly, are freshmen with high centrality scores more likely to become party or committee leaders in subsequent sessions?

We do not consider agreement scores to be a replacement for NOMINATE or IRT measures. Ultimately, these three approaches produce different estimates and scholars’ decisions about which measures are most appropriate for research should be driven by the nature of individual research questions. Agreement scores do allow us to better understand aggregate legislative outcomes as they allow for the possibility that voting coalitions of legislators will emerge, based, perhaps, in part upon their shared social relationships, etc. Agreement or centrality scores are not intended to supplant IRT estimates, but rather shed light on the legislative networks that provide further information about public policy outcomes and policy success/failure. Agreement scores, and especially centrality scores, provide insight that ideal points cannot. They do not, however, provide the type of spatial estimates with confidence intervals that the other approaches provide. Agreement scores are a viable alternative for addressing a variety of questions about legislative politics without making heroic assumptions about independence.
At the very least, examining legislative behavior through the use of agreement scores offers great potential because it sees politicians not as isolated figures but as fundamentally social creatures who build (or break) alliances and friendships in order to achieve their career goals. We believe that researchers can gain considerable traction on understanding legislative behavior by examining the ways in which relationships determine political outcomes.
REFERENCES


Table 1 - Incidence matrix

<table>
<thead>
<tr>
<th></th>
<th>B2 bomber</th>
<th>Cambodia</th>
<th>Tower</th>
<th>MLK, Jr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helms</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dole</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Nunn</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gore</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Kerry</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Actual votes, with 0 indicating “Nay” and 1 indicating “Yea” (Poole and Rosenthal 1997:12)

Table 2 - Ideal Point Scores of Simulated Actor Network

<table>
<thead>
<tr>
<th>Senator</th>
<th>Ideal Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helms</td>
<td>-0.12</td>
</tr>
<tr>
<td>Dole</td>
<td>-0.12</td>
</tr>
<tr>
<td>Nunn</td>
<td>-0.05</td>
</tr>
<tr>
<td>Gore</td>
<td>0.00</td>
</tr>
<tr>
<td>Kerry</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Scores calculated using the pscl package and ideal command with 10,000 iterations.
Table 3: Agreement rate = number of votes on which i,j agreed/number of votes on which i,j voted

<table>
<thead>
<tr>
<th></th>
<th>Helms</th>
<th>Dole</th>
<th>Nunn</th>
<th>Gore</th>
<th>Kerry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helms</td>
<td>1</td>
<td>0.75</td>
<td>0.5</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>Dole</td>
<td>0.75</td>
<td>1</td>
<td>0.75</td>
<td>0.5</td>
<td>0.25</td>
</tr>
<tr>
<td>Nunn</td>
<td>0.5</td>
<td>0.75</td>
<td>1</td>
<td>0.75</td>
<td>0.5</td>
</tr>
<tr>
<td>Gore</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>Kerry</td>
<td>0</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4: Bonacich Power Centrality

<table>
<thead>
<tr>
<th></th>
<th>Helms</th>
<th>Kerry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helms</td>
<td>0.776</td>
<td></td>
</tr>
<tr>
<td>Kerry</td>
<td>0.776</td>
<td></td>
</tr>
<tr>
<td>Gore</td>
<td>1.088</td>
<td></td>
</tr>
<tr>
<td>Dole</td>
<td>1.088</td>
<td></td>
</tr>
<tr>
<td>Nunn</td>
<td>1.195</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Sample Network

In this graph the width of lines (edges) represents the strength of voting agreement (ties) between legislators. The size of each node indicates its centrality (eigenvector centrality). The graph was created using a “spring embedded” algorithm that places nodes that are more highly related to one another closer together.
Figure 2 – Mean Agreement by Congress: 90th-110th
Figure 3 – Mean Agreement by House Majority Size (for Majority Members), 90th-110th Congresses
Figure 4 - Mean Agreement by House Majority Size (for Minority Members), 90th-110th Congresses

Legend:
- D: Minority Same Party Pair (Democratic Session)
- R: Minority Same Party Pair (Republican Session)
Figure 5  Mean Agreement by House Majority Size (for Opposite Party Members), 90th-110th Congresses
Figure 6 Agreement Score Density by Party