THE HUNT FOR PARTY DISCIPLINE IN CONGRESS[#]

Nolan McCarty*, Keith T. Poole**, and Howard Rosenthal***

2 October 2000

ABSTRACT

This paper analyzes party discipline in the House of Representatives between 1947 and 1998. The effects of party pressures can be represented in a spatial model by allowing each party to have its own cutting line on roll call votes. Adding a second cutting line makes, at best, a marginal improvement over the standard, single-line model. Analysis of legislators who switch parties shows, however, that party discipline is manifest in the location of the legislator's ideal point. In contrast to our approach, we find that the Snyder-Groseclose method of estimating the influence of party discipline is biased toward exaggerating party effects.

THE HUNT FOR PARTY DISCIPLINE IN CONGRESS

Introduction

The past several years have seen renewed scholarly investigation of how political parties and their leaders influence legislative institutions and behavior (see Aldrich 1995, Cox and McCubbins 1993, Rohde 1991, and Sinclair 1995). Much of this contemporary research investigates how parties solve the collective action problems that are inherent in the legislative and electoral processes. Cox and McCubbins, who conceptualize political parties as "cartels" that direct legislative activity to enhance the collective electoral fortunes of their members, provide a typical variant on this theme, but by no means the only one. The primary function of a cartel is to build a collective reputation for its members to run under. They argue, however, that without strong leadership members have individual incentives to engage in legislative activities (such as pork) that diminish the collective reputation.

This focus on the problems of collective action has generated much interest in the cohesiveness of parties as floor coalitions.[1] The principal prediction is that a party produces more cohesive coalitions of its members than would be possible if the members were to act on the basis of their individual preferences. Rohde (1991) uses evidence of the increase in party cohesion since 1975 to demonstrate an increasing role of party in the post-reform House. Aldrich, Berger, and Rohde (1999) also use party voting as the main dependent variable to test the predictions of "conditional party government," and Cox and McCubbins (1993) use member support on leadership votes to test for the role of leaders in creating voting coalitions. Furthermore, some scholars, including Rohde (1991), see a reassertion of party strength behind the increased cohesiveness and polarization of congressional parties since the mid-1970s (for alternative explanations, see Poole and Rosenthal 1984; McCarty, Poole, and Rosenthal 1997; and King 1998).

As Krehbiel (1993, 1998) points out, however, these empirical studies of party voting suffer from the problem that the patterns of behavior that have been uncovered are consistent with both theories of strong, influential parties *and* non-partisan models where member preferences are sorted along party lines. This dilemma is exacerbated by the problem of measuring legislative preferences. Ideally, one would like some exogenous measure of these preferences to test party theories. Voting behavior under the null hypothesis of no party influence could then be compared with actual voting behavior. The problem is that our usual measures of legislative preferences are derived from the voting behavior itself.

In this paper, we attempt to untie this Gordian knot. We begin by reviewing the evidence from work that analyzes congressional roll call votes under the maintained hypothesis of sincere spatial voting. We discuss how the evidence from this work in fact suggests the presence of some party discipline. We next develop the party discipline model of Snyder and Groseclose (2000). We argue that their estimation method both seriously biases the estimate of ideal points for ideological moderates and overestimates the extent of party discipline. In the text, we provide a compelling theoretical illustration of the bias. We detail our case further in Appendix B.

To assess party discipline (a.k.a. pressure) properly, we propose an alternative approach. The basic idea is very simple. We start with Krehbiel's (1993, 1998) proposal that the spatial model of purely preference-based voting is the appropriate benchmark for evaluation of models that incorporate party effects. In one-dimension, the spatial model asserts that, on each roll call, "Yea" and "Nay" voters are separated by a cutpoint on the liberal-conservative continuum. Now assume that the Republicans apply "pressure" to their membership. This will cause some moderate Republicans to the left of the "sincere" cutpoint to vote with the conservative wing of the party. Republicans will have a cutpoint to the left of the sincere cutpoint. Similarly, if the Democrats apply pressure, the cutpoint for Democrats will be to the right of the sincere cutpoint. That is, when one or both parties apply pressure, the voting patterns should look as if there were separate cutpoints for each party, with the Democrat cutpoint being to the *right* of the

Republican cutpoint. Consequently, if pressure is important, we should find a better fit to the data when we estimate two cutpoints than when we estimate a single cutpoint.

To keep the analysis as simple as possible, we use non-parametric optimal classification analysis where legislator ideal points and roll call cutpoints are jointly rank ordered to maximize predictive success on roll call votes (Poole, 2000). By classifying the voting of each party independently and then comparing the results to classifying both parties together, we can evaluate the maximum possible improvement in correct classification attributable to party discipline. An advantage of the cutpoint approach is that it does not require any assumptions about which specific roll calls are subject to party pressure. A particular advantage of the non-parametric approach is that it assumes only that the amount of pressure applied to individual members does not change the order of their induced ideal points. It does not require making parametric assumptions about how pressure varies with the ideal point of the individual member, such as equal pressure being applied to all. On the basis of our optimal classification analysis, we conclude that allowing for party discipline affords only a very marginal improvement over the sincere spatial model, particularly in recent Congresses.

Where, then, is party discipline? We argue that the main influence of party discipline is not on the votes on specific roll calls but on the choice of ideal point made by the representative. The smoking gun is provided by the great changes in ideological position demonstrated by those few legislators who have switched parties. Wayne Morse and Strom Thurmond are two well-known examples in the post-war Senate. The Democrats who defected to the Republicans after the 1994 election made equally dramatic shifts. Our finding that parties shape ideal points ends our hunt for party discipline in roll call voting.

Independent Voting on the Floor: The Evidence from the Spatial Model

The well-known standard spatial model provides a benchmark approach to *independent* floor voting. Poole and Rosenthal (1991, 1997) demonstrated that the spatial model is quite successful in accounting for floor decisions. With two dimensions, one can correctly predict roughly 85% of the individual decisions -- even on close roll calls -- for the period 1789-1985. McCarty, Poole, and Rosenthal (1997, p. 7) report additional results for the period 1947-1995.[2] In recent Congresses, a one-dimensional model classifies nearly 90% of the individual decisions (see figure 5).

The spatial estimates present a strong suggestion that party influence underpins much of this remarkable classification success:[3]

- In Congresses where voting is largely one-dimensional, party-line votes are along the main dimension. The distribution of ideal points is strongly bimodal. The two parties appear as two very distinct "clouds". The clouds, particularly in recent years, barely overlap. [As an illustration, see McCarty, Poole, and Rosenthal (1997, p. 11).] The presence of a "channel" between the clouds suggests that party affiliation may discipline the roll call voting behavior of members. The main dimension of political conflict clearly appears to reflect partisan conflict. Parties perhaps also influence their members' votes on specific roll calls.
- In Congresses where voting is two-dimensional, there are also two distinct clouds separated by a channel. Party-line votes are no longer on the main dimension, but a blend of the first and second dimensions. [See Poole and Rosenthal, 1991, p. 233 or 1997, p. 44 for an example.] An interpretation of such plots is that ideal points projected onto roughly a 45° line represent the ideological (liberal-conservative) dimension. The orthogonal projection, roughly at –45°, represents a party loyalty or valence dimension. Most votes occur along the main, 0° dimension. On these votes, the legislator's decision depends both on ideology and on party loyalty.

Although this evidence shows that the structure of voting coalitions in Congress coincides strongly with party affiliation, it does not prove that party *per se* has any influence on voting behavior. Party-line voting is, of course, consistent with both strong party models and ideological models where preferences are sorted by parties. In the sections that follow, we review a recent attempt to separate partisan effects from preferences and propose a method of our own.

The Snyder-Groseclose Model of Party Discipline

One of the inherent problems in identifying the effects of party is that we observe only behavior, which is presumably a mix of individual preferences and party influence. This problem is particularly acute with congressional voting data. If party discipline is exercised on floor votes, the ideal points estimated on the assumption of independent spatial voting might be very biased estimates of legislator preferences. If party influences these estimates, it is inappropriate to use them as controls for preferences when testing for a party effect. Snyder and Groseclose (2000) noted this potential for bias. They proposed both a method for first estimating unbiased ideal points and then for using the unbiased ideal points to estimate the effect of party discipline.[4]

The basics of the one-dimensional Snyder-Groseclose model are as follows:

- On roll call j, a legislator i, if a Republican, has induced ideal point $x_{ij} = x_i$
- On roll call j, a legislator i, if a Democrat, has induced ideal point $x_{ij} = x_i + \gamma_j$

In other words, the true ideal points of the Democrats, the x_i, are displaced by the amount of party pressure given by γ_j.[5] It turns out that only the relative amount of party pressure matters in the model, so the ideal points of the Republicans can just be given by their true values. For the difference in pressure to be consistent with discipline, we would expect that pressure must move Democrats in a liberal direction relative to Republicans. Thus, pressure works to increase the separation of the parties. If preferences are scales such that left is liberal and right is conservative, then we would expect γ to have a negative sign

Snyder and Groseclose argue that, because there would be little need to apply party discipline on votes not expected to be close, ideological position-taking could occur on lopsided votes. These votes, for example, those with margins over 65-35, could be used to estimate the true ideal points. On these votes, the g_j would be zero. The true ideal points could then be used to estimate the g_j on close votes, say those with margins less than 65-35.

In brief, their procedure is:

<u>Stage 1</u>. Use votes with margins greater than 65-35 to estimate the ideal points, x_i.

<u>Stage 2.</u> On the remaining votes, for each roll call j, estimate the following OLS (ordinary least squares) model:

$$Y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 D_i \tag{1}$$

where D_i = 1, if legislator i is a Democrat, = 0 if Republican; and Y_{ij} = 1, if i votes Yea, = 0 if i votes Nay. In the Appendix A, we show that if the underlying spatial utility is quadratic, then

$$\gamma_j = \beta_2 / \beta_1 . \tag{2}$$

As we noted above, the party pressure model predicts a negative estimate for γ when preferences are scaled with Democrats on the left (as we assume they are). Consequently, the two estimated β's should be of opposite sign.

Why the Method Overestimates Party Discipline

This method is likely to generate the inference that party pressure is substantial even when all voting is preference based. Consider, for example, a six-member legislature with the party affiliations and spatial preferences given in figure 1. If all voting in this legislature is spatial without error, there are only 12 possible voting configurations, which are given in figure 2.

Stage 1 of the Snyder-Groseclose method estimates a preference score using only voting patterns 1-10. But, since voters 3 and 4 cast identical votes in each of these patterns, any scaling procedure will estimate voters 3 and 4 as having the same position. *Thus stage 1 provides biased estimates of the preferences of moderates. There is not enough information in the lopsided votes to discriminate "left" moderates from "right" moderates.* The preference ordering that maximizes the classification of votes is shown in figure 3.

In stage 2, the preferences in figure 3 and party affiliation are used to explain vote patterns 11 and 12. The votes of legislators 1, 2, 5, and 6 are correctly classified on the basis of the preference estimates, but the votes of legislators 3 and 4 cannot be. However, since 3 and 4 are members of different parties, adding party to the model increases its explanatory power *even though voting is purely preference driven*.

Our example extends naturally to larger legislatures. In general with perfect spatial voting, a first stage based only on lopsided votes will produce identical preference estimates of all members in the interval between the 35th and 65th percentiles. The second stage will therefore produce a spurious party effect so long as party and ideology are correlated within this interval. Given the assumptions of no voting error and no overlap of preferences between the parties, this example is somewhat special. However, in Appendix B, we present Monte Carlo evidence that shows how this result extends to large legislatures where, as in the Snyder and Groseclose approach, there is some error in voting and the distributions of preferences of each party overlaps. We now present an alternative procedure that maintains the essential features of their model of discipline.

A Non-Parametric Model

All specifications of a spatial model of voting have two critical elements: ideal points for the legislators and cutpoints (or separating hyperplanes) for the roll calls. The Snyder-Groseclose model, with a discipline parameter to each roll call, is isomorphic with one where each party has its own cutting line. (See Appendix A.) That is, moving the ideal points for all Democrats to the left by a magnitude γ_j is equivalent to moving the cutpoint for Democrats to the right by the same amount. Party discipline generally involves getting moderates to vote with extremists.[6] Consequently, if there is party discipline, the cutpoint for the Democrats should be to the right of the cutpoint for Republicans.[7]

Consider a one-dimensional spatial configuration. If the cutpoint is constrained to be the same for both parties, this produces the standard spatial model. For example, in figure 4, with a common cutpoint, there are three classification errors, legislators 3, 11, and 15. When each party can have its own cutpoint, this produces a model that allows for party discipline. Moderate Democrats to the right of some Republicans can vote with the majority of their party. Moderate Republicans to the left of some Democrats can vote with the majority of their party. The best cutpoint for the Republicans in figure 4 remains the common cutpoint. Legislator 15 is the only R classification error. But the best cutpoint for the Democrats is to the right of the common cutpoint. The D cutpoint leaves only legislator 3 as a classification error for this party. Rather than estimate either the one-cutpoint model or the two-point model via a metric technique, such as Poole and Rosenthal's (1991) NOMINATE or Heckman and Snyder's (1997) method, one can simply find the joint rank order of legislators and cutpoints that minimizes classification error. Poole (2000) presents an efficient algorithm that very closely approximates the global maximum in correct classification.[8] Note that this method, in contrast to equation (1), does not require a uniform adjustment in the ideal points of all members of a party. Only moderates would need to be disciplined. All that is required is a displacement of the cutpoint.

We now turn to our empirical analysis. This analysis involves both the testing of implications of our methodological critique of Snyder and Groseclose as well as testing the implications of their model of party discipline utilizing our two cutpoint model. We begin by presenting evidence on three methodological predictions. In each case, these predictions are consistent with the mismeasurement of preferences in the Snyder-Groseclose framework under the hypothesis of purely spatial voting. In only one of the cases is the prediction also consistent with their theoretical model. Therefore, verification of these relationships illustrates the inability of Snyder-Groseclose to distinguish party pressure from mismeasurement of preferences. The three methodological predictions are as follows:

M1. Estimate the rank order of ideal points by one-dimensional optimal classification first using all roll call votes and then using only lopsided votes. The correlation of the all votes rank orders and the lopsided votes rank orders will be greater for extremists (the first and last thirds of the all votes distribution) than for moderates (the middle third). While this prediction is consistent with the Snyder-Groseclose assertion that party pressure primarily affects moderates, it also follows from our claim that, if there is preference-based voting, ideal points of moderates will be inaccurately recovered if only lopsided votes are used to estimate ideal points.

M2. Similarly, when the rank order is estimated first on all roll call votes and second on only close votes, the correlation of the all votes ranks and the close ranks will be greater for moderates than for extremists. The motivation for this prediction is similar to that of the first. If there is preference-based voting, the ideal points of extremists will be inaccurately recovered if only close votes are used to estimate ideal points. This prediction is inconsistent with Snyder and Groseclose as they implicitly assume that extremists will have similar preference estimates on pressured votes as they do on unpressured votes.

M3. The close-all correlations for moderates will be high if there is preference-based voting, lower if there is party discipline. The reason is that, if there is discipline only on close votes as claimed by Snyder and Groseclose, all votes estimates will mix preference-based lopsided votes and disciplined close votes. The close votes estimates will have more distortion of the true ideal points.

After these tests concerning the effects of Snyder and Groseclose's procedure on ideal point estimates, we turn to testing hypotheses from the party discipline model. In all cases, the null model of preference-based voting predicts no difference.

H1. Classification should be substantially higher with a two-point model than with a one-point model. Note that classification cannot be lower with the two-point model.

H2. The improvements in classification should be greater on close votes. Since Snyder-Groseclose predict that rational parties will whip close votes, the incremental predictive power of the two-cutpoint model should be accordingly higher on those votes.

H3. The rank order of the legislators should disclose more separation of the parties in the one-point model than in the two-point model. The reason is that the one-point model ignores party discipline. Moving Democrats to the left and Republicans to the right should pick up some of the effects of party pressure. In contrast, in the two-point model, each legislator's ideal point can take on its true rank order position, because the cutpoints can pick up the effects of party discipline.

H4. The separation of the cutpoints should be greater on close votes. The identifying assumption of the Snyder-Groseclose model is that party pressures are more likely on close votes. Therefore, under their assumptions, the distance between the Democratic and the Republican cutpoint should be greatest on those votes.

H5. The estimated cutpoint for the Democrats should be to the right of the estimated cutpoint for the Republicans.

A Caveat

Some instances of party pressure may be masked. Consider a legislature with no party overlap. All Democrats are to the left of all Republicans. Suppose that, were there no pressure, a Republican Party bill would be rejected by a majority composed of all Democrats and some moderate Republican defectors. If the Republicans then apply "pressure" to the defectors, resulting in a party-line vote, the vote will still appear to be a vote consistent with preference-based voting. Thus when ideal points are estimated correctly, the true explanatory power of party may be masked. In fact, when there is no overlap in the distribution of party ideal points and there is errorless spatial voting, it is impossible to identify party pressure effects.

This masking of party pressure is inherent to spatial analysis. It would confound a correct Snyder-Groseclose analysis as well as our optimal classification method. Albeit important, the question we can ask is limited to "Can allowing for party discipline improve on the classification of a purely preference-based model?"[9]

With our optimal classification method, it is possible to calculate an upper bound for the amount that party pressure can increase vote classification, roughly as a function of the overlap between the two parties. This upper-bound represents the classification on a strict party line vote of a two-cutpoint model (perfect classification) minus the classification of a strict party line vote using a single cutpoint.[10] When there is no overlap between the parties, a single cutpoint correctly classifies a party line vote so as noted above their can be no classification gain for the two cutpoint model. However, the greater the party overlap, the worse a one-cutpoint model does in explaining a strict party-line vote. Thus, the maximum classification gain increases in the overlap. If we use the configurations of preferences that emerge from optimal one-cutpoint classification to measure overlap, the maximum classification gain from party-pressure consistent cutpoints (i.e. D>R) ranges from 0 in the 80th House (where there is zero overlap) to 16% in the 92nd House. The average upper bound over all of the congresses we analyze is 5%. However, it is important to remember that these upper bounds are simply for roll calls consistent with party pressure (i.e. Democratic cutpoint to the right of the Republican cutpoint). Perfect classification is the upper bound if we allow other cutpoint configurations (e.g the Republican cutpoint on the right). Secondly, as we discuss below, optimal classification with a single cutpoint will underestimate party overlap which would lead to the underestimation of these upper-bounds.

Tests Using the Non-Parametric Model

One Dimensional Classification

We begin with the three predictions concerning correlations of ideal points. In order to show that the pattern of estimates we expected would arise in actual data, we first performed optimal one-point classification using *all* the roll call votes in each House from the 80th through the 105th. If the basic spatial model is correct, this procedure should produce a rank order of legislator ideal points that is very close to the true order. Next, we did optimal classification using only *lopsided* votes, those with greater than 65-35 margins. Finally, we did optimal classification using only *close* votes, those with margins of 65-35 or less.

We then computed Spearman rank order correlations between the *lopsided* vote rank orders and the *all* votes rank orders for left-wingers, the one-third of the legislators furthest to the left in the *all* votes classification; moderates, the middle one-third; and right-wingers, the one-third furthest to the right. We would expect these correlations to be high for the left-wingers and right-wingers but low for the moderates because the *lopsided* votes provide little information about the ideal points of moderates (M1). Conversely, when correlations are made between *close* votes rank orders and *all* votes rank orders, we expect the correlations to be high for moderates but low for left-wingers and right-wingers.[11]

Insert Table 1 about here

The hypothesized patterns occur, as shown in Table 1. Indeed, for the *lopsided-all* comparison, in every post-war House but one, the middle one-third correlation is lower than both that for the left-wingers and that for the right-wingers.[12] Table 1 indicates that the middle correlation is particularly low in the period preceding the passage of the major civil rights bills of the 1960s. In this period, there was an important second dimension (Poole and Rosenthal, 1997) that confounds the recovery of moderate positions on the first dimension. When the second dimension vanishes, even the middle correlations are reasonably high because the "errors" in voting provide some information about moderates. That is, for example, a relatively liberal moderate is still less likely to vote with the right-wingers than is a relatively conservative moderate, even on lopsided votes. Nonetheless, as predicted by M2, correlations for moderates are lower than for extremists.[13]

As predicted, these results reverse for the *close-all* comparison. The moderates always produce a correlation above 0.9. The left-winger and right-winger correlations are always below 0.9, usually much below, and in one case, the correlation is negative.

The *close-all* correlations for moderates are strikingly high, predicted by preference-based voting but not by voting subject to party discipline (M3). If the party discipline effect were important, we would expect lower rank order correlations, particularly for Houses before 1980, when there was still considerable overlap in the ideal point distributions of the two parties.

Classification with Two Cutpoints

In this section, we assess the ability of a party discipline model to improve on a preference-based model. Our criterion is percentage of votes correctly classified.

To find the highest classification possible for a party discipline model, there is a simple solution: just classify each party separately. This allows the cutpoint on each roll call to adjust to pressures internal to the party. *Because the cutpoints can adjust, one will find the true intra-party rank order of the ideal points.* The classification from this model can be compared to a single cutpoint model.

The results of this exercise appear in figure 5, which shows results for one-, two-, and six-dimensional models. We used a six-dimensional model to parallel the high dimensionality used by Snyder and Groseclose in their empirical work. With multiple dimensions, the cutpoint is replaced with a separating hyper-plane.

In one dimension, it is apparent that a two-party model adds little, particularly in recent Congresses. The improvement in the earlier Houses is at the level that results when a two-dimensional model with one cutting line is used. In one dimension, the two cutpoints allow for Southern Democrats to vote with Northern Democrats on some issues, but they also allow for votes where a coalition of conservative Republicans and Southern Democrats opposes liberal Republicans and Northern Democrats. Since the Democrats were the majority party in the conservative coalition era, these votes demonstrate breakdowns of party discipline that would be exactly opposite to the basic assumption of the Snyder-Groseclose model.

In two or six dimensions, allowing for two (as against one) separating hyperplanes results in even less improvement than in the one-dimensional case. In fact, the improvement is almost always less than 1% for all post-war Houses. The smaller improvement occurs because, in one dimension, "party" was picking up some effects than can be accounted for just as well by a higher-dimensional preference-based voting model. The strength of the results in figure 5 is further emphasized by two observations. First, some of the increase in fit is simply noise-fitting due to the extra degrees of freedom. Second, classifying each party separately allowed for "both ends against the middle" voting where liberal Democrats and conservative Republicans vote together. This last problem and other considerations lead us to adopt a slightly different approach.

The remainder of the analysis in this section uses a two-stage procedure:

1. Using optimal classification, we estimate a *one-dimensional spatial model that has a single cutpoint, common to both parties*.
2. *Holding the rank order positions of the legislators constant* at the positions produced by step 1, we then estimate *separate cutpoints* for the two parties. The two cutpoints must be placed to maintain polarity. That is, unlike in the separate scalings reported in figure 5, we did not consider improving classification by allowing moderates to be opposed by extremists at both ends of the spectrum.

Bob Barr and Maxine Waters can't vote together against Connie Morella. This constraint is fully consistent with the Snyder-Groseclose approach that calls for an order-preserving shift in a party's ideal point distribution but not for a flip-flop.

The motivation for this two-step approach is that it is not possible to estimate jointly a single order for the legislators and two cutpoints for each roll call. The reason is that the rank order of the legislators within each party is pinned down only by the cutpoints for that party. Consequently, it is impossible to rank order either the legislators of a party or the cutpoints for that party with those for the other party. In contrast, once we fix the rank order of the legislators, we can estimate separate cutpoints and test theoretical predictions about these cutpoints. We cannot directly test **H3**, however, that preferences will show less party overlap in a one-point model than in a two-point model. That hypothesis could be tested only indirectly, by our test of **M3**.

To justify holding the legislators constant, we computed within-party Spearman rank order correlations between the rankings of the single cutpoint model and the rankings when optimal classification is applied to the party separately. Recall that this separate classification is consistent with a party pressure model—there is a true underlying order of ideal points but cutpoints are adjusted to reflect party pressure. As table 2 shows, these correlations are remarkably high. For both parties, the correlations are above 0.95 since the mid-1960s. [Previously, some correlations were lower as a consequence of the presence of an important second dimension.] Consequently, the single cutpoint ratings, particularly for the past 30 years, are likely to provide accurate rankings of the "true" ideal points within each party.

Insert Table 2 about here

Note that the analysis presented in table 2 informs us that the relative order of legislators *within* each of the two parties is insensitive to whether we just assume pure preference-based voting or explicitly account for party pressure. The result does not rule out party pressure; it just tells us, consistent with equation (1), that party pressure is unlikely to change the relative order of induced ideal points. The result does not rule out party pressures polarizing one party relative to the other. The lack of overlap we observe in the 1990s might, for example, be the result of party pressure. We return to this point presently.

The two-point model creates only minor gains in classification of roll call votes. As the second dimension has diminished in importance, these gains, as shown in figure 6, have declined to under 0.5% in the last 8 Congresses. In other words, adding a second cutpoint typically allows correct classification of only an additional 2 of the 435 representatives (assuming full turnout). Note that (1) the classification must get better with a second cutpoint, (2) the second cutpoint can just fit noise in the data (see Poole and Rosenthal, 1997, p. 156), and (3) much of the improvement in classification occurs from using two cutpoints that have the Democratic cutpoint counter-hypothesis to, that is *left* of, the Republican cutpoint (see table 3). Thus, the improvements of under 1 percent are truly small potatoes.[14] **H1** is not supported.

Figure 7 shows the results for close and lopsided roll calls and contains a wee bit of good news for advocates of party pressure theories. The classification gain is greater on close roll calls than on all roll calls, but only since the mid-1960s. The evidence for the earlier Congresses reinforces our contention that the larger improvements in classification for these Congresses shown in figures 5 and 6 are the work of a second dimension. If discipline were producing the gain, the gain should not occur on lopsided roll calls. There is a systematic difference in the gain on close and lopsided roll calls in later Congresses. However, some of the gain on close roll calls must result from non-discipline factors—such as noise fitting—that affect lopsided as well as close votes. The difference between the gain on close and lopsided votes is roughly one percent. The gap of only one percent suggests that party pressures are changing not more than about 4 votes per roll call on the close votes. At best, **H2** is weakly supported.

The fourth hypothesis derived from the party pressure model is also weakly supported. To test **H4** we computed the average of the difference between the rank of the Democratic cutpoint and the rank of the Republican cutpoint and then divided by the number of legislators serving in the House. This procedure normalized the difference in the rank orders to a –1 to +1 scale so that the Houses could be more easily compared. We used the difference rather than the *distance* (absolute difference) between the ranks because the pressure model predicts that the Democratic cutpoint should be greater than the Republican cutpoint (D > R).

We classified all roll calls into three types. Our first type includes roll calls where, in line with the fifth hypothesis, the Democrat cutpoint was greater than the Republican cutpoint (D>R). Note that whenever there is some overlap in the ideal point ranks of the two parties, straight party-line votes are counted as D>R. Our second type is clearly counter-hypothesis roll calls with R> D that satisfied this condition. Finally, for many roll calls (see table 3), the relative locations of the two party cutpoints were ambiguous. We term this third type "Undecided". Note that cutpoints which are interior to the legislators of a party can be identified for only a subset of roll calls. A portion of our analysis will be restricted to such roll calls. [15]

When the ideal point distributions of the two parties have no overlap, as happened in the 80th House (1947-48), we cannot identify any roll calls as D>R so the average difference *must* be less than zero. In contrast, when there is substantial party overlap, as in the 1970s, the party pressure model predicts that the average difference for close votes should be greater than zero and be greater than the average difference for lopsided roll calls. The average difference for lopsided roll calls should be close to zero. The results, computed for all roll calls with interior cutpoints in both parties, appear in figure 8.

The average difference for the close roll calls is indeed above zero for 19 of the 26 Houses. Since the 91st House, however, the average difference is *very* close to zero – hovering around .02 or an average difference of about 8 to 9 ranks. In only three Houses, all in the two-dimensional 1950s and 1960s, does it exceed 0.1 or 10% of the House membership. To benchmark this difference, the normalized difference or overlap between the third rightmost Democrat and the third leftmost Republican averages 46% of the House membership for the 26 Houses we analyze; it exceeds 32% in all but the 80th, 84th, and 100th to 105th Houses. Moreover, note that this average difference is highly biased in favor of the party pressure model in that it does not include “Undecided” roll calls. These include, for example, all roll calls on which the Republicans are unanimous but the Democrat cutpoint is to the left of the leftmost Republican. Such roll calls are most likely ones where party discipline broke down among the Democrats so that D<R.

The average difference for the lopsided roll calls is negative for all 26 Houses. That the difference is negative probably reflects instances of “both ends against the middle” voting. If the six most liberal Democrats and the six most conservative Republicans cast protest votes on final passage and they are the only negative votes, with fixed polarity, one of the party cutpoints will be near an end of the dimension while the other party cutpoint will be near the middle of the dimension. Consequently, the difference in ranks will be negative and large in magnitude. The negative differences can reflect a few conservative Republicans and a few liberal Democrats voting against a lopsided majority.

Finally, **H5**'s prediction that the Democrat cutpoint would be to the right of the Republican cutpoint is not supported, as shown in table 3. The pattern, except for the no overlap or low overlap Congresses 80 and 103-105, is quite stable, so we present results in tabular form. Recall that in low overlap Congresses, there are very few or no roll calls with D>R. But even in Congresses with overlap, the pattern runs counter to the Snyder-Groseclose model, with R>D roll calls outnumbering the “pressure” D>R roll calls by more than 3 to 2.

Insert Table 3 about here

Table 3 is much less favorable to the party pressure model than figure 8 because for many of the

Houses a handful of Southern Democrats were in the midst of the Republicans and a handful of liberal Republicans were in the midst of the Democrats. Consequently, on party-line or near party-line votes, D>R and the difference in ranks was quite large. The differences in ranks are smaller in magnitude on counter-hypothesis R>D votes, but such votes are typically a majority of the roll calls.[16]

Some of the counter-hypothesis R>D votes almost certainly indicate a true breakdown of party discipline. A breakdown of party discipline can occur, for example, when the majority is subject to a few defections of its own moderates but offers bills or makes promises that buy the support of moderates of the opposite party. The seduction of minority moderates is a scenario that seems to fit the two Gingrich Houses, where, in the single cutpoint analysis, the modal cutpoint fell interior to the Democratic Party (see McCarty, Poole, and Rosenthal, 1997, 12). [The two Gingrich Houses are the last two points in every plot.]

These results about cutpoints are, however, subject to the warning that the single cutpoint estimation of ideal point ranks might possibly show too much separation of the parties. We therefore calculated how far the ideal points of Republicans would have to shift leftward until the average difference for lopsided roll calls was zero. Once each House has been shifted, a new version of figure 8 would have a flat line through zero for lopsided votes. That is, the shift forces the average pattern for lopsided votes to match the theoretical level in the Snyder-Groseclose model.[17]

The results of this exercise are shown in figure 9. When the lopsided vote difference is just slightly negative, as in the late sixties, very few ranks need to be shifted. In these cases the close vote difference is near zero and R>D roll calls outnumber D>R, so the Snyder-Groseclose model is not supported. Where the lopsided vote difference is sharply negative (see figure 8), in the late forties and in the nineties, many ranks have to be shifted to force the lopsided votes to show a zero average. In the most recent Congresses in our time series, the order of change is of 100 ranks, or about half the Democratic membership. Nevertheless, placing the “true” ideal points of the most moderate Republicans in the middle of the Democratic Party is seriously lacking in face validity. The amount of overlap in the ideal point distribution is just too great to make a party pressure model credible. In fact, the amount of shifting needed matches the decrease in party polarization in the post-war period and its increase since the late 1960s (McCarty, Poole and Rosenthal, 1997) as measured via NOMINATE scores. The increasing separation of the parties one sees is, in our view, much more likely to reflect fundamental political changes, such as a large increase in southern Republican representatives, than an increase in party discipline within Congress.

Since our initial ideal point distribution has greater face validity than the shifted distribution, we use the initial distribution to ask a final question in this section. Does discipline make a difference in outcomes? We assess this in two ways:

A. We assume the true cutpoint is the minority cutpoint. Pressured voters are those majority party voters with ideal points between the minority and majority cutpoints. This reflects a benchmark where all pressure is exerted by the majority party. Would the outcome have changed were their votes reversed?

B. We assume that the true cutpoint is the average of the two party cutpoints, reflecting a scenario in which both parties exert equal pressure. Pressured voters are those voters with ideal points between their party cutpoint and the average. Would the outcome have changed were their votes reversed?

The results vary substantially from one Congress to the next, in part a function of the separation of party ideal points. We find that, averaged across Congresses, discipline makes a difference, for assumption A, on 16.97% of close roll calls (Std. Dev. 9.17) and, for assumption B, 11.07% (Std. Dev. 8.03%). While these numbers are substantial, they are well below the proportion of significant t-statistics reported by Snyder and Groseclose. Moreover, they are almost certainly overestimates. One qualification is that assumption A is extreme, since it assumes only the majority party exerts pressure. Another is that some of the pressured voters may not have changed their votes were "pressure" removed. This is because under the null hypothesis of a single cutpoint, errors in voting will result in there being legislators on the yea side of the cutpoint who vote nay, and vice versa. Similarly, under the alternative hypothesis of two cutpoints there will be two types of legislators between the cutpoints – those who are pressured and those who voted with their party for idiosyncratic reasons. Scenario A mistakenly counts both types of legislators as pressured.

This section, in summary, has established that:

- Allowing for party discipline does not make an important contribution to classification.
- Those improvements in classification that do occur are, more frequently than not, the result of using cutpoints that are inconsistent with the party pressure model.

Ideal Point Changes in Party Switchers

If there is little evidence that many ideal points are displaced on individual votes, there is very substantial evidence that party affiliation has a strong influence on ideal points. To see this, we used the procedure of Poole (2000) to obtain rank orders of the ideal points in separate estimations for the House and Senate using all roll calls from 1947 to 1998.[18] Each member was constrained to have a constant ideological position in his or her career, except that party switchers were allowed to have two positions, one before and one after the switch. There were 472 senators and 2,326 representatives (counting the party switchers as two individuals). The orderings were normalized to 0-1 by dividing the raw ranks by 472 for the Senate and 2,326 for the House.

When legislators switch from R to D, they should have a lower rank. The reverse should hold for D to R switchers. There were 19 legislators who both switched parties and remained in the same house in the period of our analysis. They are listed in table 4. In 18 of 19 cases the rank changed as expected. The only exception is Strom Thurmond. His slightly more moderate position as a Republican is a reflection of his more moderate views on race relations in the past 20 years. A simple sign test is overwhelmingly significant. Induced ideal points respond to party affiliation.[19]

We have shown that party switchers generally move in the theoretically expected direction. Did they move very much? The average rank movement was 0.28; thus a switch induced a jump over more than one-fourth of all legislators serving in the period. To benchmark this movement, we reran the analysis for the House allowing two positions not only for the party switchers but also for some legislators who never changed party. Specifically, we picked in the legislator file every 500th legislator among moderates — that is those with ideal points between -0.3 and +0.3 — who served in at least 2 Houses.[20] There were 15 such representatives, matching the number of actual switchers in the House. For each group of 15 we computed the average partisan switch. That is, for Democrats the switch was just the change in the coordinate, as Democrat switchers are expected to increase their ranks. But for Republicans, we used the negative of the change. Actual switchers moved substantially, a change in normalized rank of 0.281. On average, non-switchers barely budged, moving only 0.026 in ranks. The (one-tail) t-statistic for the differences in the means indicate a high level of statistical significance.

Insert Table 4 about here

This evidence is consistent with a party effect, but a couple of caveats are in order. First, it is silent on the mechanism that generates this effect. Therefore, the source may not be internal to the legislature. Switchers after all have to adapt to a new set of primary constituents and contributors as well as legislative leaders. Second, party switchers are obviously not a random sample of all legislators. In the 104th House, southern Democrats switched to the Republican Party for a reason -- they wanted to reflect the increasingly conservative temperament of their districts.[21] Therefore, selection bias precludes us from suggesting that the shift in ideal points is an unbiased estimate of party pressure. But even if the selection bias were severe, it is telling that changing party labels was deemed necessary to reflect changing district sentiment.[22] Third, the estimates based on party switchers are almost certainly an upwardly biased measure of the average amount of discipline. Those members who do not switch probably have more congruence between their personal/constituency position and the party's desires.[23] In particular, those representatives close to the party median are likely to vote "correctly" without any discipline.

Conclusion

In the past decade, theorizing about the influence of parties and leaders on legislative behavior has outstripped progress in solving difficult methodological and measurement problems necessary to test these theories. In this paper, we have addressed the problems associated with distinguishing party effects from a null hypothesis of individual preference-driven behavior. We began by demonstrating the unattractiveness of regression-based procedures such as that of Snyder and Groseclose. We find that these methods of estimating the effects of party discipline on individual roll call votes are biased toward exaggerating the effect of party discipline. To remedy these statistical problems, we incorporate the theoretical insight of Snyder and Groseclose into the spatial model of voting, which we estimated non-parametrically. We find that empirically, a party discipline approach makes, at best, a marginal improvement over the standard spatial model.

We do not conclude, however, that party is irrelevant. Voting behavior changes fairly dramatically when members change parties. Party discipline, we conclude, is manifest in the location of the legislator's ideal point in the standard spatial model. It is not a strategic variable manipulated by party whips, but a part of a legislator's overall environment that forms her induced preferences. The legislator, in choosing a spatial location, may be responding as much to the external pressures of campaign donors and primary races as to the internal pressures of the party.

On the other hand, the evidence we presented does not suggest that a resurgence of party or party-induced institutional changes is responsible for the greater voting cohesiveness of parties and the emergence of polarized politics in Congress. Having distinct cutting lines (or separating hyperplanes) for the Democrats and Republicans never adds substantially to the classification success of the spatial model in the post World War II period. Indeed the additions have fallen throughout this period, both during the period of declining polarization (1947-circa 1975) and during the more recent surge in polarization.[24]

Appendices

Appendix A: Shifts in Ideal Points

Let z_{yj} and z_{nj} be the "yea" and "nay" outcomes of roll call j. In both the Heckman-Snyder and NOMINATE methods for estimating the spatial model, the non-random portion of the utility a legislator i has for roll call outcome $z_j \in \{z_{yj}, z_{nj}\}$, can be expressed as:

$$U_{ijz} = f\left(d_{ijz}^2\right) \tag{A1}$$

where f is a negative monotonic function and d_{ijz} denotes the Euclidean distance from x_i, i's ideal point, to z_j.

Now let the "party-pressured" ideological coordinates for Democrats equal $x_i + \gamma_j$. We obtain:

$$d_{ijz}^2 = \left(x_i + \gamma_j - z_j\right)^2 \tag{A2}$$

But this expression for distance is identical to the expression we would have if the ideal point were unchanged but the yea roll call outcome were changed to $z_{yj} - \gamma_j$. The distance to z_{nj} would also be unaffected if it were also changed to $z_{nj} - \gamma_j$. Shifting both roll call outcomes by γ_j also shifts the midpoint $\left(z_{yj} + z_{nj}\right)/2$ by γ_j. So, for example, a leftward shift in the ideal points for all Democrats is equivalent to a rightward shift in the outcome locations and midpoint for Democrats. The argument extends readily to multi-dimensional shifts. Since for every ideal point shift there is an equivalent outcome shift, neither Heckman-Snyder nor NOMINATE can discriminate between a model where a party alters ideal points on a roll call and one where each party has its own midpoint or separating hyperplane on each roll call.

Now consider the more general situation where the amount of pressure is not equal for all members but where the pressured ideal points maintain the same order as the original members and the magnitude of the pressure, for Democrats, is increasing in spatial position. Moderates are pressured more than liberals are. Since the pressure is not uniform, the shift in ideal points can no longer be captured by a simple shift in outcome locations. Nonetheless, in the map from the pressured ideal points back to the original ideal points, there will continue to be a point where a party member is indifferent between voting Yea and Nay. Let this point be the pressured midpoint for the party on the roll call. Optimal classification should be reasonably robust in identifying the pressured midpoint as long as the form of pressure does not depart too strongly from uniform pressure.

Appendix B: Monte Carlo Analysis of the Snyder-Groseclose Approach

To demonstrate that the Snyder-Groseclose method is likely to reject the null hypothesis of preference-based voting when it is true, we conduct a number of Monte Carlo experiments. The Monte Carlo data are generated by one-dimensional spatial voting with error. Snyder and Groseclose use the scaling method of Heckman and Snyder (1997). Our specification of the underlying random-utility model is therefore identical to that assumed by Heckman and Snyder. Legislator i votes **Y**ea rather than **N**ay if and only if

$$-\left(x_i - z_{yj}\right)^2 + \varepsilon_{ijy} \geq -\left(x_i - z_{nj}\right)^2 + \varepsilon_{ijn} \qquad \text{(A3)}$$

where x_i is the ideal point of legislator i, z_{yj} and z_{nj} are the positions of the Yea and Nay voting alternatives, and the ε are random shocks. Let $z_{Mj} = (z_{yj} + z_{nj})/2$ be the midpoint of the roll call and $d_j = (z_{yj} - z_{nj})/2$ be half the (directional) distance between the yea and nay outcomes. To simulate realistic values for the yea and nay positions, we assume that z_{Mj} is distributed on [-1,1] according to the density $f(z) = 1-|z|$ which produces a modal voting margin of 50%-50%. Further we assume that d_j is distributed uniformly on $[-1,-.05]\cup[.05,1]$. The "gap" from -.05 to .05 prevents votes in which the yea and nay outcomes are too similar so that voting is purely random.

We divide the 435 members of our House of Representatives into 218 Democrats and 217

Republicans. The ideal points of the Democrats are distributed uniformly across the interval $[-1,r]$ and those of the Republicans are distributed across $[-r,1]$. (See columns **(a)**, **(b)** of table A1.) The variable r controls the extent of overlap in the ideal points of members of the two parties. If $r = 0$, then the parties are perfectly spatially separated. If $r = 1$, the parties are drawn independently from the same distribution. In general, in expectation, a fraction $2r/(1+r)$ of each party overlaps with the other party. In our experiments, we let $r \in \{0.1, 0.2, 0.3\}$ so that the corresponding correlations between party and preferences take on the values of -0.82, -0.76, and -0.68. These are consistent with measures of party overlap and correlation in the post-war House of Representatives.

Also following Heckman and Snyder, we assume that

$$\eta_{ij} \equiv \varepsilon_{ijy} - \varepsilon_{ijn} \tag{A4}$$

is drawn from $U[-m,m]$. We let $m \in \{0.2, 0.4, 0.6\}$. (See column **(c)** of table A1.) These values are chosen to be consistent with the range of goodness-of-fit measures such as classification success reported in Heckman and Snyder (1997). In the experiments that follow, the correct classification of voting decisions following Heckman-Snyder estimation of the ideal points ranged from 81% to 92%. Finally, the experiments are conducted with 1000 roll calls.[25] This is roughly the number of actual roll calls in recent Houses.[26]

For each set of experiments, we produced two sets of estimates for ideal points using the Heckman-Snyder scaling method:

1. The Snyder-Groseclose estimates using only roll calls with margins greater than 65-35.
2. “Naïve” estimates using all the votes.[27]

Note that the Heckman-Snyder method should estimate ideal points very close to the true ideal points when the naïve model is used, because the Monte Carlo experiments generate the artificial data from a preference-based voting model.

Each specification was run ten times, so that 10,000 second-stage regressions were performed for each. In table A1, we present the percentage of times the null hypothesis of no party voting [that is, β_2=0 in equation (1)] was rejected at the 1% level (one-tail) using White's heteroscedasticity-consistent standard errors. Snyder and Groseclose also used the 1% criterion and White's standard errors, but whether the tests were one-tailed or two-tailed is unclear. That is, because preference-based voting generated the data, we expect to find a "significant" β_2 in only 1% of the simulated roll call votes. The actual results are strikingly different.

The extent of over-rejection for *close roll calls* (column **(d)**) is enormous. Under the most favorable conditions, shown in the last three rows of table A1 — large party overlap — the Snyder-Groseclose model rejects the null at approximately the expected 1% rate. However, in the least favorable conditions[28] — less overlap and precise voting — shown in the first row, the over-rejection rate is 73.1%.

The example shown in figure 2 indicates that the naïve method should lead to lower levels of rejection than the Snyder-Groseclose method because the better estimation of ideal points using all votes will leave less room for the party dummy to act as a proxy for the ideal points. This intuition is borne out. In all of the 9 matches of cells in table A1, the rejection rate for *close* roll calls (column **(f)**) is lower using the naïve method than using the Snyder-Groseclose method.[29] In the intermediate cases of rows 4-6 where Snyder-Groseclose rejects over 5 times the expected rate, the naïve method rejects at just about the expected rate.

One explanation for the Snyder-Groseclose bias on close roll calls is that the ideal points are recovered incorrectly as we argued with figure 2. Figure 2 was based on voting without errors. Errors in voting are not sufficient, even with large numbers of roll calls, to permit accurate recovery of legislator positions. Compare columns **(i)** and **(j)** of table A1. The correlations for the middle sixth of the legislature are systematically less using only close votes to estimate the ideal points than using all votes. That is, the effect we illustrated with figure 2 occurs even when both error is present and the number of roll calls is very large.

In many of our simulations the Heckman-Snyder estimates contradict the assumption of the underlying linear probability model. In particular, many of the voting probabilities lie outside the [0,1] interval. We indicate the percentage of these probabilities in column **(h)** of Table A.1. While it is true that the over-rejection rate is increasing in the number of improper probabilities, the variation in these proportions is too small to generate the large variation in the over-rejection rates. Furthermore, the percentage of extreme probabilities is approximately what one finds in applications of the Snyder-Groseclose method to actual roll calls from the House of Representatives.

At this point, the reader may have noticed an apparent anomaly. Under the naïve model, we should expect to get about 1% of the coefficients significant at the 1% level. The results are not too far off both for lopsided votes (column **(g)**) and for close votes (column **(f)**) where there is considerable party overlap. On the other hand, there are far too many significant coefficients for other close votes, particularly those in the first rows of the tables, where there is little overlap and only a small amount of randomness in voting.

There is an additional anomaly in our Monte Carlo experiments. For lopsided votes, both the näive and Snyder-Groseclose methods produce a large number of statistically significant coefficients at the one-tailed 1% level, but with the wrong sign.[30] Table A2 presents the percentage of "wrong" coefficients for the experiments on 1000 vote legislatures. Note that the problem is worst for the naïve model with little party overlap and lopsided votes.

Both of these anomalies arise because the Snyder-Groseclose second stage provides biased estimates of the party effect, *even when the ideal points have been correctly estimated in the first stage*. The intuition for both anomalies is provided by considering the case of a uniform distribution of ideal points on [-1,+1] with $r = 0$, no overlap. Moreover, assume, errorless voting, that is, $m = 0$. (And continue to assume no party pressure.)

Consider midpoints c in the interval [-1,+1]. A straightforward calculation shows that the coefficient on the party dummy is given by $\beta_2 = -1 + 4|c| - 3c^2$. Thus, β_2 is -1 for $c = 0$, the quintessential close 50-50 vote opposing Ds and Rs. From equation (2), the estimate of the extent of party discipline is $\gamma = -1/0 = -\infty$. On the other hand we get a wrong sign with $\beta_2 = 1/4$ for $c = 1/2$, that is, for a lopsided 75-25 split. More generally, the party coefficient is of the wrong sign for party-pressure voting when $1 > |c| > 1/3$. Although the coefficient should always be zero for preference-based voting, the coefficient is 0 only when the magnitude of c is exactly 1/3. When the magnitude of c is 1/3, we would get a 67-33 split. Like Snyder and Groseclose, we chose 65-35, very close to 67-33, to differentiate lopsided from close votes. The results in the tables, "correct" signs for close votes when the true coefficient should be zero and "wrong" signs for lopsided votes, conform to this theoretical analysis.

The theoretical example can be extended to allow for both overlap in party ideal points and for errors in voting. We focus on c=0, or predicted 50-50 splits, since this is the situation where Snyder and Groseclose expect the greatest party pressure. We begin by showing that allowing for overlap does not eliminate bias.

Introduce overlap in the party positions as follows. Let the left-most 25% of the legislature, those with ideal points in [–1, –½) be Democrats, the next 25% in [–½,0) be Republicans, then another 25%, in [0, ½), be Democrats and the rightmost 25%, in [½,1] be Republicans. On average, the Democrats are still the left, with a mean position of –¼ and the Republicans are at ¼. This is more overlap than appears in any Congress in the last two decades.

For this overlap case, the coefficient on the dummy is +6/13, showing an incorrect sign when voting is purely preference based.

In the no overlap example, the coefficient on the dummy was –1, indicating strong party pressure when there was none. Obviously, as the overlap increases the coefficient on the party dummy increases. There is some amount of overlap that will make the coefficient on the dummy 0, but this would be knife-edge.

Does error save the day? Yes and No. To see this, let the probability of voting Yes be linear in the ideal point between $-w$, $+w$, with

Prob(Yes vote|*x*£-*w*) = 0
Prob(Yes vote|$-w<x<w$) = .5+ $x/(2w)$
Prob(Yes vote|*x*³*w*) = 1.

Again assume a uniform distribution of (or equally spaced) ideal points and a legislature that is 50% D.

After calculating the appropriate variances and covariances and then plugging into the standard formula for the regression coefficient with two independent variables, we can compute values for the coefficient in both the overlap and no overlap cases. The results appear in Table A3.

In the no overlap case, we have a "correct" sign when the coefficient should be zero. In the overlap case, we have a "wrong" sign. The bias falls as the amount of error increases. For w of 0.8 or 0.9, which correspond to the error levels likely to occur with actual data, the bias is quite small. However, since the expected value is not zero, there still should be a disproportionate number of "significant" coefficients in reasonably large samples—such as the US House.

The no overlap case is more disturbing, since the bias does not fall as fast. With w of 0.8, the coefficient is -.04, which corresponds, in the example, to 2% of the legislators being switched by non-existent pressures. Thus, for no overlap or very low levels of overlap, one is quite likely to incorrectly conclude that there is some pressure when none exists.

More generally, the expected value of the party dummy coefficient, for a fixed non-random distribution of true ideal points and party affiliations, is a linear combination of the expected value of two covariances

$$E(\beta) = aE\left(Cov(vote, dummy)\right) + bE\left(Cov(vote, x)\right)$$

The linear coefficients depend on the variances of the ideal points and the party dummy and their covariance. The expected covariances depend on the error process, the distributions of the dummy and the ideal points, and the true cutting line for the roll call. Therefore, the sign and magnitude of the party dummy will depend in a complex way on both the distribution of ideal points and the distribution of errors. Only in special cases will the coefficient on the dummy have an expectation of zero when voting is based solely on spatial preferences and stochastic errors.

To illustrate how the patterns uncovered in the Monte Carlo experiments reappear in actual voting data, we replicate the analysis of Snyder and Groseclose for a number of Congresses.[31] Table A4 contains those results. In addition to the reported percentages of significant party coefficients, we also report the percentage of "wrong signs". Note that the pattern of the Monte Carlo experiments is echoed in the actual data. The number of "correct" significant coefficients is consistently higher for the Snyder-Groseclose model than the naïve model, and the number of "wrong" coefficients is higher for the naïve model. The differences are most striking with respect to correct signs on close votes. Parallel to the Monte Carlo work, the Snyder-Groseclose method produces many more significant instances of party discipline than does the naïve method.

In conclusion, our results demonstrate that the Snyder-Groseclose technique is heavily biased toward rejecting the null hypothesis of preference-based voting. Even if Snyder and Groseclose were able to estimate ideal points correctly in the first stage, they would get too many "significant" coefficients with the correct sign in the second stage on close votes and too many with the wrong sign on lopsided votes. The bias arises because they use OLS in the second stage. The bias is attenuated and becomes unimportant when there is a high degree of party overlap. Even when there is substantial party overlap, however, the Snyder-Groseclose method is biased toward finding too many "significant" coefficients because the first stage provides biased estimates of ideal points.

Appendix C: Procedure for Computing the Difference in Cutpoint Ranks

We used the following procedure for determining roll call cutpoints, classifying roll calls into the three categories, and computing the differences in ranks:

1. Optimally classify all legislators using a single cutpoint. Rank order the legislators from 1 to N, starting at the left.

2. Estimate the two cutpoint model for roll calls using the rank order of legislators from step "1". (Note that the estimation must "maintain polarity": Classification is optimal subject to making the same prediction for Ds and Rs to the left of their party's cutpoint.)

3. Every interior Democrat cutpoint must be between two Democratic legislators. Let their ranks be i and j. The rank of the roll call cutpoint is then given as $c_D = (i+j)/2$. When the cutpoint is to the right of the rightmost Democrat, denote the cutpoint by c_D=d_R=rank of rightmost Democrat. When the cutpoint is to the left of the leftmost Democrat, denote the cutpoint by 1. The Republicans are treated similarly; when the cutpoints is to the left of the leftmost Republican, denote the cutpoint by c_R=r_L=rank of leftmost Republican, to the right of the rightmost Republican, denote the cutpoint by N.

4. Score the roll calls as follows:

    a. If c_D=1 and c_R>r_L or if 1<c_D<d_R and c_R>r_L and c_D < c_R, score the roll call D<R.

b. If $1<c_D$ and $c_R<N$ and $c_D>c_R$, score the roll call D>R.

c. Otherwise, the roll call is "undecided".

5. For roll calls with interior cutpoints in both parties, the difference in ranks is c_D-c_R. Roll calls with one or more party cutpoints exterior are excluded from the difference in ranks computations (figure 8). (Thus, more roll calls are included in the ordinal comparisons under "4" above.)

References

Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Political Parties in America*. Chicago: University of Chicago Press.

Aldrich, John H., Mark M. Berger, and David W. Rohde. 1999. "The Historical Variability in Conditional Party Government, 1877-1986." Paper presented at the Conference of the History of Congress, Stanford University.

Cox, Gary, and Mathew McCubbins. 1993. *Legislative Leviathan*. Berkeley: University of California Press.

Doberman, Matthew C. 1997. *Congressional Responsiveness to Constituent Change*. Bachelor of Arts Honors Thesis. The Department of Government, Harvard College.

Fiorina, Morris P. 1999. "What Happened to the Median Voter?" paper presented at MIT conference on Political Parties in Congress.

Heckman, James, and James Snyder. 1997. "Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators." *Rand Journal of Economics*. Special Issue: S142-89.

Jenkins, Jeffrey. 1999. "Examining the Bonding Effects of Party: A Comparative Analysis of Roll Call Voting in the U.S. and Confederate Houses." *American Journal of Political Science* 43(4):1144-65.

Lowry, William R. and Charles R. Shipan. 2000. "The Causes of Party Ideological Differentiation in Congress," Typescript, University of Iowa.

King, David C. 1998. "Party Competition and Polarization in American Politics." Paper presented at the 1998 Midwest Political Science Association.

Krehbiel, Keith. 1993. "Where's the Party?" *British Journal of Political Science* 23(1): 235-66.

Krehbiel, Keith. 1998. *Pivotal Politics: A Theory of U.S. Lawmaking*. Chicago: University of Chicago

Press.

Levitt, Steven D. 1996. "How Do Senators Vote? Disentangling the Role of Voter Preferences, Party Affiliation, and Senator Ideology." *American Economic Review* 86(3): 425-41.

McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal. 1997. *Income Redistribution and the Realignment of American Politics*. Washington: AEI Press.

Nokken, Timothy P. 2000. "Dynamics of Congressional Loyalty: Party Defection and Roll Call Behavior, 1947-1997." *Legislative Studies Quarterly*, 25(3):417-444.

Poole, Keith T. 2000. "A Non-Parametric Unfolding of Binary Choice Data." *Political Analysis* 8(3):211-237.

Poole, Keith T. and Thomas Romer. 1993. "Ideology Shirking and Representation." *Public Choice* 77(Sept): 185-196.

Poole, Keith T., and Howard Rosenthal. 1984. "The Polarization of American Politics." *Journal of Politics* 46(4): 1061-79.

Poole, Keith T., and Howard Rosenthal. 1991. "Patterns of Congressional Voting." *American Journal of Political Science* 35(2): 228-278.

Poole, Keith T. and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York, Oxford University Press.

Rohde, David W. 1991. *Parties and Leaders in the Post-Reform House*. Chicago: University of Chicago Press.

Sinclair, Barbara. 1995. *Legislators, Leaders, and Lawmakers: The U.S. House of Representatives in the Post-Reform Era*. Baltimore: Johns Hopkins University Press.

Snyder, James and Timothy Groseclose. 2000. "Estimating Party Influence in Congressional Roll-Call Voting," *American Journal of Political Science* 44(2):193-211

Table 1. Correlations of Legislator Ideal Points from Optimal Classification Analyses

	Average Spearman Rank Order Correlations of Legislator Ideal Point Ranks (Standard Deviations in Parentheses)					
	Lopsided Vote Estimates vs. All Votes			**Close Vote Estimates vs. All Votes**		
Houses	Left 1/3	Middle 1/3	Right 1/3	Left 1/3	Middle 1/3	Right 1/3
80-90	.86	.44	.91	.58	.97	.51
(1947-68)	(.07)	(.30)	(.05)	(.13)	(.03)	(.17)
91-105	.94	.77	.94	.54	.97	.60
(1969-98)	(.01)	(.07)	(.07)	(.14)	(.02)	(.17)
80-105	.90	.63	.93	.56	.97	.56
(1947-98)	(.07)	(.26)	(.06)	(.14)	(.03)	(.17)

Note to Table 1. In each House, each 1/3 represents an N of at least 145. Actual N's are typically slightly larger because of deaths, replacements, etc. The averages are then computed as unweighted averages across the indicated set of Houses.

Table 2. Correlations of Legislator Ideal Points from One and Two Point Models

Average Within-Party Rank Order Correlations (Standard Deviations in Parentheses)		
Houses	Democrats	Republicans
80-90 (1947-68)	.94 (.06)	.93 (.03)
91-105 (1969-98)	.99 (.01)	.98 (.01)
80-105 (1947-98)	.97 (.05)	.96 (.03)

Note to table 2. The averages are unweighted averages across the indicated set of Houses.

Table 3. Order of Cutpoints on Close Roll Calls

House	D>R	Undecided	R>D
81-102	41.7%	2.7	55.6
80, 103-105	18.0%	40.2	41.7
All (80-105)	38.1%	8.5	53.4

Note to table 3. Entries are the percentages of close roll calls exhibiting the indicated pattern. E.g. 53.4% of all alocse roll call votes in the 80th through the 105th Congresses had an R>D pattern.

Table 4. Rank Order Positions of Legislators in Old and New Parties

Last Congress in Old Party	First Congress in New Party	Party Old	Party New	State	Name	Normalized Rank Old	Normalized Rank New
					Senate		
82	83	R	D	OR	Morse, W	.127	.004
88	89	D	R	SC	Thurmond, S	.988	.824
103	104	D	R	AL	Shelby, R	.439	.856
104	104	D	R	CO	Campbell, B	.355	.603
					House		
89	90	D	R	SC	Watson	.840	.860
92	93	R	D	NY	Reid, O	.085	.082
93	94	D	R	OK	Jarman, J	.517	.796
94	96	R	D	NY	Peyser	.269	.172
96	97	D	R	PA	Atkinson, E	.473	.539
97	98	D	R	AZ	Stump, B	.898	.991
97	98	D	R	TX	Gramm, P	.918	.969
98	99	D	R	FL	Ireland, A	.495	.884
100	101	D	R	FL	Grant, B	.442	.641
100	101	D	R	AR	Robinson, T	.486	.659
104	104	D	R	LA	Hayes, J	.480	.890
104	104	D	R	TX	Laughlin	.455	.881
104	104	D	R	MS	Parker	.500	.876
104	104	D	R	GA	Deal	.474	.925
104	104	D	R	LA	Tauzin	.506	.872

Note: See text for details of computation of normalized ranks.

Table A1. Preference-Based Monte Carlo Data for Legislatures with 435 Legislators, 1000 Roll Calls, 10 Replications

Distribution of Ideal Points		Voting Error	Ideal Points Estimated by Heckman-Snyder Method Applied to: Close Votes Only (Snyder/Groseclose Method) — Percentage of 10000 Roll Calls With Party Pressure Effect Significant at 1% Level		All Votes (Naïve Method)		Close Votes — %Vote Probs. 0utside [0,1]	Lopsided Votes — Correlation of True and Estimated Ideal Points, Middle Sixth	All Votes	Optimal Classification — % Gain in Correct Classification, Two Cutpoint Model Over One Cutpoint Model
Rep. (a)	Dem. (b)	(c)	Close Votes (d)	Lopsided Votes (e)	Close Votes (f)	Lopsided Votes (g)	(h)	(i)	(j)	(k)
U[-.1, 1]	U[-1, .1]	U[-.2,.2]	73.1	0.2	19.1	0.8	24	0.66	0.97	.10
U[-.1, 1]	U[-1, .1]	U[-.4,.4]	47.8	0.5	9.4	1.8	21	0.80	0.94	.16
U[-.1, 1]	U[-1, .1]	U[-.6,.6]	23.4	0.6	4.7	1.5	17	0.82	0.90	.23
U[-.2,1]	U[-1,.2]	U[-.2,.2]	34.0	0.4	0.7	1.2	24	0.79	0.98	.16
U[-.2,1]	U[-1,.2]	U[-.4,.4]	13.1	0.6	0.9	0.9	20	0.86	0.95	.24
U[-.2,1]	U[-1,.2]	U[-.6,.6]	5.8	0.9	1.3	1.5	16	0.88	0.92	.30
U[-.3,1]	U[-1,.3]	U[-.2,.2]	0.6	0.7	0.5	15.7	23	0.82	0.98	.21
U[-.3,1]	U[-1,.3]	U[-.4,.4]	0.7	0.9	0.5	3.2	19	0.89	0.96	.30
U[-.3,1]	U[-1,.3]	U[-.6,.6]	1.2	0.9	0.9	1.3	16	0.87	0.93	.36

Table A2. "Wrong Sign" Party Coefficients from Application of the Snyder-Groseclose Second Stage to Preference-Based Monte Carlo Data

Distribution of Ideal Points		**Voting Error**	**Percentage of 10000 Roll Calls With "Wrong" Party Pressure Effect Significant at 1% Level**			
			Snyder / Groseclose Method in First Stage		**Naïve Method in First Stage**	
Republicans	Democrats		Close Votes	Lopsided Votes	Close Votes	Lopsided Votes
U[-.1, 1]	U[-1, .1]	U[-.2,.2]	13.2	38.1	0.0	56.1
U[-.1, 1]	U[-1, .1]	U[-.4,.4]	7.2	30.5	0.1	42.8
U[-.1, 1]	U[-1, .1]	U[-.2,.6]	3.7	21.6	0.3	29.1
U[-.2,1]	U[-1,.2]	U[-.2,.2]	7.1	33.0	0.4	10.6
U[-.2,1]	U[-1,.2]	U[-.4,.4]	2.4	20.6	0.2	10.7
U[-.2,1]	U[-1,.2]	U[-.6,.6]	1.2	10.8	0.5	7.5
U[-.3,1]	U[-1,.3]	U[-.2,.2]	0.8	14.1	9.8	0.1
U[-.3,1]	U[-1,.3]	U[-.4,.4]	0.6	4.3	1.4	0.2
U[-.3,1]	U[-1,.3]	U[-.6,.6]	0.9	1.7	0.9	0.8

Table A.3 Value of Party Dummy Coefficient (b_2) When There is No Party Pressure

	Party Coefficient	
w	Overlap	No Overlap
0.0	0.4615	-1.0000
0.1	0.3985	-0.8100
0.2	0.3323	-0.6400
0.3	0.2631	-0.4900
0.4	0.1908	-0.3600
0.5	0.1155	-0.2500
0.6	0.0574	-0.1600
0.7	0.0257	-0.0900
0.8	0.0092	-0.0400
0.9	0.0019	-0.0100
1.0	0.0000	0.0000

Table A4. Replication of Snyder-Groseclose Using Six Dimensional Model

House	*Number of Roll Calls*	*Number of Close Votes*	*Proportion Significan. at 1%, Correct Sign, Close Votes*	*Proportion Significant. at 1%, Correct Sign, Lopsided Votes*	*Proportion Significant at 1%, Wrong Sign, Close Votes*	*Proportion Significant at 1%, Wrong Sign, Lopsided Votes*
85th (Naïve)	175	99	.101	.013	.121	.158
85th (S-G)	175	99	.394	.039	.141	.144
90th (Naïve)	409	158	.203	.060	.127	.068
90th (S-G)	409	158	.601	.056	.025	.143
95th (Naïve)	1348	649	.068	.064	.085	.043
95th (S-G)	1348	649	.317	.019	.045	.081
100th(Naïve)	803	335	.143	.049	.197	.246
100th (S-G)	803	335	.394	.077	.107	.098
105th(Naïve)	550	275	.000	.015	.021	.113
105th (S-G)	550	275	.767	.127	.054	.189
Total (Naïve)	3285	1516	.088	.050	.105	.116
Total (S-G)	3285	1516	.450	.057	.065	.114

Figure 1. A Six-Member Legislature

	Left					Right
Legislator	1	2	3	4	5	6
Party	D	D	D	R	R	R

Figure 2. Perfect Spatial Voting in a Six Member Legislature

	Legislator					
	Left					**Right**
	1	**2**	**3**	**4**	**5**	**6**
Vote on Roll Call:						
1	Y	Y	Y	Y	Y	Y
2	N	N	N	N	N	N
3	Y	N	N	N	N	N
4	N	Y	Y	Y	Y	Y
5	Y	Y	N	N	N	N
6	N	N	Y	Y	Y	Y
7	Y	Y	Y	Y	N	N
8	N	N	N	N	Y	Y
9	Y	Y	Y	Y	Y	N
10	N	N	N	N	N	Y
11	Y	Y	Y	N	N	N
12	N	N	N	Y	Y	Y

Note: Y=Yea Vote, N=Nay Vote

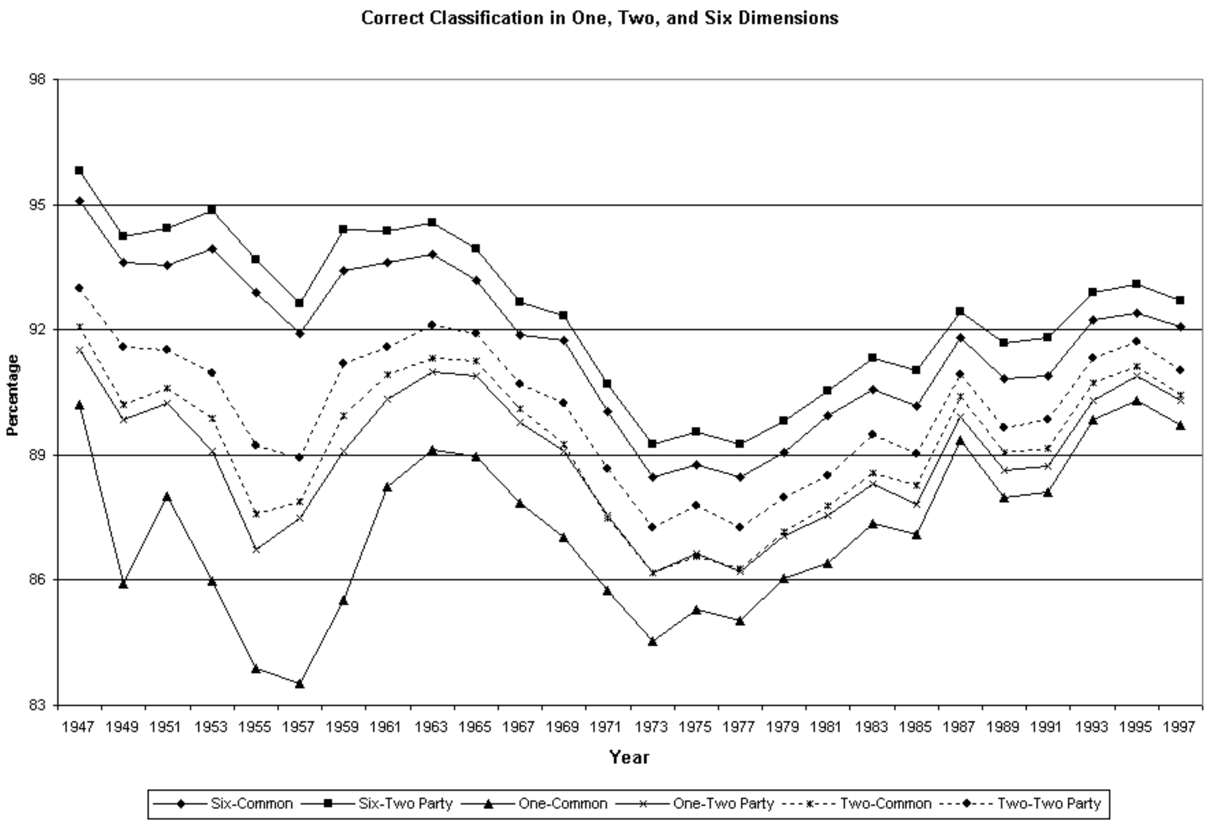
Figure 3. Preference Order Based on Lopsided Votes

			4		
Legislator	1	2	3	5	6

Figure 4. Cutpoint Models

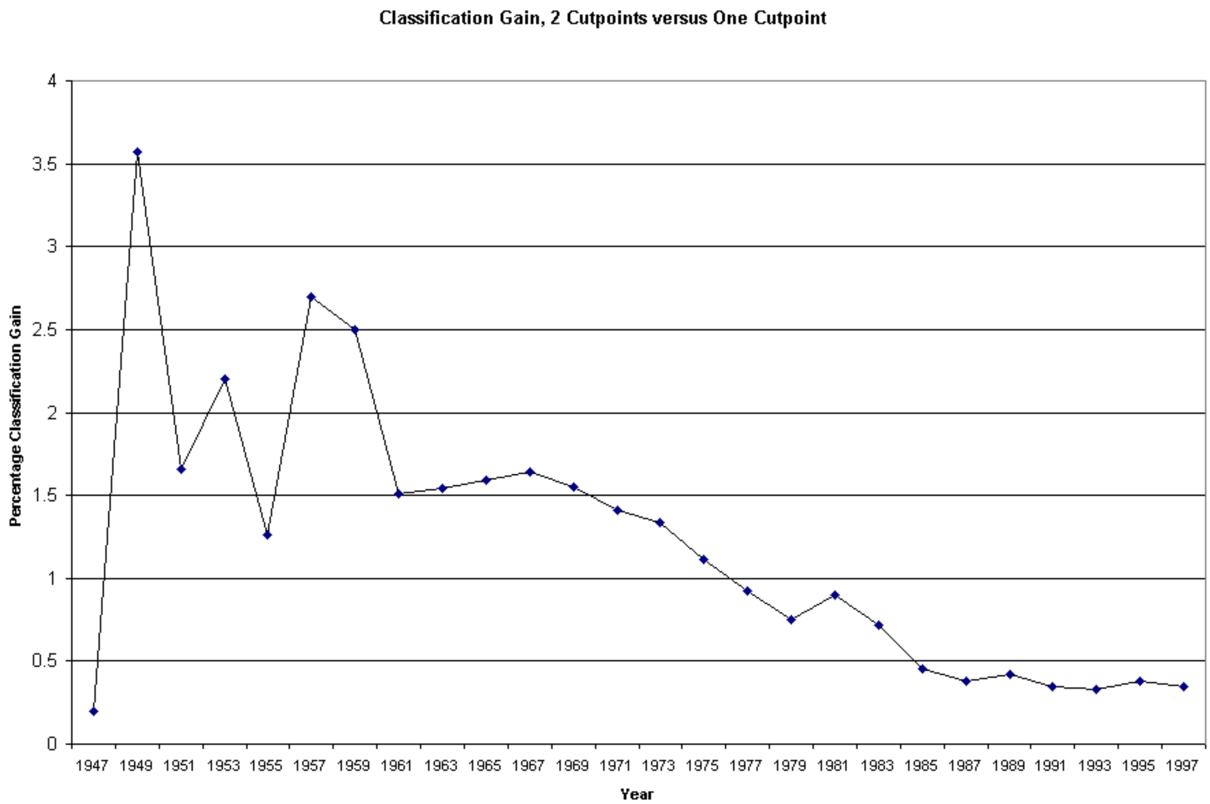
Legislator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Party	D	D	D	D	R	D	R	D	R	R	D	D	R	R	R	R	R
Vote	Y	Y	N	Y	Y	Y	Y	Y	N	N	Y	N	N	N	Y	N	N
Common Cutpoint								*Predicted Yea*	*Predicted Nay*								
Rep. Cutpoint								*Predicted Yea*	*Predicted Nay*								
Dem. Cutpoint											*Predicted Yea*	*Predicted Nay*					

Figure 5.



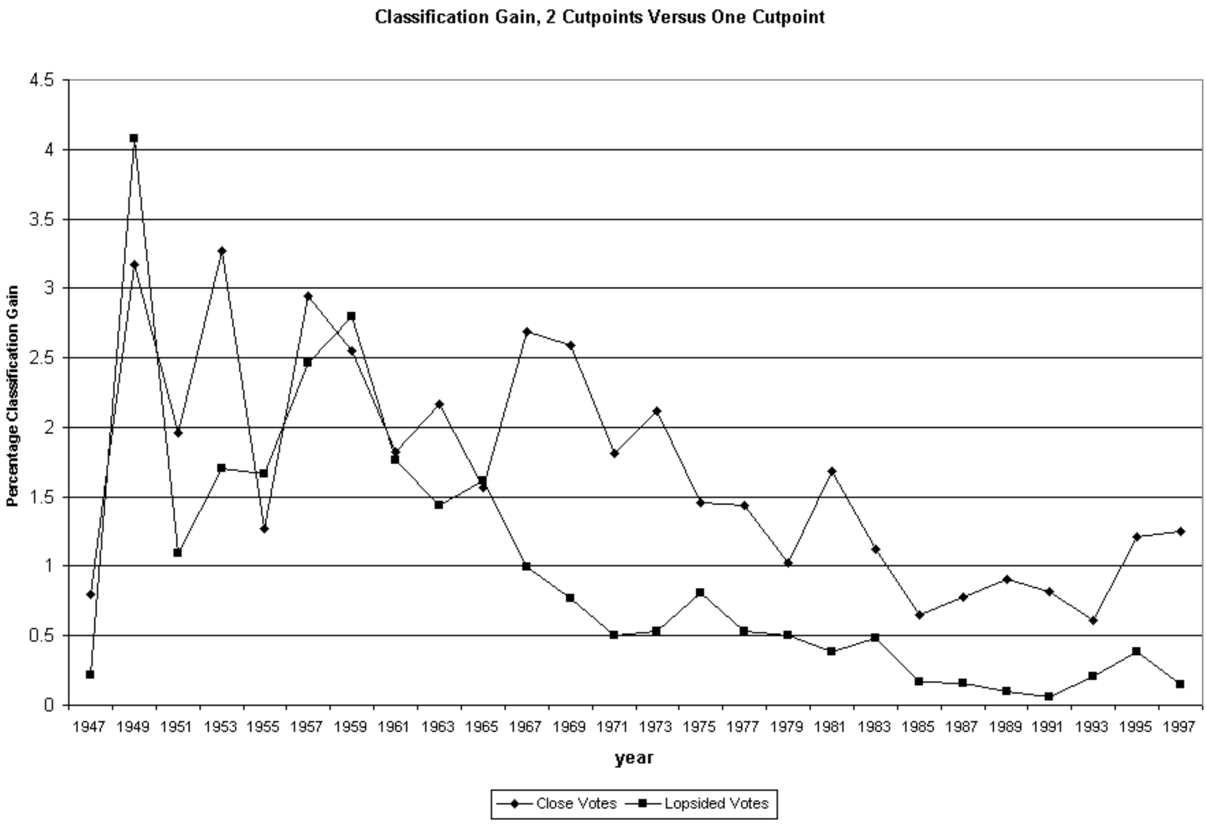
Note to figure 5. “Common” refers to optimal classification when all representatives are scaled together, “2-party” to optimal classification of each party separately. The total classifications for the two cases are equal. The percentage correct for the “2-party” must exceed the percentage correct for the “common”.

Figure 6.



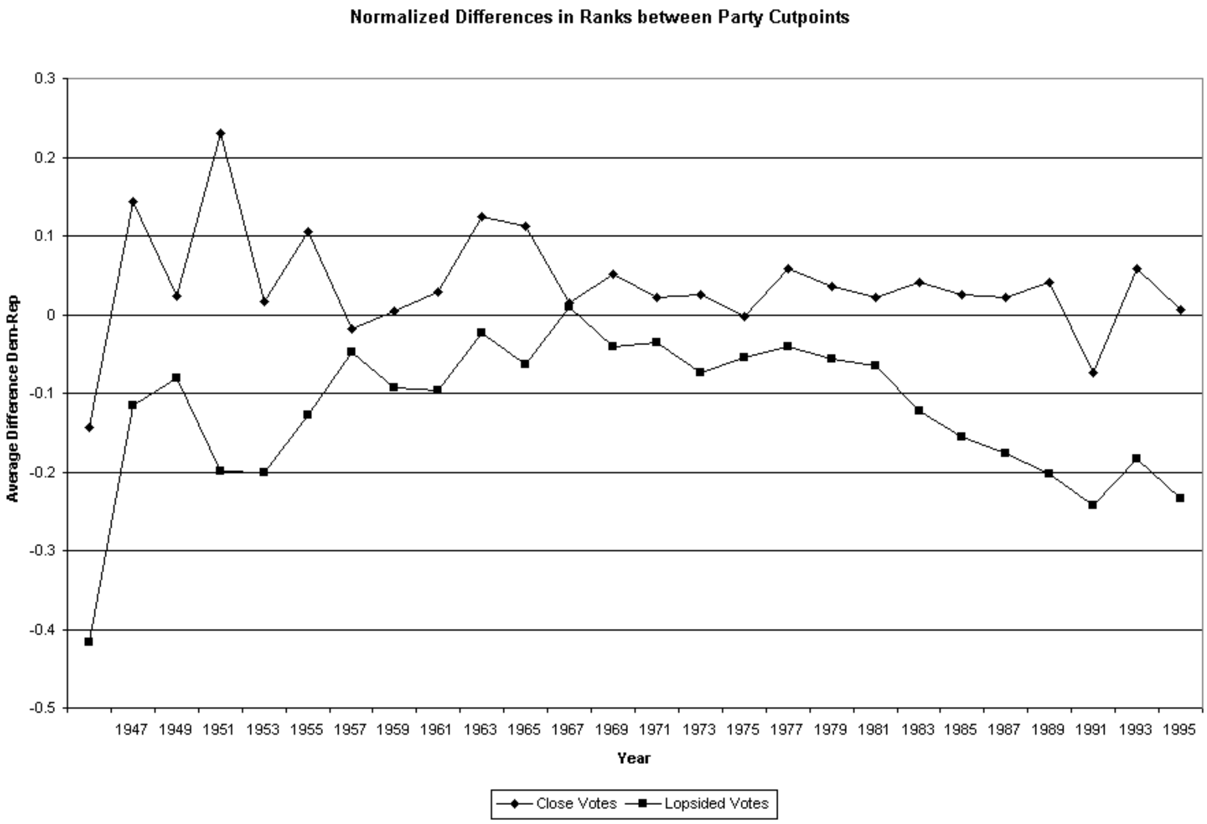
Note to figure 6. . The classification gains are for a one-dimensional voting model. All representatives were scaled together, as in the "Common" scaling of figure 4. With the ideal point orders from the "Common" scaling held fixed, a separate cutpoint was then estimated for each party. Comparison to figure 4 shows that the classification gains are similar to those in the "2-party" scalings where each party has an independent rank order of ideal points as well as a separate cutpoint.

Figure 7



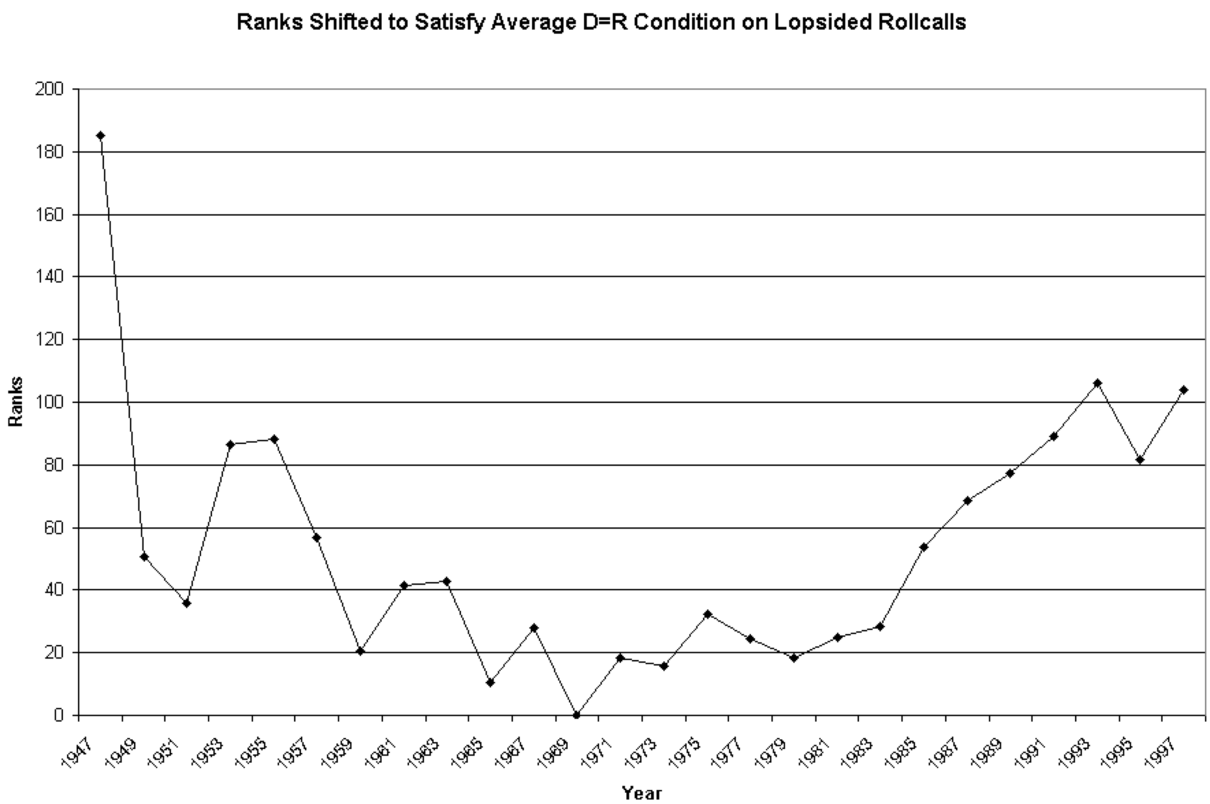
Note to figure 7. The gains from the common scaling(see note to figure 6) have been broken down into those for close votes and those for lopsided votes.

Figure 8.



Note to figure 8. For each House, the rank orders were normalized to run from 0 to 1. For example, if 438 legislators served in a House, the rank order was normalized to 0/437… 437/437. The "average difference" is the average of the differences between the normalized rank of the Democrat cutpoint and the normalized rank of the Republican cutpoint.

Figure 9.


Ranks Shifted to Satisfy Average D=R Condition on Lopsided Rollcalls
200
180
160
140
120
100
80
60
40
20
0
Ranks
1947
1949
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
Year

Notes

Poole's research was supported by NSF grant SBR-973035. Rosenthal's research was supported by NSF grant SBR-973053. This paper was written while Rosenthal was a fellow of the Center for Advanced Study in the Behavioral Sciences. Rosenthal's work at the Center was also supported by NSF grant SBR-9022192. We thank Kathleen Much and Elspeth Wilson for editing and Larry Bartels, Tim Groseclose, Keith Krehbiel, Jim Snyder, and seminar participants at George Washington, Princeton, Stanford, Rochester for comments and suggestions.

* Columbia University
** University of Houston
*** Princeton University

[1] Cohesiveness is not the only focus, or even the most active focus, of this line of inquiry. Much of the work has focused on the role of party leaders in setting the legislative agenda.

[2] These authors use the Poole and Rosenthal (1991, 1997) NOMINATE methods. Both NOMINATE and the Heckman and Snyder (1997) method are parametric. The results of the two methods are very similar, particularly on the first and second dimensions.

[3] The spatial model does, however, strongly outperform a model of straight party-line voting.

[4] See Jenkins (1999) for an application of this method.

[5] Snyder and Groseclose (2000) allow the displacement to be other than a constant, but their empirical work relies on the simple constant displacement model. They also allow for multiple dimensions, but the unidimensional case gives the intuition of their more general model.

[6] Poole and Rosenthal (1997, pp. 155-157) document that there are very few "both ends against the middle" votes where extremists defect.

[7] The predicted order of cutpoints is equivalent to the prediction that γ is negative.

[8] Although the underlying assumptions are very different, in one dimension this method is essentially

equivalent to classical Guttman scaling.

[9] This focus is consistent with a key point of Krehbiel (1998). He argues that the main empirical question should not be whether parties have influence on legislative behavior, but whether partisan models represent a substantial improvement over those that assume autonomous legislators.

[10] Pressure beyond that necessary to generate a strict party line vote cannot further increase classification.

[11] We focus on the results using rank-order correlations since they are most consistent with our optimal classification approach. However, we have conducted each of these experiments using standard correlations and have found there to be little substantive difference.

[12] Thus there would be overwhelming statistical significance using a simple sign test for the observation of 25 successes in 26 trials.

[13] Since prediction M1 may be consistent with either party or preference voting, we generated Monte Carlo data imposing preference voting without party discipline. These results listed in columns i and j of Table A1 show that under pure preference voting the correlation between the true and estimated preferences are lower when only lopsided votes are used.

[14] An improvement of 0.5% may well be statistically significant. In column **(k)** of table A1, which shows simulations for preference-based voting, we show classification gains for various one-dimensional specifications. In the first three, low overlap rows, similar to actual overlap in the past 8 Congresses, the gains range from 0.10% to 0.26%, all considerably less than 0.5%. Of course, the gains from "fitting" an extra hyperplane in a multi-dimensional model would be expected to be even higher. In any event, a 0.5% improvement may lack substantive import.

[15] In appendix C, we outline our procedure for determining roll call cutpoints, classifying roll calls into the three categories, and computing the differences in ranks.

[16] In addition, the ordinal comparisons involve some roll calls with exterior cutpoints. See Appendix C.

[17] We thank Tim Groseclose for suggesting the adjustment. The algorithm we developed to implement the suggestion is as follows. If the average difference in ranks for *lopsided* roll calls is non-negative, no shift is required. Otherwise, shift every Republican leftward by a number of ranks equal to the average difference in ranks. This procedure implicitly assumes that the ranks are interval measurements. By shifting the Republicans leftward, we are compressing the space. In the original estimates, the unnormalized space will extend from 1 to N, where N is the number of scaled legislators in the House. In the shifted estimates, the space will run from 1 to N-A, where A is the number of ranks shifted.

[18] In an earlier version of this paper, we used the McCarty, Poole, and Rosenthal (1997) DW-NOMINATE procedure for estimation to obtain metric estimates of the magnitude of changes induced by party switching. The metric assumptions in the NOMINATE procedure lead to sensible results--for example, there is less distance between the median and the 9th decile in the Gingrich Houses than between the 1st decile and the median. We conducted the metric analysis in two dimensions. Switchers from R to D were expected to become more negative on the first dimension and more positive on the second and vice-versa for D to R switchers. All movement on both dimensions was in accord with the hypothesis. For more details, the paper can be accessed at http://porkrind.pols.columbia.edu/discip.pdf.

[19] These results are consistent with Nokken (2000), who also finds significant changes in congressional behavior following a party switch.

[20] The representatives selected were Boland (D-MA), Johnson (D-CA), J. Melcher (D-MT), Button (R-NY), Fallon (D-MD), Traficant (D-OH), Matthews (D-FL), Morella (R-MD), Fountain (D-NC), Taft (R-OH), Lloyd (R-UT), Kasten (R-WI), Haley (D-FL), T. Corcoran (R-IL), and Zion (R-IN).

[21] On the other hand, studies show that constituency changes do not have much impact on the ideal points of legislators. See Poole and Romer (1993) for House redistricting and Doberman (1997) for House members who moved to the Senate.

[22] Levitt (1996) provides some indication of the relative effects of party vs. constituency factors in determining the ideal points of switchers. For the Senate, Levitt models each senator's ideal point (as proxied by ADA rating) as a weighted average of personal ideology, overall state characteristics, support group characteristics, and the "national party line". While all four of these factors might change for switchers, the main changes are likely to be in the new national party line and in the new support group that is relevant to campaign funding and primaries. Levitt's results put about equal weight on these two factors. Consequently, about half of the change in the ideal point would reflect forces internal to Congress.

[23] We thank Larry Bartels for this observation.

[24] On polarization in Congress, see Fiorina (1999), King (1998), Lowry and Shipan (2000), McCarty, Poole and Rosenthal (1997), and Poole and Rosenthal (1984, 1997, 1999).

[25] Legislatures with 500 roll calls were also examined, but the results are very similar. These may be found in an earlier version of this paper available at http://porkrind.pols.columbia.edu/discip.pdf.

[26] The setup of these Monte Carlo experiments is very similar to the Monte Carlo experiments reported in the published version of Snyder and Groseclose (2000) which were conducted in response to our original working paper. The major difference is that Snyder and Groseclose impose restrictions on the distribution of voting error to eliminate probabilities outside the [0,1] interval in their linear probability setup. This arbitrary assumption makes their results a best case for their model by assuming away one of the sources of mistaken inferences that we identify below.

[27] Like Heckman and Snyder (1997), we excluded all votes with less than 1% voting on the minority side.

[28] McCarty, Poole, and Rosenthal (1997) present evidence that these conditions are likely to prevail in recent Congresses.

[29] One might argue that in the case of large party overlap, the naïve model somewhat under rejects the null.

[30] That is, the party coefficient implies that Democrats are under pressure to vote in the conservative way. We coded the coefficients "wrong" if β_1 and β_2 had the same sign (recall preferences are scaled so that conservatives score higher). In the case of quadratic preferences, this is equivalent to a finding of $\gamma > 0$. We found that explicitly testing the hypothesis $\gamma < 0$ produced results substantively similar to coding the expected sign of β_2 based on the sign of β_1.

[31] Our replications appear to match their results with a few caveats. First, Snyder and Groseclose present their results in a line-graph so verifying an exact match is impossible. Second, they do not indicate how wrong -signed coefficients were treated or how many tails were used in their hypothesis tests. Finally, another potential factor for discrepancy is that they do not indicate the dimensionality of the preference model they used on each roll call. They outline a Monte Carlo procedure for determining the right number of dimensions to retain and indicate they retained a "few more" that this number. Rather than replicate this analysis (which would be imperfect due to the use of a Monte Carlo test statistic), we included six dimensions which is approximately the average used by Snyder and Groseclose.