Bridging Institutions and Time: Creating Comparable Preference Estimates for Presidents, Senators, Representatives and Justices, 1950-2002*

Michael Bailey
Georgetown University

Working Draft

July 2005

Abstract

Difficulty in comparing preferences across time and institutional contexts hinders the empirical testing of many important theories in political science. In this paper, I characterize these difficulties and provide a measurement approach that relies on inter-temporal and inter-institutional "bridge" observations and Bayesian Markov chain simulation methods. I generate preference estimates for Presidents, Senators, Representatives and Supreme Court Justices that are comparable across time and across institutions. Such preference estimates are indispensable in a variety of important research projects, including research on statutory interpretation, executive influence on the Supreme Court and Senate influence on court appointments.

*Prepared for the Summer Meeting of the Society for Political Methodology at Florida State University. I acknowledge the helpful support of the National Science Foundation (SES-0351469). Keith Poole generously provided congressional and presidential vote data for this project. I am also grateful to Kelly Chang, Eric Lawrence, Forrest Maltzman, Barry Weingast and participants at seminars at Washington University in St. Louis, New York University Law School, Dartmouth George Washington and Georgetown for very helpful comments. I also appreciate excellent research assistance from Laura Miller, Chris Drewry, Stephen de Man, Caroline Wells, Mike Griffin, Alexis Teagarden, Cynthia Fleming and Matt Hard. Comments welcome to baileyma@georgetown.edu. Analysis is provisional, as data collection is still underway.

Accurate measurement is central to theory testing in political science. Institutionalist scholars claiming that certain political actors or institutions influence outcomes need to accurately characterize who wanted what and how they did or did not get it. For example, scholars testing theories of congressional influence on the court (see, e.g., Ferejohn and Weingast 1992; Segal 1997; Moraski and Shipan 1999), need to show what members of Congress desired with regard to the ideology or decisions of Supreme Court Justices. Behaviorists claiming connections between societal and political phenomena need to characterize outcomes at different times in a comparable manner. For example, researchers investigating the influence of public opinion need to track preference and output measures over time (see, e.g., Stimson, Mackuen and Erikson 1995; Mishler and Sheehan 1992).

These kinds of comparisons are challenging because it is difficult to disentangle agenda differences from preference differences. For example, scholars comparing preferences across institutions must confront the fact that agenda differences across institutions make it impossible to directly compare standard preference measures: in short, even the best measures of congressional preferences based only on votes in Congress are not directly comparable to even the best measures of judicial preferences based only on Supreme Court votes. Or, scholars interested in comparing preferences across time worry about mistaking agenda change for preference change: if they observe an increase in conservative votes, is it it because preferences have shifted or because the agenda has shifted?

If researchers do not address these issues directly, they risk compromising their empirical tests with distorted preference comparisons. This is more than a technical theoretical concern: I show below that estimates produced by some of the best methods in the litera-

ture imply that pro-segregationist southern Senators in the 1950s had the same preferences as moderate Democrats in the 1990s and that the Supreme Court that produced the *Roe* decision was one of the most *conservative* courts of the modern era.

This paper addresses these problems by using "bridging" data that allows us to make inter-institutional and inter-temporal comparisons. To bridge across institutions, I incorporate an extensive and substantially original data set of observations of presidents and members of Congress taking positions on Supreme Court cases. To bridge across time, I incorporate information on the movement of the agenda across time and observations of individuals taking positions on cases and votes in the past. I use flexible and powerful Bayesian Markov chain simulation statistical methods (Albert and Chib 1993; Martin and Quinn 2002; Clinton, Jackman and Rivers 2004) to generate preference estimates for Presidents, Senators, Representatives and Supreme Court Justices from 1950 to 2002.

This paper proceeds as follows. Section 1 discusses the need for, and difficulty of, generating preference estimates that are comparable across institutions and time. Section 2 presents the estimation routine. Section 3 discusses the original data collection. Section 4 presents the results.

1 Comparing Preferences across Institutions and Time

Not long ago scholars were content to use raw interest group scores and percent liberal scores to measure congressional and judicial preferences. Baum (1988), Snyder (1992) and others identified devastating flaws in such measures. Poole and Rosenthal's (1991; 1997) work provided a major breakthrough by providing a serious, model-based approach to measurement

of preferences for Congress and the president. Others provide additional approaches for legislatures (Heckman and Snyder 1997; Clinton, Jackman and Rivers 2004). Bailey and Chang (2001) and Martin and Quinn (2002) provide model-based methods and estimates for the court.

The modern preference models are based on spatial theory in which the preferences of an individual are defined by a single-peaked utility function that is maximized at the individual's "ideal point." Policy alternatives are defined as points in that preference space. When given a choice among alternatives, the individual is assumed to choose the alternative that is spatially closer. Much of the literature, including this paper, focuses on one-dimensional models. For example, a judge with an ideal point of 10 ruling on a case in which ruling for the plaintiff yields an outcome of 2 and ruling for the respondent yields an outcome of 11 will rule for the respondent. Votes are often discussed in terms of the midpoint between the two alternatives, as individuals with ideal points above the midpoint will vote for the higher alternative and individuals with ideal points lower than the midpoint will vote for the lower alternative. I refer to this midpoint as a vote "cutpoint."

1.1 Comparing preferences across institutions

One challenge in measuring these spatial preferences is comparing preferences across institutions. Simply put, no matter how well preferences are estimated within an institution, they are not comparable across institutions without clear points of reference. Figure 1 makes this point graphically. At the top is a hypothetical court with judges with ideal points at J_1, J_2 and J_3. This court ruled on two cases: in the first, the cutpoint between the two alternatives

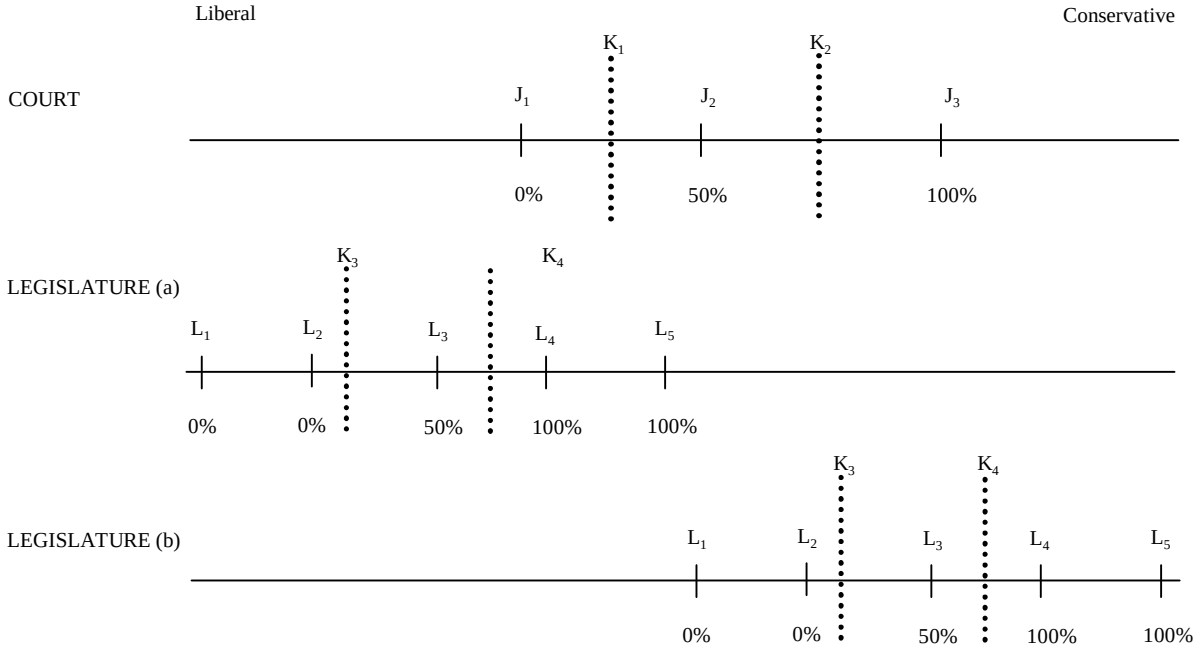

Figure 1: The challenge of making inter-institutional preference comparisons

was K_1; in the second, the cutpoint was K_2. The ideal point of the most conservative judge (J_3) is greater than both cutpoints, leading the judge to vote conservatively 100 percent of the time. The ideal point of the median judge (J_2) is greater than K_1, but less than K_2, leading the judge to vote liberally once and conservatively once. The ideal point of the most liberal judge is less than both cutpoints, meaning he or she voted conservatively 0 percent of the time.

In the next two panels are hypothetical legislatures with legislators with ideal points at L_1 through L_5. The legislature voted on legislation that had cutpoints at K_3 and K_4. The two most conservative legislators had ideal points above both cutpoints, implying a 100 percent conservative voting record. The median legislator's ideal point was higher than the first cutpoint (implying one conservative vote) and lower than the second cutpoint (implying

one liberal vote) producing a 50 percent conservative voting record. The two most liberal legislators' ideal points were below both cutpoints, implying a 0 percent conservative voting record.

The problem is that based simply on voting patterns within the two institutions, we cannot know whether the depiction in the middle panel (where the legislative median is far to the left of the court median) or the bottom panel (where the legislative median is far to the right of the court median) is correct.[1] Without knowing the location of K_1 and K_2 relative to K_3 and K_4, either depiction is logically possible, even as they differ dramatically from each other. More sophisticated within-institution preference estimation may produce better comparisons within an institution, but will not solve this fundamental inter-institutional problem.

Existing inter-institutional analyses address this problem in an ad hoc manner, usually by assuming single-institution scores are directly comparable. Segal (1997) assumed that preference measures for Justices (based on Segal-Cover scores) were directly comparable to ADA scores.[2] Moraski and Shipan (1999) assumed that percent-liberal scores for justices were directly comparable to adjusted ADA scores from Groseclose, Levitt and Snyder (1999). These assumptions are, tacitly, assumptions that the distribution of votes facing the senate and courts are identical.

In order to illustrate the practical problem with ad hoc assumptions of direct compara-

[1] Consider the following extreme example. Suppose that the agenda of one institution consists of very conservative proposals and that only a very conservative individual would vote conservatively on half of the proposals (imagine, for example, voting on platform proposals at the Republican National Convention). If one compares that to another institution that faces very liberal proposals (imagine voting on the platform proposals at the Democratic National Convention), there could be a situation in which a moderate would vote nearly zero percent conservative in one context and nearly 100 percent conservative in the other. Comparing percent liberal scores across these contexts would be meaningless.

[2] Segal is well aware of the problem. In fact, his clear statement of the problem is one of the motivations of this project.

bility, let us compare two plausible versions of this approach. The first is to assume percent liberal judicial scores are comparable to ADA scores. The second is to assume that Poole and Rosenthal Common Space scores for senators and presidents are comparable to Martin and Quinn scores for justices. (This approach to inter- institutional preference measurement is not endorsed by Poole, Rosenthal, Martin or Quinn; I am presenting it simply as an example of assuming direct comparability across institutions). Figure 2 presents the estimated preferences of the court median, senate median and the president for these two approaches over time.[3]

Even casual observation suggests stark differences between the two approaches. In the fifties, the ADA measure indicates the court was the most liberal and the senate was the most conservative. The Poole and Rosenthal/Martin and Quinn (PRMQ) measure, on the other hand, has the opposite, with the senate the most liberal and the court the most conservative. In the early seventies and again in the eighties, the ADA measure indicates the court was more liberal than the Senate while the PRMQ measure indicates the senate was substantially more liberal than the court. In the nineties, the ADA measure indicates the court median was between the president and senate median, while the PRMQ measure indicates the court median was substantially to the right of both. The point is that scholars who use an ad hoc approach to inter-institutional preference measurement are entering highly risky territory. The analysis differs dramatically depending on which ad hoc approach they use. It is in precisely this kind of situation in which we need a carefully modeled statistical approach.

[3]The ADA and percent liberal measures have been subtracted from one in order to give them the same ideological polarity as the PRMQ measures.

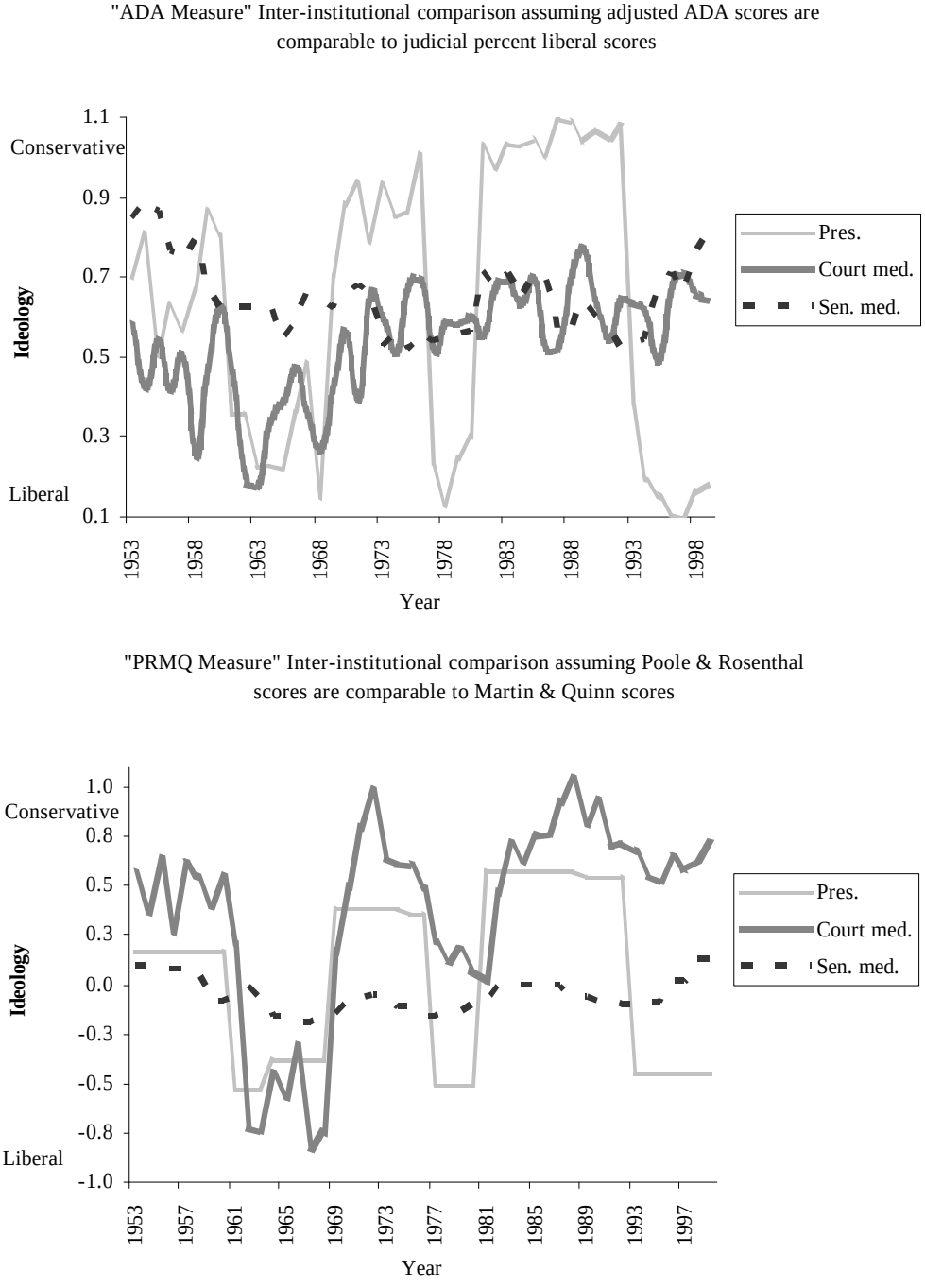

Figure 2: Inter-institutional comparisons assuming direct comparability

1.2 Comparing preferences over time

The other major challenge facing preference estimation is generating measures of preferences that are valid over time. Scholars frequently care directly about evolution of preferences over time, as with studies of opinion and policy (see, e.g., Stimson, Mackuen and Erikson 1995; Mishler and Sheehan 1992) or they require preferences as either key independent variables or important control variables in studies that span several decades (such as in studies of the control of the bureaucracy or of legislative output). The problem is that agendas change – often times endogenously (Mouw and Mackuen 1992)– making it very hard to disentangle potentially changing preferences from potentially changing agendas.

To explicate this challenge, suppose we wish to compare the conservativism of two courts: one that voted seven to two in favor of a liberal outcome on *Roe v. Wade* 410 U.S. 113 (1973) and another that voted five to four in favor of a conservative outcome on *Webster v. Reproductive Health* 492 U.S. 490 (1989) (allowing Missouri to ban abortions in public facilities, prohibit abortion counseling and require fetal viability tests for women seeking abortions who are more than twenty weeks pregnant). One possibility is that court was indeed more conservative on *Webster*. Another possibility is that the case content shifted, such that even the *Roe* Court would have voted as the *Webster* court.

Figure 3 presents a hypothetical depiction of preferences of the justices on each case.[4] In the top panel, the seven justices who voted liberally on *Roe* are to the left of the cutpoint and the two who voted conservatively are on the right. In the next two panels are hypothetical depictions of the preferences of the justices who voted on *Webster* case. Again, those who

[4] For simplicity, we will consider only one case and assume a deterministic model of voting; the logic extends to a situation in which we look at many cases in one statistical model and allow for randomness in case-specific ideal points.

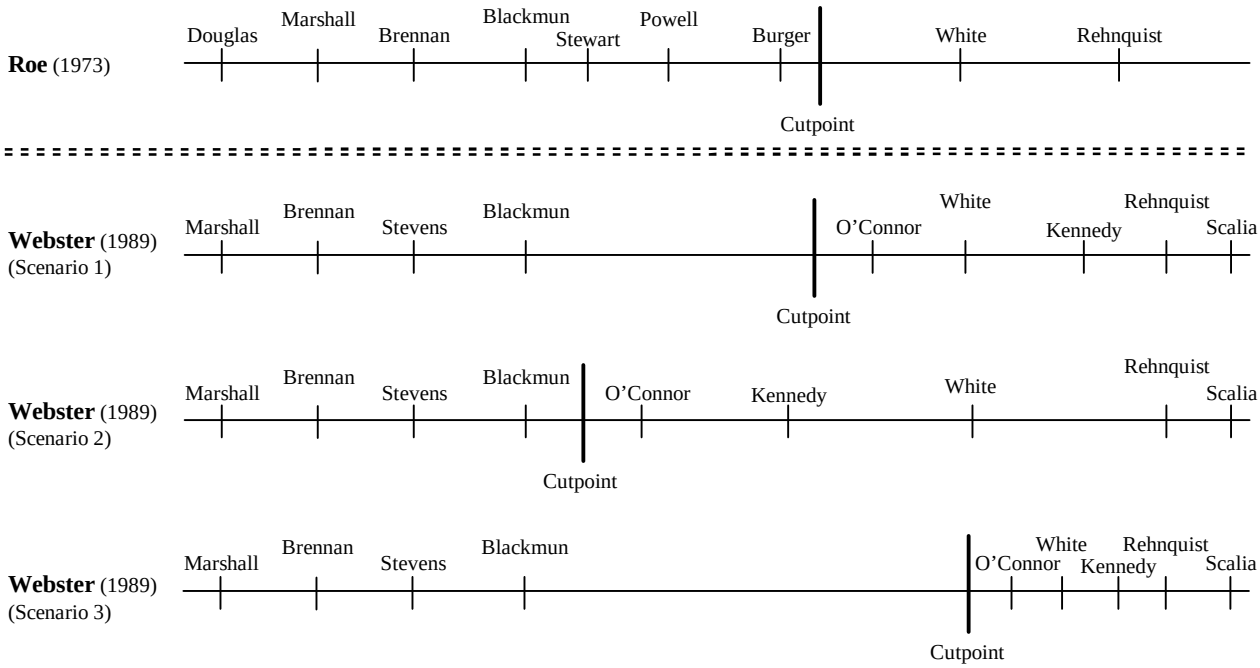

Figure 3: Difficulty in Identifying Preference Change or Cutpoint Change

voted liberally are to the left of the cutpoint and those who voted conservatively are to the right. Here, though, we see the problem: do we believe the second panel in which the vote cutpoint is similar to that of *Roe*? Or do we believe the third panel in which the vote cutpoint has shifted to the left? Or the fourth panel in which the cutpoint has shifted to the right? The stakes are high for estimating preference change over time: if we believe the second panel, for example, there has been a significant rightward shift on the court. In that depiction, O'Connor (the court median for *Webster*) is considerably more conservative than Stewart (the court median for *Roe*). If we believe the third panel, however, these two cases indicate little change in court preferences; in fact, O'Connor in this depiction has similar preferences as Stewart did for *Roe*.

The challenge of identifying preference change is not merely theoretical. Consider, for

example, the following apparent anomaly in Martin and Quinn's preference estimates: according to their model, the median ideology of the Supreme Court nearly reaches its post-war conservative peak in 1973 (see Figure 2). If their estimates are correct, the court that produced *Roe* and struck down the death penalty in *Furman v. Georgia* 408 U.S. 238 (1972) was actually one of the most conservative in the modern era and was more conservative than the court today. While it is possible that the court was indeed extremely conservative in 1973, such a finding conflicts with conventional wisdom enough to merit careful examination.[5]

One possibility is that Martin and Quinn's "sticky preference" model does not capture all preference change over time. In their model, individual judicial preferences vary from term to term. An informative prior based on the estimated preference in the previous term constrains preferences from moving too much from one year to another. Two factors raise questions whether such a model can pin down preference change over time. First, assuming sticky preferences in this manner does not place any restrictions on the directionality of change. Without information on the relative location of vote cutpoints, the estimation has no reason to favor any one of the bottom three panels in the figure, meaning that, on average, the best assumption is that the cutpoint has not changed.[6] If there is a rapid increase in the number of conservative votes (as occurred in the early 1970s) and cutpoint distribution is by default to be assumed to be the same or to the right, this means the estimation procedure will lean towards the alignment in scenario (3) even if, in fact, scenario (1) or (2) reflects reality. Second, identifying preference change over time with a sticky preference assumption

[5] Columnist Michael Kinsley refers to *Roe* as "the high-water mark of liberal judicial activism" (Kinsley 2005).

[6] With a prior on the cutpoint distribution, there will be more weight placed on cutpoints closer to the prior mean. If the cutpoints tended to be to the left for the very liberal Warren Court, this would mean that the cutpoints in the succeeding courts would be assumed to move toward the right.

is very hard given the limited number of individuals voting on each case. As highlighted by Londregan (2000), we need many observations or case-specific data to pin down case cutpoints with any precision. In the figure, we see, for example, that even an assumption of absolutely fixed preferences would not be very helpful in pinning down preference change over time as the crucial justices in the middle did not serve on both the *Roe* and *Webster* courts.

The challenge of inter-temporal preference estimation extends even to Congress where much more data is available. Poole and co-authors create several sets of preference estimates, including "common space scores" that "place the members of the House and Senate in the same space... [allowing] members to be compared across Chambers and across Congresses" (Poole 2005; Poole and Rosenthal 1998; Poole 2003 and McCarty and Poole 1995). For these estimates, they assume members of Congress have fixed preferences over time (even if they moved between the House and Senate).

Poole (2005) encourages caution in the use of these estimates. An examination of selected members of Congress makes it clear why. Figure 4 depicts the Common Space scores of selected members. The first dimension captures "party loyalty" and explains most votes in Congress; the second dimension generally is associated with "race-related matters" (Poole and Rosenthal 1997, 46-48). It was important in the fifties and sixties, but faded considerably with the Republican realignment in the South in the eighties. Average preferences are roughly (0, 0). Senators Helms and Kennedy are included as reference points, as these two typically are treated as anchors on the conservative and liberal extremes on both dimensions. So far, so good. However, when the southern segregationist senators are compared to modern moderate

Democrats we see a striking anomaly: the two groups are virtually indistinguishable on both dimensions. If we were to accept these characterizations, we would be saying, for example, that Sen. Hollings (D-SC) was more conservative on race in 2004 than Sen. Harry Byrd (D-VA), who advocated "massive resistance" to civil rights rulings by the court at a time when African-Americans were routinely denied voting rights and segregated in public as a matter of course. Byrd (not to be confused with Sen. Robert Byrd (D-WV)) stated that that "we should exclude the Negro population" from voting and that "non-segregation would be most unwise and I will certainly oppose it in every way I can" (Wilkinson 1968, 63; Heineman 1996, 318).[7]

It seems inconceivable that Hollings, a Senator who, among other things, voted to override President Bush's 1988 veto of the Civil Rights Restoration Act and voted for the 1991 Civil Rights Act, has the similar preferences on race as the pro-segregationist members of Congress who preceded him.[8]

The explanations for this result all point to a need for continued research in this area. First, there has been a change in the underlying dimensionality congressional voting over time. For other estimates, Poole (2005) cautions users that "cross-Congress comparisons should be conducted only between Congresses occurring during one of the stable 2-party

[7]On this measure, Hollings is also similar as Sen. Eastland who had stated "I assert that the Negro race is an inferior race. I say quite frankly that I am proud of the white race. I am proud that the purest of white blood flows through my veins. I know that the white race is a superior race. It has ruled the world It is responsible for all the progress on earth." (DeParle 2004, 32, citing *Congressional Record* 79th Congress 1st Session vol 91 (June 29, 1945) p. s7000). Note that the segregation position was consistent with public sentiment in the south: in 1964 51 percent of Southern whites supported segregation and another 39 percent supported "something in-between" desegregation and strict segregation (see the National Election Study 1964 variable 0815).

[8]In fact, Hollings makes it clear that his views have evolved considerably from the time of *Brown v. Board of Education*: "It certainly is the most significant judicial decision of that century ...There is no question in my mind that was for the good. I had my doubts at that particular time... [The plaintiffs] really understood the Constitution in America better than this particular Senator" (*Congressional Record* 5/13/2004, p. S5457). Contrast Hollings views in 2004 on the case to the pro-segregation Senators in Figure 4 who signed the Southern Manifesto which described *Brown* as "an abuse of judicial power", "contrary to the Constitution" and "unwarranted."

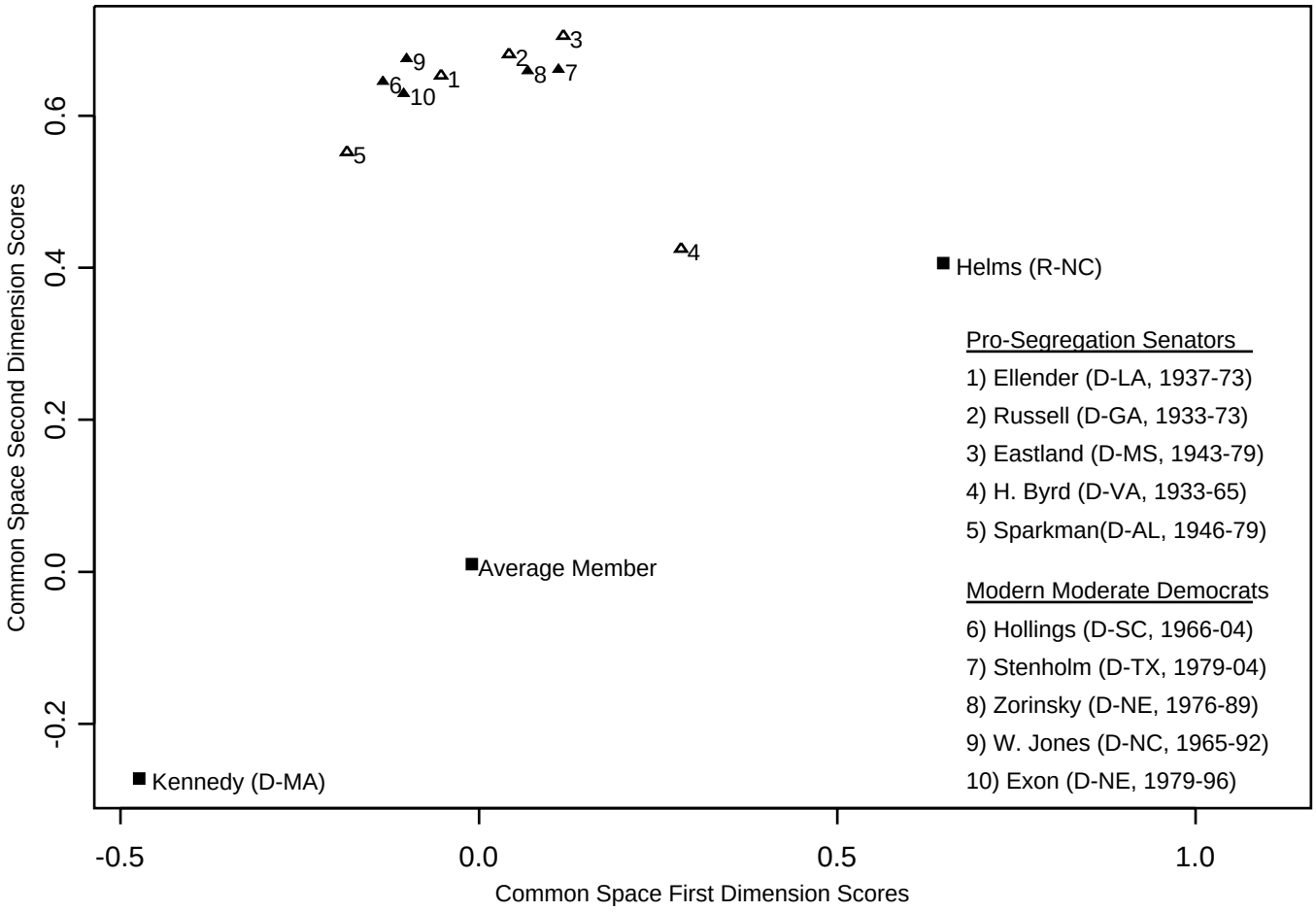

Figure 4: Common Space Scores of Selected Members of Congress

periods of American History." For the Common Space scores which are produced for 1937-2000, it is not clear if that period is to be considered a stable period or, if not, what the dividing lines should be. Of course, many scholars are interested in time periods that span this time period (or longer), meaning that this advice is at best cold comfort.[9] Second, preferences may have changed over time; Poole (2005) explicitly identifies the assumption of fixed preferences as a reason for caution in using the Common Space scores. Later, I provide estimates that reflect Hollings' preference evolution over the course of his career. Third, the whole political spectrum may have moved left such that the ordering of preferences has been stable, but the content of votes has changed. That is, the nature of debate on racial matters shifted from explicit de jure discrimination in the fifties to affirmative action in education and the workplace in the nineties. If the voting patterns remain similar on even as the agenda changes, it will be very hard to identify changes in preferences. Such a possibility is plausible if vote cutpoints tend to be toward the middle of the contemporary political spectrum. As the spectrum moves left, so too do the cutpoints making the vote patterns appears similar across time even as vote content differs substantially. These kind of movements cannot be picked up by the various NOMINATE measures because, as Poole and Rosenthal (1997, 6)

[9] One may consider other Poole and Rosenthal estimates such as DW-NOMINATE which are "directly comparable" across Congresses for Congresses within stable two-party periods (Poole 2005). Even with these estimates, inter-temporal comparisons are tricky: the second (race-related) dimension looks very much like the Common Space second dimension, with the segregationist Southern Democrats having scores very similar to modern moderate Democrats. These groups differ on the first dimension, although the DW-NOMINATE first dimension scores will not be useful for studies in which race politics are important as they do not reflect differences in preferences in that area. For example, Sen. Hubert Humphrey (D-MN) (whose rise to national prominence was with his 1948 Democratic National Convention speech in which he urged the Democratic party "to get out of the shadow of states' rights and walk forthrightly into the bright sunshine of human rights!") was more conservative on the DW-NOMINATE dimension one than Sen. William Fulbright (D-AR) who signed the Southern Manifesto and opposed major civil rights legislation. (Their DW-NOMINATE second dimension preferences differ substantially.) Other examples why it is a bad idea to use DW-NOMINATE first dimension scores for anything to do with civil rights is that estimated preferences of the southern segregationist Senators (such as Smathers (D-FL), Hill (D-AL), Sparkman (D-AL), Ellender (D-LA), Talmadge (D-GA) and Stennis (D-MS)) were more liberal on the first dimension than northern Republicans who were, in fact, very liberal on race (such as Hatfield (R-OR), Percy (R-IL) Weicker (R-CT) and Packwood (R-OR)). (Again, their second dimension scores differ substantially in the direction one would expect.) Or, one may consider the "turbo-ADA" scores from Groseclose, Levitt and Snyder 1999). These scores, however, were not designed to fix the problems inherent in interest group ratings, but were instead designed to stretch and shift the scores so that an individual's scores in two Congresses would, on average, map to the same score. These scores assume no preference change for individuals.

are careful to note, the NOMINATE procedures are "blind to the substance of the roll call vote(s)."

This discussion highlights two important points about preference estimation across time. First, without fixed reference points, preference changes over time are not statistically identified; we simply cannot determine whether the agenda changed or preferences changed. Second, inter-temporal preference estimation may be problematic even with an assumption of fixed or sticky preferences as with widely used preference estimates for the Court and Congress. Failure to address these points may render preference estimates misleading as in the example of the southern Senators. The next section presents an estimation strategy and statistical model which seeks to avoid pitfalls related to inter-termporal and inter-institutional preference comparisons.

2 Generating Comparable Preferences

2.1 Using "bridge" observations and vote information to identify preferences

The central problem here is that identifying preferences across institutions or across time, requires identifying fixed and common reference points.[10] This paper achieves inter-institutional and inter-temporal comparability by making use of two kinds of external information that do not typically enter into voting analyses. The first is the use of "bridge" observations of actors taking positions on issues before another institution (Bailey and Chang 2001). These bridge observations provide fixed references against which the preferences of actors across

[10] King, Murray, Salomon, and Tandon (2004) use answers to highly specific vignettes as fixed reference points that can be used to better understand international responses to more abstract questions about health status; in terms of the model here, the use the vignettes to provide a common cutpoint across all cultural contexts.

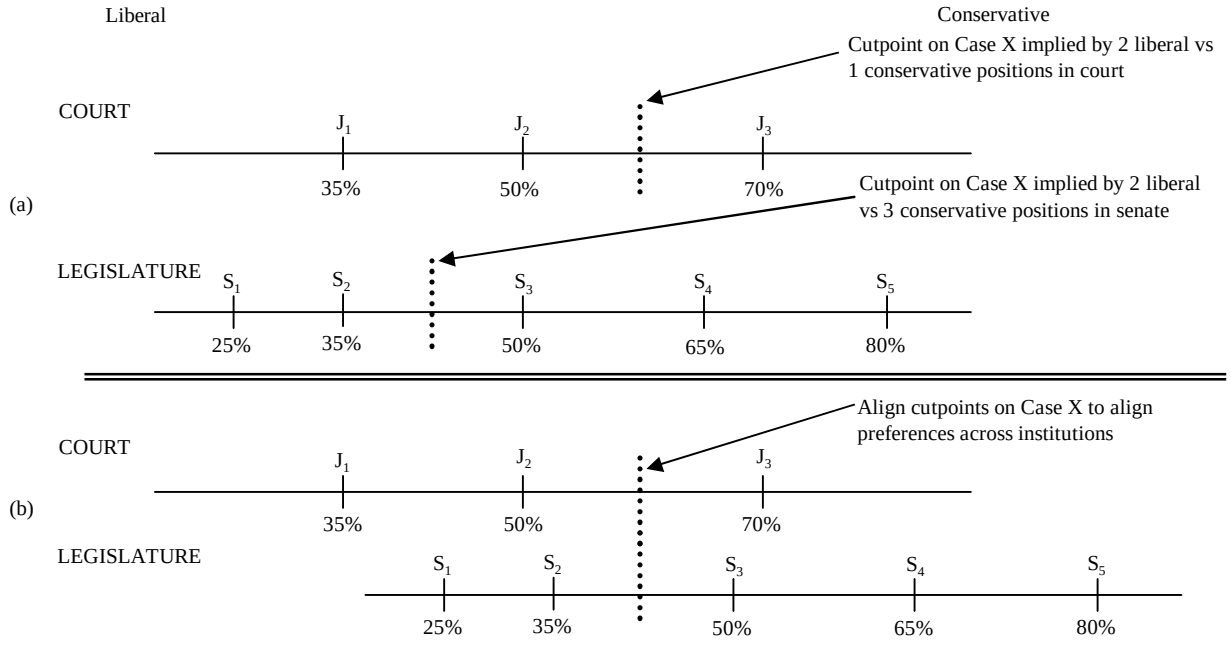

Figure 5: Identifying Preferences in an Inter-Institutional Context

institutional boundaries can be judged. The second is use of inter-temporal bridges including information about the relationship of vote cutpoints across time and position-taking by individuals on earlier votes in Congress and the court.

To see conceptually how inter-institutional preferences can be identified with bridge observations, suppose we are interested in comparing preferences of a three-person court (with justices J_1, J_2 and J_3) to a five-person senate (with senators S_1, S_2, S_3, S_4 and S_5). These individuals are aligned from liberal to conservative in each institutional context in Figure 5(a). I include hypothetical percent conservative scores to illustrate again the weakness of assuming direct comparability of such scores. As discussed earlier, we cannot calibrate preferences across the two contexts based only on votes within the respective contexts. However, if we observe the position of the justices and senators on "Case X" we have information that is very helpful. In the example, two justices were liberal and one was conservative on Case X

while two senators were liberal and three were conservative on the same case. Using the cut-point of Case X as a fixed reference, we can heuristically align the preferences across the two institutions as in Figure 5(b). In essence, the case serves the same function as the anchoring vignette that King, Murray, Salomon and Tandon (2004) use to identify cross-cultural differences in self-perceived health. In this case, the senate median is more conservative than the court median even though both had the same percent conservative scores within their respective institutions. Using this basic insight, the statistical model below incorporates such information in a large-scale fully-specified dichotomous choice statistical model.

There are two approaches to dealing with inter-temporal preference comparability. One is simply to focus on a short enough time period that it is safe to assume preferences are fixed. This solves the problem of identifying cutpoints relative to one and other and is what is typically done in within year preference estimation that is common in the literature (see, e.g., Poole and Rosenthal's W-NOMINATE scores, Clinton, Jackman and Rivers).

The other approach is to find fixed reference points that allow preferences to be estimated even when preferences change over time. This is the approach pursued here and it follows a similar reasoning as for the inter-institutional bridges. To see the logic, first suppose that instead of having two separate institutions, we have the same institution at two separate points in time. There may be some overlap of membership, but if we allow preferences to change over time, we will not be able to align preferences across institutions without additional information or assumptions. One very useful source of information is the existence of cases (such as "Case X") on which individuals at both points in time took positions. This produces comparability just as in Figure 5. To provide a concrete example, Justice Thomas

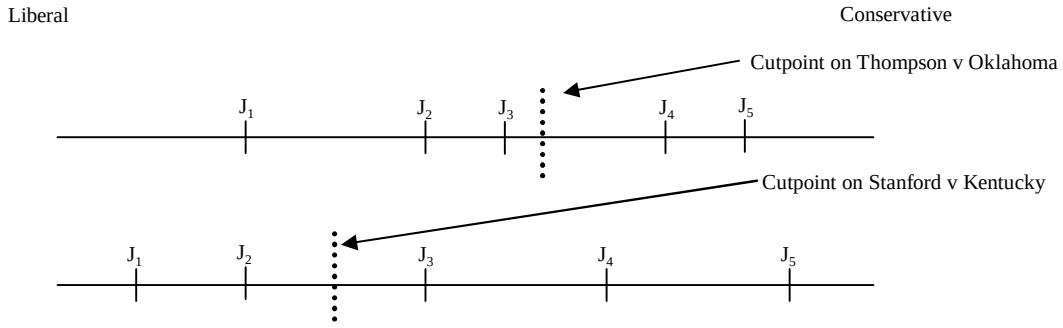

Figure 6: Identifying Relative Cutpoint Locations

provided such a data point when he wrote in *Planned Parenthood v. Casey* 505 U.S. 833 (1992) that *Roe* was "wrongly decided." This provided an indication of Thomas's preferences in 1992 relative to a case decided well before he came to the court by justices with whom he did not generally overlap.[11] Later I describe the collection and coding of such data in greater detail.

Second, we can also use information about the relative position of vote cutpoints. Figure 6 illustrates the relative position of vote cutpoints for *Thompson v. Oklahoma* 487 U.S. 815 (1988) and *Stanford v. Kentucky* 492 U.S. 361 (1989). In *Stanford*, the court assessed whether execution of people between 16 and 18 years old was permissible; in *Thompson*, the court assessed whether execution of people under 16 was permissible. Allowing execution of minors under 16 logically implies execution of individuals over 16 is acceptable. This means that we can infer from the substance of the cases that a justice who was conservative on the Thompson case would be conservative on the Stanford case which in turn implies that the cutpoint on *Thompson* is the right of the cutpoint for *Stanford*.

Incorporating information about vote characteristics has the additional salutary effect of

[11] Note that I am measuring revealed preferences, preferences that may be affected by more than the justices true ideological preferences. For example, the justice may be affected by *stare decisis*, by public opinion, by strategic considerations with regard to Congress and so forth. The estimation infrastructure presented here provides a framework for creating a full structural decision-making model that tests directly for the effect of such factors.

increasing information about case parameters. As Londregan (2000) emphasized, cutpoint estimates for institutions with a small number of actors will be poorly estimated. The best way to mitigate the problem is to add "votes" whenever possible and to incorporate other sources of information about vote parameters (see also Clinton and Meirowitz 2001).

2.2 Statistical model

The model builds on the canonical formulation of latent preferences in the ideal point estimation literature (see, e.g., Bailey 2001, Bailey and Rivers 1997, Baker 1992; Clinton, Jackman and Rivers 2004; Lewis 2001; Londregan 1999). As detailed in the appendix, a random utility model simplifies to the following formulation of the latent preference of individual i at time t on vote v:

$$y^*_{itv} = \alpha_v(\theta_{it} - \kappa_v) + \epsilon_{iv} \tag{1}$$

where α_v is the vote "discrimination parameter" (described in the appendix), θ_{it} is the ideal point of individual i at time t, κ_v is the vote "cutpoint" and ϵ_{iv} is a random shock. The discrimination parameter increases in the distance between the yea and nay alternatives and decreases in the vote specific variance of error; the higher the value, the more a vote distinguishes well between individuals with ideal points above and below the cutpoint. The cutpoint is the midway point between the spatial location of the yea and nay alternatives. The random shock is a mean zero random variable with variance σ_i^2. The errors are assumed to be normally distributed and uncorrelated with each other and with the vote and preference parameters. If $y^*_{itv} > 0$ we observe $y_{itv} = 1$ and if $y^*_{itv} \leq 0$ we observe $y_{itv} = 0$.

I allow the ideal points of individuals to vary over time in order to account for the

ideological evolution of justices and long-serving senators and representatives. For justices, there is a broad consensus that at least some individuals exhibited substantial preference evolution over the course of their service (see, e.g., Bailey and Chang 1999; Martin and Quinn 2002; Epstein, Hoekstra, Segal and Spaeth 1998); for members of Congress, this is more debatable, as Poole and Rosenthal (1997) find little preference change. For the reasons discussed above with regard to preference change over time, however, it may be that their method does a better job of identifying preferences of members of Congress relative to each other than over time. As we shall see, in fact, the data is consistent with substantial preference evolution of members of Congress, as well (which is consistent with the anecdotal evidence on Sen. Hollings presented above).

For long-serving individuals (individuals who served more than 18 years), I assume that the ideology of individual i at time t is

$$\theta_{it} = \gamma_{0i} + \gamma_{1i}X_{it} + \gamma_{2i}X_{it}^2 + \gamma_{3i}X_{it}^3 + \gamma_{4i}X_{it}^4 \tag{2}$$

where the γ parameters are preference parameters to be estimated, and X_{it} is the years the individual has been in office.[12] For individuals (including Presidents) who serve between seven and 18 years, I estimate their preferences with a quadratic equation (meaning I estimate γ_{0i}, γ_{1i} and γ_{2i}). I assume members of Congress who served less than six years and Presidents Eisenhower, Kennedy and Ford have fixed preferences (meaning I estimate only γ_{0i}).[13] For Presidents Johnson, Nixon, Carter, Bush and Bush and justices with less than six years on the bench, I estimate γ_{0i} and γ_{1i}.

[12] The years of service data are expressed in terms of deviations from mean years of service for computational convenience.

[13] Because I have relatively little data for Eisenhower and the data is concentrated in the later years of his tenure, I do not attempt to model preference change for him. Ford was in office only three years.

This functional form represents a tradeoff between flexibility and computation. The fourth order polynomial can represent highly non-linear patterns of ideal point evolution; it requires five parameters. In contrast, Martin and Quinn estimate preferences for each term, with a Bayesian prior that preferences are similar to those in the previous term. This provides more flexibility, but increases complexity and computational time. In addition, the patterns of preference evolution that they do find with their method seem generally explicable in terms of quadratic preferences, let alone a fourth order polynomial specification.

The model is estimated with a Markov chain simulation methods. While Bailey and Chang (2001) used an EM algorithm, the sampling Bayes approach more readily estimates standard errors (Jackman 2000; although the Lewis and Poole (2004) approach could provide standard errors for EM models). The strategy is to repeatedly sample from the posterior density of the parameter distribution. The mode, mean and standard error of the distribution of the parameters can then easily be derived from the mode, mean and standard error of the sampled observations. The appendix and references provide additional explanation.

3 Data

Identification of preferences across institutions and time relies on "bridge" observations and case-specific information. In Bailey and Chang (2001) amicus filings by the Solicitors General and presidential positions on senate roll calls served as the sole source of bridging institutions. A limitation of that approach is that there are only 298 bridge observations from only nine individuals (presidents from Eisenhower to Clinton).[14] In this paper, I vastly expand

[14] The approach here improves on Bailey and Chang (2001) in four significant respects: it identifies preference change across time, it increases the number of bridge observations by almost thirty times, it includes the House of Representatives and it uses

the number of bridge observations by using the additional data sources described below. This facilitates estimation of inter-institutionally comparable preferences from Eisenhower to George W. Bush. All told, the number of bridge observations ($8,907$) is more than 29 times greater than in previous work; in addition, I include additional information based on 501 case/vote cutpoints and 364 judicial comments on previous cases.[15]

For both the Congress and the Supreme Court, I look only at votes and cases related to civil rights and civil liberties. Because these issues have dominated Supreme Court politics from 1950 to 2002, most of the inter-institutional observations are related to these matters. In recent years, positions on these issues have correlated highly with positions on economic and other matters. For example, Martin and Quinn (2001) report that cases on the Burger Court overwhelmingly divided justices according to preferences on the civil rights dimension. These methods can be extended to federalism or economic issues, although a paucity of bridging data may be a problem.

Presidents Presidents' positions on Supreme Court cases are drawn from two sources. One is a set of all statements by presidents on Supreme Court cases. These data are available from presidential Public Papers, presidential library web sites and other sources. These statements indicate whether the president sided with the plaintiff or defendant in a case. For example, George H.W. Bush on June 24, 1992 stated he was "very disappointed" by the Supreme Court's ruling in *Lee v. Weisman* that religious figures could not deliver invocations at public schools graduation ceremonies. The comments reflect the preference of the actor

flexible and powerful Markov chain simulation methods.

[15] The numbers reported here exclude repeat observations (that occur, for example, when a member of Congress files an amicus brief and takes a position on a court case).

at the time the statement is made.

The second source is a database of Solicitors General amicus filings. Given the influence of the president in the selection of the Solicitor General and the power of the president to overrule or remove him or her, these can be treated as administration positions. Bailey and Chang (2001) elaborate on the data and the congruence between presidential and Solicitor General positions.

Presidential positions on Senate and House votes are based on congressional quarterly data provided by Keith Poole. McCarty and Poole (1995) were the first to estimate presidential preferences simultaneously with members of Congress by including their positions on roll call votes. I limit congressional roll calls to those dealing with civil liberties, broadly construed.

Senators and Representatives Congressional positions on Supreme Court cases are based on four sources. First, I have gathered statements in support of or in opposition to specific civil liberties decisions by the Supreme Court. Most observations are from an extensive search of the *Congressional Record*. For example, Sen. Paul Douglas (D, IL) characterized *Brown* as a "correct and noble decision" (*Congressional Record* 110: 20910) while twenty-one Southern Democrats signed the Southern Manifesto stating the decision was "a clear abuse of judicial power" (*Congressional Record* 102: 4459). Second, I have data from amicus filings by senators. Third, I searched for sponsors of legislation that explicitly or implicitly took a position on Supreme Court cases. For example, in 1982, Sen. Nickles (R, OK) introduced S1741, a bill "to provide that human life shall be deemed to exist from conception." Other examples include many proposals to limit the court's ability to restrict prayer in schools. Fourth, I gathered data based on roll call votes that explicitly took a position on specific

Supreme Court cases. For the reasons discussed in the appendix, these are rare. The best example is an October 1999 amendment stating *Roe* was "an appropriate decision and secures an important constitutional right."

In order to have an adequate number of observations for each senator, the collection of roll call votes discussed above has been supplemented with civil liberties votes on which the president did not take a position.

Justices Data on Supreme Court voting from 1950 to 2002 is available from Spaeth (2002) and Spaeth (1999). I include all civil liberties cases on which a president or senator took a position and "important" civil liberties cases.[16] Randomly selected civil liberties cases were included for years with few cases that fit the above requirements in order to ensure an adequate number of observations for every year (something that is important given that I allow preferences to vary across time).

The observations of justices taking positions on cases from previous courts were taken from written opinions. Opinions were identified by (1) searching for the phrases such as "wrongly decided" or "correctly decided" (2) examining every case that overturned precedent and (3) working through issue-specific discussions in legal reference books. An example is Justice Thomas's position on *Roe* discussed above. In cases where a case clearly and directly overturned a precedent, a vote in favor of overturning the precedent was also coded as a vote against the original decision. For example, in *Wolf v. Colorado* 338 U.S. 25 (1949), the Court held that the Fourteenth Amendment does not forbid the admission of evidence

[16] Important cases are those for which at least one of the following is true: discussed directly in the *Harvard Law Review*'s annual court review, included as a landmark case in the Legal Information Institute's database of cases (see supct.law.cornell.edu/supct/cases/name.htm), coded as a salient case in Epstein and Segal (2000), included in the CQ's key cases list, a president or senator or future justice took a position on the case, the case has clear cutpoint relation to another case.

obtained by an unreasonable search and seizure in a prosecution in a State court for a State crime. In *Mapp v. Ohio* 367 U.S. 643 (1961) the Court explicitly overturned this holding by extending Fourteenth Amendment to state courts. A vote in favor of *Mapp* was coded as a vote against *Wolf*. On some occasions a justice changed his or her position on a case; such observations are very useful in gauging the ideological evolution of individual justices.

Vote Parameters Observations relating to the relative position of case cutpoints were drawn from analysis of cases based on issue specific legal reference materials. For example, the abortion issue provides several cases for which cutpoints have clear relations. The Court decided in *Roe* that there is a constitutional right to abortion in the first two trimesters of pregnancy. In many of the abortion cases which followed (including, for example, *Webster*) the court was asked to either rule on legislation that regulated, but did not outlaw abortion. If one thought that states could outlaw abortion, then logically, states must be able to regulate it in a manner that stops short of outlawing it, as was true in all these cases. Hence, someone to the right of *Roe* must be to the right on these cases that do less than outlaw abortion, implying the cutpoints of these cases must be to the left of *Roe*.

Data Validity The use of bridge observations across institutions and time raises issues that do not arise in conventional studies limited to analysis of voting within single institutions. A first question is whether public comments may be less consequential than votes and thus provide less valid measures of preferences. I do not believe this is a fundamental problem. Public position-taking on Supreme Court cases by presidents or senators has clear electoral and political consequences. No contemporary politician would treat his or her position on *Roe*

as a trivial act, nor would politicians in the fifties and sixties treat their public pronouncements on *Brown* or busing cases as inconsequential. Even comments on less prominent cases can be politically relevant, as were Sen. Santorum's (R-PA) remarks on *Texas v. Lawrence* 539 U.S. 558 (2003) (Loughlin 2003). Indeed, it is the importance of such statements that have made the use of non-voting data for the purpose of preference measurement routine. For example, presidential NOMINATE and ADA scores are based primarily on presidential position taking. Likewise, Ansolabehere, Snyder and Stewart (2001) use comment data from candidate surveys to identify legislators' preferences.[17]

A second question is whether comment data may suffer from selection bias because actors have more latitude over whether they will make a comment than they do over whether to vote. This critique is rather subtle: for the results to be contaminated by selection bias, the error in the selection equation must be correlated with the error in the preference equation (Greene 2000, 976). Selection bias is not induced simply if ideology enters in the selection equation. This means that, for example, the tendency of relatively extreme members to take positions does not in and of itself create selection bias; selection bias occurs if individuals are more likely to take positions when they are more liberal (or conservative) than usual. To assess whether selection bias contaminates our results, I used Senate data to calculate two separate ideal point estimates. One was based only on Senate roll call votes. The other was based on court data and voluntary senate data such as public statements and *amicus* filings made by members of the Senate (e.g. no roll call votes were included). Clear differences

[17] In addition, we must not overstate the consequences of most roll call votes. Because most roll call votes are decided by more than one vote, legislators have considerable leeway to vote based on position-taking rather than substance. Mayhew's (1974, 61) famous discussion of position taking begins: "the third activity congressmen engage in may be called position taking, defined here as the public enunciation of a judgmental statement on anything likely to be of interest to political actors. The statement may take the form of a roll call vote." See also Poole and Rosenthal (1997, 69) for evidence that "the bulk of roll call voting is concerned with position-taking rather than policy-making."

in preference ordering across these two estimation procedures would indicate that senatorial behavior on voluntary observations were markedly different from senatorial behavior on non-voluntary roll-call observations. This does not appear to be the case, as the correlation between the two preference estimates is 0.89.

A third question is simply whether the bridge observations provide enough data to pin down the relative preferences across the institutions. To address this question, I simulated ideal points and vote parameters for the actual numbers of individuals and votes in the data set. Based on these parameters, I then simulated votes such that I have a simulated observation for every case in which I have an actual observation in the data set, including bridge observations. This simulated data has the same data structure as the actual data, only here the true parameter values are known. I then estimated the model for this simulated world five times. The results indicate that one can identify inter-institutional preferences quite accurately with this data structure. For example, in the five simulations, the estimated γ_1 parameter correlates with true γ_1 at 0.94 once and 0.95 four times. Averaging the estimated γ_1 parameters across the five simulations yields an estimate that correlates with the true values at 0.97.

4 Preference Estimates

In this section I report the estimates for actors in the three branches of government, focusing on the inter-institutional and cross-temporal aspects of the preference estimates. These results and additional analysis will be made available on my personal web-site.

4.1 Inter-institutional preference comparisons

The real payoff of this exercise is the creation of preference estimates that allow for comparison of preferences across institutions. Figure 7 plots the estimated preferences of the senate median, the president and the court median. In the fifties, the court median moved liberal, in the direction of the president and senate median and then hovered near the senate median. For comparison, consider how different the findings would be if we used the measures from Figure 2: in the top panel, the court moves more liberal even as the senate and president wanted it to be more conservative, while in the bottom panel, the court stays more conservative even as the senate median and president preferred more liberal policy positions.

In the sixties, the court moves firmly in the liberal direction, especially after Goldberg replaced Frankfurter in 1962. In the seventies, the court trends more conservative throughout the decade; this is in contrast to the percent liberal based measure in Figure 2 in which the court bounces around or the PRMQ measure in which the court starts out substantially more conservative than the president (who is, in turn, substantially more conservative than the senate median) and trends liberal. In the eighties, the estimates here indicate that the court median was between the senate and president and trending conservative; again this contrasts with the ad hoc measures in which the court is either quite volatile or in which its median is far more conservative than the president or senate. In the nineties, my estimates indicate that the court median was more conservative than the president and senate median, although with the election of the Republican Congress in 1994, the senate median came quite close to the court median.

One of the interesting patterns we see is that only in brief periods was the court median

Figure 7: Court, Senate and House Medians and Presidents Civil Rights Preferences

outside of the range between the president and the senate, something we would expect given the influence these branches have. In the early period the (more) political branches desired a more liberal court and got one that more or less matched the senate median's preferences. In the sixties, the president was able to move the court strongly to the left, but even then the court remained between the preferences of the president and senate. Nixon's appointments moved the court to the right, but still not beyond Nixon's ideal point. When Carter was president, both he and the senate median preferred a more liberal court, but they were not given the opportunity to re-make the court with appointments. In the Reagan and Bush years, the court continued its rightward drift, but always remained between the presidential and senate median preferences. For a brief period in the early Clinton years, the court was to the right of both the senate median and president, but it stabilized and the senate median eventually moved right.

4.2 Cross-temporal preference comparisons

Earlier, I raised concerns about the ability of existing approaches to capture preference change over time and noted several striking anomalies in the major preference estimation projects. I now revisit these apparent anomalies to see if they persist with the measures developed here. The short answer is that they do not. First, consider the preferences of the Supreme Court median. As discussed earlier, Martin and Quinn's estimates imply that the court median moved decisively more conservative in the early 1970s, reaching almost its conservative peak in 1972 (see Figure 2). According to my estimates, the court moved consistently to the right in the seventies and eighties, but was more liberal in the early

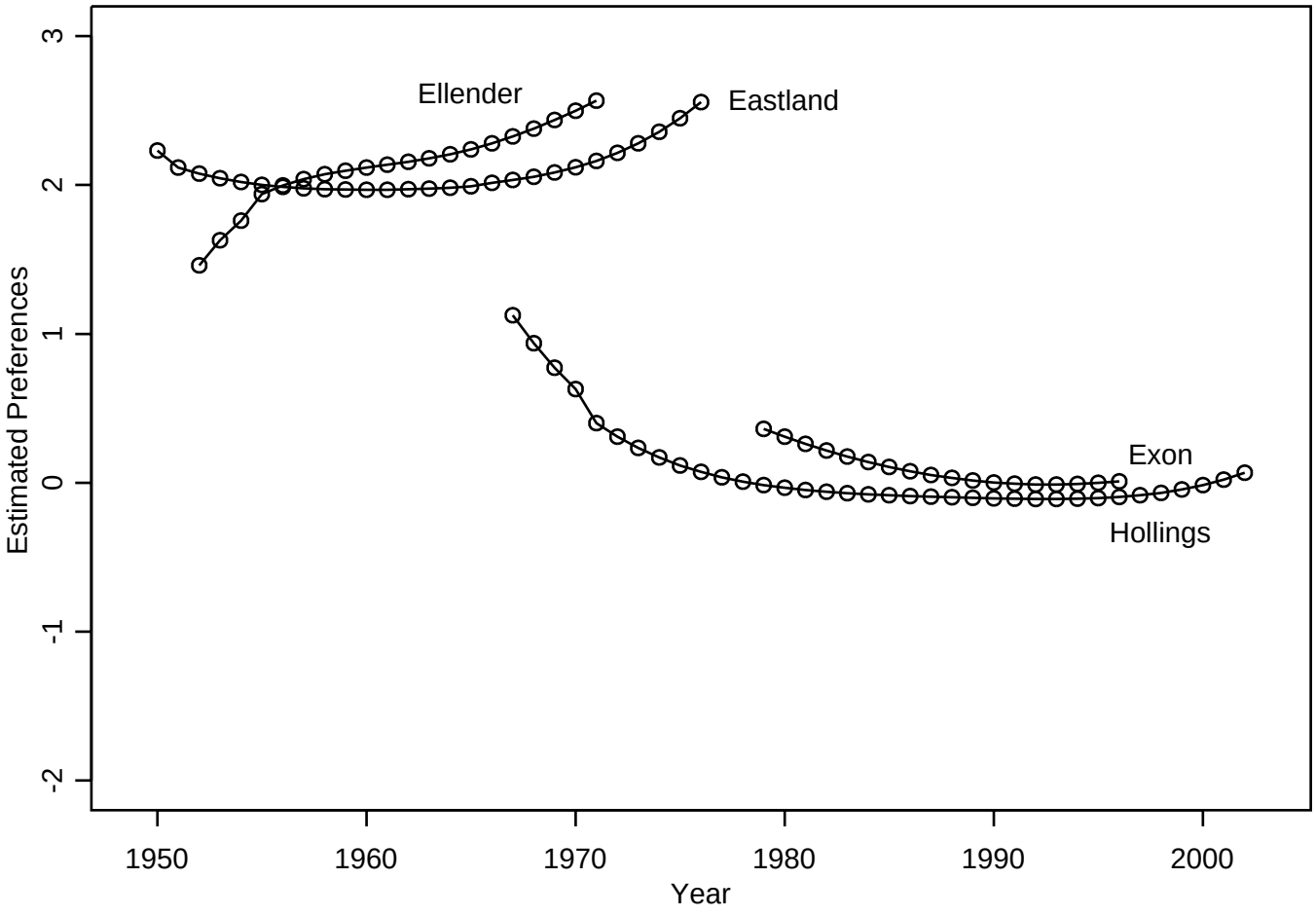

Figure 8: Estimated Civil Rights Preferences over time for selected Senators

seventies than it is today (see Figure 7).[18]

Second, consider the preferences of southern Democrats over time. According to the Common Space scores discussed earlier, modern southern Democrat are hardly distinguishable from their segregationist predecessors. This anomaly disappears with the data and approach used here. Figure 8 depicts the preferences over time of two southern segregationist Senators (Eastland and Ellender) and two moderate modern Democrats (Hollings and Exon). There is a clear difference between the modern folks and the segregationists and indications that for some senators at least, preferences changed over time.

[18] While there are clear differences in the longitudinal preference estimates, the within-term preference estimates correlate highly with Martin and Quinn's results. The overall correlation of my results with their dynamic preference estimates for Supreme Court justices is quite high. On a term-by-term basis, the correlation is above 0.9 since 1975; before 1975, the average within term correlation is lower.

Finally, consider the estimated ideal points of selected Supreme Court justices, as displayed in Figure 9. (Bayesian 90 percent confidence intervals are indicated with solid lines around the estimates.) Liberals are toward the bottom and conservatives are toward the top. These results also accord with intuition, Rehnquist and Thomas at the conservative end of the spectrum, O'Connor toward the middle and Stevens and Souter at the liberal end. Souter began his career on the Supreme Court very close to O'Connor, but moved left over time. Harlan's estimated ideal points were generally between those estimated for O'Connor and Souter, while Brennan's were consistently to the left of Stevens and the rest.

5 Conclusion

Accurate measurement is essential for quantitative theory testing. Testing the most sophisticated theoretical model with the most advanced statistical methods is worth little if the data that goes into the process is not valid. For many questions, this means that many tests of spatial theories must wait on development of valid preference measures that are comparable over time and across institutions.

This paper makes two contributions on this topic. The first is a critique of existing approaches. I highlight theoretical challenges in making preference comparisons across time and institutions and then show how these challenges can matter. In the context of inter-institutional comparisons, I show that ad hoc and equally plausible approaches yield starkly different conclusions about the relative policy preferences of the president and congressional medians. In the context of inter-temporal comparisons, I provide specific examples of preference estimates that strain credulity.

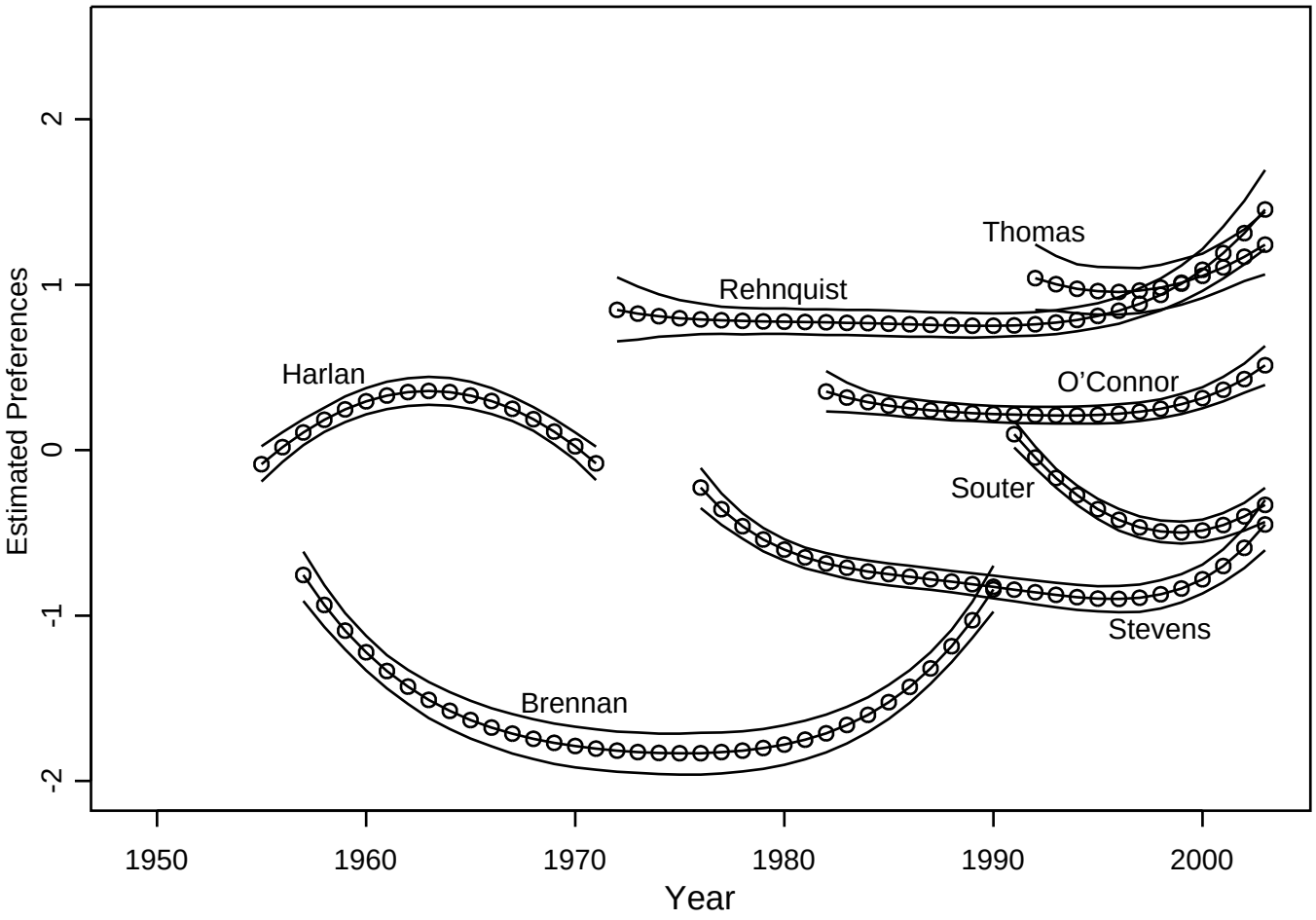

Figure 9: Preference Estimates for Selected Supreme Court Justices

The second contribution of the paper is to provide preference estimates that can be used for preference comparisons across time and institutions. I use two types of data: "bridge" observations of actors taking positions on cases or votes in another institution or in a previous time period and "linkage" information about the relative position of vote cutpoints over time and, sometimes, across institutions. I incorporate this data into a spatial ideal point model estimated via Bayesian Markov chain simulation methods. The immediate payoff is that the method produces preference estimates that make more sense. No longer is the Sen. Hollings of 2002 portrayed as more conservative on race as his southern segregationist forebears and no longer is the Supreme Court of 1973 portrayed as being more conservative than the court of today.

The longer term payoff is that the method provides data and methods that can be applied to a very large number of applications, including studies on judicial appointments, judicial independence and other matters. For example, these preference estimates may be useful in assessing whether the court is potentially constrained by the political branches (Harvey and Friedman 2003; Eskrigde, Ferejohn and Gandhi 2003). These estimates can also be used to assess the Supreme Court appointment process (Bailey and Chang 2003), to track the court's ideology and agenda over time and to analyze judicial responses to Congress and vice versa.

This research – in combination with the broader literature on preference estimation – is also helpful for developing several rules of thumb for scholars who need to make inter-institutional or inter-temporal preference comparisons. First, scholars should be very cautious when making preference comparisons across institutions or time. One needs a fixed

reference point that is both justifiable and the source of sufficient data to identify preference differences. Sometimes, such a fixed reference point simply does not exist: in these situations, we must be willing to admit that we cannot say much about the questions at hand. Second, one should use a preference estimation approach that is as narrowly tailored as possible. If one's model can be tested with data based on one term in one institution based on one issue, then one should be reluctant to drag in other terms, institutions or 2 issues. The risk of slippage is too significant unless one does the (typically) heavy lifting of ensuring that there are enough fixed reference points to make comparisons across issues or terms or institutions. Third, one should use structural models when possible (Herron 2001; Clinton and Meirowitz 2003). That is, instead of using preference estimates from this or another routine as variables, one should write down a statistical model in which the preferences and their determinants or effects are simultaneously determined. This is not always feasible, but the methods developed here and in Bailey (2001) naturally accommodate covariates and, as a general matter, so too do all Bayesian hierarchical models.

Appendix

Data The data on amicus filings come from Gibson (1997) for the period 1953 through 1987 and from Lexis-Nexis Academic Universe and the Solicitor General's website thereafter. Only amicus filings on merit are included. For more details on the selection of amicus filing and Senate roll calls, see Bailey and Chang (2001).

Comments in the Senate and House were taken from the *Congressional Record*. For 1989 to present, I used the Thomas.gov database to search for entries with "Supreme Court." For years before that my research assistants and I investigated every entry under "Supreme Court" in the annual indices. I also investigated other sources, such as Eskridge (1991), *Congressional Almanac* and sources cited therein.

One must be careful when using roll call votes to ascertain senators' positions on Supreme Court cases. First, provisions that address court cases are often embedded in broader legislation. This makes it impossible to know if the vote indicates an opinion on the court case or some other matter. Second, votes are seldom directly on a court ruling. For example, in the 1957 Mallory case the court threw out a conviction of a rapist due to delay before his arraignment. Sen. John Marshall Butler (R, MD) sought to overrule the court with legislation providing that evidence could not be ruled inadmissible due to delay. This legislation was amended to say evidence could not be deemed inadmissible because of "reasonable delay." Butler voted against the amended legislation, deeming it too weak, even as most of the supporters of the legislation viewed it as a vote against the court decision (Murphy 1962, 195, 207).

Statistical Details The latent variable specification in Equation 5 is derived from a random utility framework. Let i = 1,...,N index individuals and v = 1,...,V index votes. The utility of actor i of voting for the conservative alternative is

$$u_i(\lambda_t^C) \;=\; -(\theta_{it} - \lambda_v^C)^2 + \eta_{iv}^C \tag{3}$$

where λ_v^C is the spatial location of the conservative alternative, θ_{it} is the ideal point of the actor at the time of proposal t and η_{iv}^C is a random shock. The utility of voting for the liberal alternative with spatial location of λ_v^L is analogous.

Let $\tilde{y}_{itv}^*$ be the utility difference between the conservative and liberal alternatives. It is

$$\begin{aligned} \tilde{y}_{itv}^* &= -(\theta_{it} - \lambda_v^C)^2 + \eta_{iv}^C + (\theta_{it} - \lambda_v^L)^2 - \eta_{iv}^L \\ &= 2\theta_{it}(\lambda_v^C - \lambda_v^L) + \lambda_v^{L2} - \lambda_v^{C2} + \eta_{iv}^C - \eta_{iv}^L \\ &= (\lambda_v^C - \lambda_v^L)(2\theta_{it} - (\lambda_v^L + \lambda_v^C)) + \eta_{iv}^C - \eta_{iv}^L \end{aligned} \tag{4}$$

Let $\hat{\epsilon}_{iv} = \eta_{iv}^C - \eta_{iv}^L$ be a mean-zero random variable with variance $= \sigma_v^2\sigma_i^2$. The variance has been decomposed into an element associated with the vote (σ_v^2) and an element associated with the individual (σ_i^2). This is useful in order to allow the variance of the shock to vary across votes (and, by extension, across institutions) and also produces substantively interesting estimates of individual level variation that are used, for example, in Bailey and Chang (2003). Dividing Equation 4 through by σ_v yields

$$y_{itv}^* \;=\; \alpha_v(\theta_{it} - \kappa_v) + \epsilon_{iv} \tag{5}$$

where $y^*_{itv} = \frac{\tilde{y}^*_{itv}}{\sigma_v}$, $\kappa_v = \frac{\lambda^L_v + \lambda^C_v}{2}$ is the vote "cutpoint," $\alpha_v = \frac{2(\lambda^C_v - \lambda^L_v)}{\sigma_v}$ is the vote "discrimination parameter" and $\epsilon_{iv} = \frac{\eta^C_{iv} - \eta^L_{iv}}{\sigma_v}$ is a mean-zero random variable with variance σ^2_i.

Observed votes (as opposed to unobserved latent values above) are denoted by y_{itv}. I address rotational identification (e.g. are liberals can have high values or low values) by coding votes in the conservative direction as $y_{itv} = 1$. I identify the location and scale of ideal points by assuming they have mean 0 and variance 1; this is equivalent to fixing two individuals at arbitrary points (see, e.g., Bafumi et al 2004).

The estimation process uses a Gibbs sampler algorithm. This algorithm allows us to draw samples from the posterior distribution of the parameters (Gelman et al 1995, 326; see also Johnson and Albert 1999, 194-197). To start the process, I set provisional ideal points to be the percent of conservative votes by the individual minus 0.5; this anchors conservative ideologies as high values and liberal ideologies as low values. After a "burn in" period, the following iterative procedure will produce random samples from the underlying posterior distribution.

1. Equation 5 implies that y^*_{itv} (where i indicates individual, t indicates term and v indicates vote) will be distributed according to one of the two truncated distributions (see e.g. Jackman 2000, 311)

$$y^*_{itv}|y_{itv} = 1 \quad \sim \quad N(\alpha_v(\theta_{it} - \kappa_v), \sigma^2_i) I(y^*_{itv} > 0) \tag{6}$$
$$y^*_{itv}|y_{itv} = 0 \quad \sim \quad N(\alpha_v(\theta_{it} - \kappa_v), \sigma^2_i) I(y^*_{itv} \leq 0) \tag{7}$$

   where I is an indicator function that serves to truncate distributions above or below zero.

2. Generate γ_i (a 5 x 1 vector) on an individual-by-individual basis. Substituting Equation 2 into Equation 5 and rearranging yields

$$\frac{y^*_{itv}}{\alpha_v} + \kappa_v = X'_{it}\gamma_i + \frac{\epsilon_{itv}}{\alpha_v}. \tag{8}$$

   Because of the heteroscedasticity implied by the above, use GLS results to calculate the distribution of γ to be

$$\gamma_i \quad \sim \quad N((X'_i\Sigma^{-1}_i X_i)^{-1} X'_i \Sigma^{-1}_i \tilde{y}, \sigma^2_i (X'_i \Sigma^{-1}_i X_i)^{-1}) \tag{9}$$

   where X_i is a $V_i \times 5$ matrix of covariates for individual i (and V_i the number of observations for individual i), $\tilde{y} = \frac{y^*_{itv}}{\alpha_v} + \kappa_v$ and Σ_i is a $V_i \times V_i$ covariance matrix with $\frac{1}{\alpha^2_v}$ down the diagonal. I impose a N(0,Ω) prior on γ to identify the preferences of individuals who vote conservatively or liberally all the time. Without this prior, their estimated ideal points would become unbounded. The implementation of the prior follows Gelman et al (1995, 260). As discussed in the paper, I restrict the higher order elements of γ to be 0 for individuals who served relatively short periods of time. For individuals with fixed preferences Equation 9 plus the prior simplifies to $\gamma_{0i} \sim N(\frac{\sum \alpha^2_v \tilde{y}}{1+\sum \alpha^2_v}, \frac{1}{1+\sum \alpha^2_v})$.

3. Simulate σ^2_i for each individual from an inverse χ^2 distribution. A problem with variance parameters is that they can become unbounded (creating a situation in which the

variance is very high and the preference parameters are essentially meaningless). To prevent this, I include an Inv-$\chi^2(n_{0i}, 0.36)$ prior which is equivalent to having observed n_{0i} observations with variance equal 0.36 (Gelman, Carlin, Stern and Rubin (1995, 261). I let n_{0i} equal 10 percent of the total observations for individual i. A variance of 0.36 implies that an individual with an ideal point of 1 (toward the conservative edge of the spectrum) would have about a 5 percent chance of voting liberal on a vote for which an individual with an ideal point of 0 would have a 50 percent chance of voting liberal.

The posterior distribution is σ_i^2 Inv $\chi^2(n_{0i} + n_i, \frac{n_{0i}0.36+n_i s^2}{n_{0i}+n_i}$. To draw from this distribution, draw Z_i from the $\chi^2_{n_{0i}+n_i}$ and let $\sigma_i^2 = (n_{0i}+n_i)s^2/Z$ (where $s^2 = \frac{\sum(y^*_{itv}-\alpha_v\theta_{it}+\alpha_v\kappa_v)^2}{n_i}$) (see Gelman, Carlin, Stern and Rubin 1995, 480). The code for generating the χ^2 random variable is from Marsaglia and Tsang (2000).

4. Generate $\alpha, \alpha\kappa$ on a vote-by-vote basis. If we let $\beta_v = [\alpha_v, \alpha_v\kappa_v]'$ and $\Theta_v = [\theta_{it}, -1]$ (indicating the preference parameter of individual i for vote v which occurred during term t)we can re-write Equation 5 as

$$y^*_{itv} = \Theta_v\beta_v + \epsilon_{iv}. \tag{10}$$

By standard GLS results,

$$\beta_v \sim N((\Theta_v'\Sigma_v^{-1}\Theta_v)^{-1}\Theta_v'\Sigma_v^{-1}y_v^*), (\Theta_v'\Sigma_v^{-1}\Theta_v)^{-1})$$

where Σ_v is a $V_v \times V_v$ covariance matrix with the individual level variance parameters (σ_i^2) of the individuals who voted on vote v on the diagonal and zero elsewhere (with V_v being the number of votes cast on vote v), and y_v^* is a vector of all votes for vote v.

The discrimination parameter is, in part, a measure of vote-specific variance and, as a variance parameter is subject to becoming unbounded as discussed above (see also Baker 1992, 97-98; Mislevy and Bock 1990, 8). I therefore incorporate normal priors on α following Gelman, Carlin, Stern and Rubin (1995, 254, 260).

The knowledge we have about the relations between vote cutpoints is incorporated in the following manner. Knowing that a case had a cutpoint lower than another (in the manner discussed in the body of the paper) implies that the cutpoint parameter distribution is truncated at the cutpoint of that other case. Given that cutpoints are jointly distributed with discrimination parameters, I sample from this truncated distribution by drawing from the truncated joint distribution of both vote parameters via rejection sampling.

5.1 Identification

A model is unidentified "if the same likelihood function is obtained for more than one choice of the model parameters" Gelman, et al, (1995, 422). There are several sources of non-identification that must be addressed with additional assumptions. For fixed-preference one-dimensional models, it is sufficient to fix polarity (meaning, for example, conservative preferences are high values and liberal preferences are low values) and two observations (which I do by setting the mean $\theta = 0$ and variance of $\theta = 1$).

When preferences are allowed to change – as they are here – we need more identifying restrictions. To see this, suppose we make the identifying assumption that Kennedy's ideal point in his first year equals -1 and Helms's ideal point in his first year equals 1 (assume here for simplicity that Kennedy and Helms served identical years). If we observe one vote in their first year and one vote in their second year, we could get the same likelihood by allowing the cutpoint to be 0 in both years of by having the cutpoint be zero in the first year and adding some number to both their ideal points and the cutpoint in the second year. The problem is that the identifying assumption here does not pin down all the cutpoints, something it does do when preferences are assumed to be fixed.

Showing this lack of identification even when ideal points are constrained to be distributed with mean zero and variance one is more involved, but involves around the same issues. Suppose that there is a three person legislature voting on two roll calls at two different points in time. In the first configuration, preferences and the cutpoints are the same for both votes. In the second configuration, preferences and the cutpoint are shifted to the left for the first vote and to the right for the second vote. Because the preferences must (by assumption) satisfy the mean zero, variance one assumption, they cannot simply be shifted, but they must be shrunk; the discrimination parameter can change as well, however, to offset this shrinkage in the likelihood. Hence, if the shifted ideal points and cutpoints are divided by the standard deviation of the shifted ideal points and the discrimination parameter is set to this standard deviation, we have a situation in which two different choices of model parameters yield the same likelihood. (Mathematically, this is $\alpha\sigma(\frac{\theta-\delta}{\sigma}-\frac{\kappa-\delta}{\sigma}) = \alpha(\theta-\kappa)$.) For example, the probability of a conservative vote for individual one on vote 1 in configuration 1 is $\Phi(1(-1.22--0.5)) = \Phi(-0.72)$; the probability of a conservative vote for individual one on vote 1 in configuration 2 is $\Phi(\sigma(\frac{-1.22-0.5}{\sigma}-\frac{-0.5-0.5}{\sigma})) = \Phi(-0.72)$.

	Configuration 1		Configuration 2	
	Vote 1	Vote 2	Vote 1	Vote 2
α	1	1	σ	σ
κ	-0.5	-0.5	(-0.5-0.5)/σ	(-0.5+0.5)/σ
θ_1	-1.22	-1.22	(-1.22 -0.5)/σ	(-1.22 +0.5)/σ
θ_2	0.00	0.00	(0 -0.5)/σ	(0 +0.5)/σ
θ_3	1.22	1.22	(1.22 -0.5)/σ	(1.22 +0.5)/σ

Therefore, an additional identifying assumption is required. I fix two case cutpoints: *Brown* at 1.5 and *Roe* at 0. This makes it impossible to shift and stretch in such a way as to yield an identical likelihood from different parameter values. Compared to the innocuous identifying assumptions about mean and variance of the ideal point distribution, this assumption contains substantive content. This is unavoidable, but another reason one should be cautious when attempting to measure changing preferences.

I deal with unbounded discrimination and preference parameters by using a prior and setting maximum values. Missing values are not imputed as a computational convenience that does not affect estimation. In order to facilitate convergence to the true conditional densities I deleted the first $30,000$ iterations (often referred to as the "burn in" period) and took every 40th sample produced thereafter until I had 500 MCMC samples. The final analysis will be based on longer runs of the simulation process.

References

Albert, James and Siddhartha Chib. 1993. Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association* 88: 669 - 679.

Ansolabehere, Stephen, James Snyder, Jr. and Charles Stewart, III. 2001. Candidate Positioning in U.S. House Elections. *American Journal of Political Science* 45: 136-159.

Bafumi, Joseph, Andrew Gelman, David Park, Noah Kaplan. 2004. Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation. Manuscript, Columbia University.

Bailey, Michael. 2001. Ideal Point Estimation with a Small Number of Votes: A Random Effects Approach. *Political Analysis* 9, 3 (Summer): 192-210.

Bailey, Michael and Doug Rivers. 1997. Ideal Point Estimation: A Survey. Manuscript. Stanford University.

Bailey, Michael and Kelly Chang. 2001. Comparing Presidents, Senators, and Justices: Inter-institutional Preference Estimation. *Journal of Law, Economics and Organization* 17, 2: 477-506.

Bailey, Michael and Kelly Chang. 2003. Extremists on the Court: The Inter-Institutional Politics of Supreme Court Appointments. Paper presented at the American Political Science Association Meeting.

Bailey, Michael, Brian Kamoie and Forrest Maltzman. 2005. Signals from the Tenth Justice: The Political Role of the Solicitor General in Supreme Court Decision-making. *American Journal of Political Science* (January) 49, 1: 72-85.

Baker, Frank. 1992. *Item Response Theory.* New York: Marcel Dekker.

Baum, Lawrence. 1988. Measuring Policy Change in the U.S. Supreme Court. 82 *American Political Science Review* 905-12.

Clinton, Josh, Simon Jackman and Doug Rivers. 2004. The Statistical Analysis of Legislative Roll Call Data. *American Political Science Review* 98 (2): 355-370.

Clinton, Josh and Adam Meirowitz. 2001. Agenda Constrained Legislator Ideal Points and the Spatial Voting Model. *Political Analysis* 9: 242 - 259.

Clinton, Josh and Adam Meirowitz. 2003. Integrating Voting Theory and Roll Call Analysis: A Framework. *Political Analysis* 11: 381 - 396.

Epstein, Lee and Jeffrey A. Segal. 2000. Measuring Issue Salience. *American Journal of Political Science*, 44 (1): 66-83.

Epstein, Lee, Valerie Hoekstra, Jeffrey Segal and Harold Spaeth. 1998. Do Political Preferences Change? A Longitudinal Study of U.S. Supreme Court Justices. 60 *Journal of Politics* 801-18.

Eskridge, William. 1991. Overriding Supreme Court Statutory Interpretation Decisions. *Yale Law School Journal* 101 (November): 331- 365.

Ferejohn, John and Barry Weingast. 1992. A Positive Theory of Statutory Interpretation. *International Review of Law and Economics* 12 (June) 263-279.

Harvey, Anna and Barry Friedman. 2003. The Limits of Judicial Independence: The Supreme Court's Constitutional Rulings. Manuscript, New York University.

Gelman, Andrew, John Carlin, Hal Stern and Donald Rubin. 1995. *Bayesian Data Analysis.* Boca Raton: Chapman Hall.

Gibson, James. 1997. *United States Supreme Court Judicial Database, Phase II 1953-1993*. First ICPSR version. ICPSR Number 6987.

Greene, William. 2000. *Econometric Analysis*. 3rd edition. New York: Prentice Hall.

Groseclose, Timothy, Steven Levitt and James Snyder. 1999. Comparing Interest Group Scores across Time and Chambers: Adjusted ADA Scores for the U.S. Congress. 93 *American Political Science Review* 33-50.

Heckman, James and James Snyder. 1997. Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators. 28 (Special Issue) *RAND Journal of Economics* S142-89.

Heinemann, Ronald. 1996. *Harry Byrd of Virginia*. Charlottesville: University of Virginia Press.

Herron, Michael. 2001. Interest Group Ratings and Regression Inconsistency. *Political Analysis* 9(3): 260-274.

Jackman, Simon. 2000. Estimation and Inference Are Missing Data Problems: Unifying Social Science Statistics via Bayesian Simulation. *Political Analysis* 8, 4: 307-332.

Jackman, Simon. 2000. Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo. *American Journal of Political Science* 44: 375-404.

Johnson, Valen and James Albert. 1999. *Ordinal Data Modeling*. New York: Springer.

King, Gary, Christopher J.L. Murray, Joshua A. Salomon, and Ajay Tandon. 2004. Enhancing the Validity and Cross-cultural Comparability of Survey Research. *American Political Science Review*, forthcoming.

Kinsley, Michael. 2005. Activism, Ripe for The Takings. *Washington Post*. June 26, 2005, page B07.

Lewis, Jeffrey. 2001. Estimating Voter Preference Distributions from Individual Level Data. *Political Analysis* 9, 3: 275-297.

Lewis, Jeff and Keith Poole. 2004. Measuring Bias and Uncertainty in Ideal Point Estimates via the Parametric Bootstrap. *Political Analysis* 12: 105-127.

Londregan, John. 2000. Estimating Legislators' Preferred Points. *Political Analysis* 8: 35-56.

Loughlin, Sean. 2003. Santorum under fire for comments on homosexuality. *CNN.com/Inside Politics* Accessed at cnn.com/2003/ALLPOLITICS/04/22/santorum.gays/.

Marsaglia, G. and W.W. Tsang. 2000. A Simple Method for Generating Gamma Variables. *ACM Transactions on Mathematical Software* 26: 363-372.

Martin, Andrew and Kevin M. Quinn. The Dimensions of Supreme Court Decision Making: Again Revisiting The Judicial Mind. Presented at the 2001 meeting of the Midwest Political Science Association.

Martin, Andrew and Kevin Quinn. 2002. Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999. *Political Analysis* 10, 2: 134-153.

Mayhew, David. 1974. *Congress: The Electoral Connection* New Haven: Yale University Press.

McCarty, Nolan and Keith Poole. 1995. Veto Power and Legislation: An Empirical Analysis of Executive and Legislative Bargaining from 1961 to 1986. 11 *Journal of Law, Economics and Organization* 282-312.

Mishler, William, and Reginald Sheehan. 1992. The Supreme Court as a Countermajoritarian Institution? The Impact of Public Opinion on Supreme Court Decisions. *American Political Science Review.* 87:87-101.

Moraski, Bryon and Charles Shipan. 1999. The Politics of Supreme Court Nominations: A Theory of Institutional Constraints and Choices. 43 *American Journal of Political Science* 1069-1095.

Mouw, Calvin J. and Michael B. Mackuen. 1992. The Strategic Agenda in Legislative Politics. *The American Political Science Review.* 86, 1: 87-105.

Murphy, Walter. 1962. *Congress and the Court.* Chicago: University of Chicago Press.

Poole, Keith 2005. Description of NOMINATE data. http://voteview.com/page2a.htm.

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

Segal, Jeffrey. 1997. Separation of Powers Games in the Positive Theory of Congress and Courts. 91 *American Political Science Review* 28-44.

Snyder, James. 1992. Artificial Extremism in Interest Group Ratings. 17 *Legislative Studies Quarterly* 319-45.

Spaeth, Harold. 2002. *United States Supreme Court Judicial Database, 1953-2000 Terms.* from www.ssc.msu.edu/ pls/pljp/, updated: May 15, 2002.

Spaeth, Harold. 1999. *Expanded United States Supreme Court Judicial Database, 1946-1968 Terms* [Computer file]. 4th ICPSR version. East Lansing, MI: Michigan State University, Dept. of Political Science [producer], 1998. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Stimson, James, Michael Mackuen and Robert Erikson. 1995. Dynamic Representation. *The American Political Science Review* 89, 3: 543-565.

Wilkinson, J. Harvey III. 1968. *Harry Byrd and the Changing Face of Virginia Politics 1945-1966*. Charlottesville: University of Virginia Press.