

Is Today's Court the Most Conservative in Sixty Years? Challenges and Opportunities in Measuring Judicial Preferences

Michael A. Bailey Georgetown University

Court scholars have a voracious appetite for Supreme Court preference measures. Several articles question whether widely used Martin and Quinn (2002, 2011) scores provide valid intertemporal measures, calling into question virtually an entire generation of quantitative research on the Court. This article discusses the challenges of intertemporal preference estimation and revises, updates, and extends Bailey and Maltzman (2011) to present Supreme Court preference estimates that are more defensibly comparable across time and institutions.

Many scholars use Martin and Quinn (2002, 2011)¹ scores to measure the preferences of U.S. Supreme Court justices. Not everyone is convinced these measures capture preference change over time, however. The scores imply the Court was very conservative when it decided *Roe* and *Furman*, two of the most liberal decisions in the history of the Court. These scores also imply the Court today is the most conservative in more than 60 years, a more plausible, yet debatable proposition.

How can we assess such claims? Ho and Quinn (2010, 846) say we cannot. They note the strong assumptions required to use Martin and Quinn scores in a cardinal sense and encourage scholars to use the scores as ordinal measures. Following Ho and Quinn's admonition would undermine virtually a generation of empirical court research and leave us unable to test many important questions in judicial politics.

Are Court preferences so difficult to estimate that we are doomed to using only ordinal, rather than cardinal, measures? To generate intertemporally comparable preference estimates, we must distinguish agenda change from preference change. When we observe an increase in conservative votes on the bench, is it

because the justices have moved to the right? Or does this occur because of changes in characteristics of cases considered by the Court?

This article addresses this challenge by using bridging information that aids intertemporal comparability. The results differ markedly from Martin and Quinn regarding the ideological location of the Court in the 1970s and in the current era. The methods are distinct from Bailey (2007) and Bailey and Maltzman (2008, 2011) in that they use a more flexible approach to modeling preference change (following the logic of Martin and Quinn), cover a longer time period that includes Justices Kagen and Sotomayor and more explicitly address issues that have arisen in the literature such as the coding of ideological polarity (Harvey and Woodruff 2011). The article also presents alternative specifications that control for some (but certainly not all) nonpolicy values that may influence judicial voting, including factors that have not previously been assessed in ideal-point models.

This article proceeds as follows. The first section discusses judicial-preference measurement. The second section presents the methods and data, and the third presents results, including a discussion of the contemporary court as a case that sheds light on

¹An online appendix for this article is available at <http://journals.cambridge.org/jop> containing supplemental details and analyses. Data and supporting materials necessary to reproduce the results in the article will be made available at <http://www9.georgetown.edu/faculty/baileyma/JOPIdealPointsJan2013.htm> upon publication. The Martin and Quinn scores will be referred to hereafter (2002, 2011) as simply "Martin and Quinn scores."

broader measurement challenges and opportunities. The last section concludes.

Challenge of Preference Measurement

Martin and Quinn (2002) generate measures of justices' ideologies via an item-response-theory model of preferences to data on voting by individual justices on all cases from 1937 onward. The method requires no coding of cases as liberal or conservative and estimates a different preference for each justice for each term. The evolution of preferences for individual justices are smoothed over time via a "prior" that shrinks justices' ideal-point estimates in a given term toward their preference estimates in previous terms.

Martin and Quinn scores have been used to explain opinion writing (Bonneau et al. 2007), variation in ideological voting (Bartels 2009), oral arguments (Johnson, Wahlbeck, and Spriggs 2006), appointments (Krehbiel 2007), congressional influence on the Court (Sala and Spriggs 2004), intercircuit conflict (Lindquist and Klein 2006), and much more.

A major concern about Martin and Quinn scores is whether they are comparable across time. Many applications that use Martin and Quinn scores use them because they are "on a comparable scale over time" (Martin and Quinn 2007, 366) and "allow for comparisons of ideological positions for justices who never served with each other" (Bartels 2009, 490).

Such statements are true only if "the distribution of case characteristics is constant over time" (Ho and Quinn 2010, 845; Bailey 2007). Figure 1 illustrates the role this assumption plays for Martin and Quinn scores. The top row shows ideal points of three justices on "Case 1," a case on which two of the justices voted liberally and one voted conservatively. Suppose we assume that justices' ideal points can vary over time (as most work on the Court does) and consider possible ideal points on "Case 2," a case on which one of the justices voted liberally and two voted conservatively. In the first scenario, the case cut point on Case 2 is the same as for Case 1, and this means that justice 2 has moved to the right. However, the second scenario shows a situation in which the case cut point has moved left, and justice 2 has not moved at all. Both scenarios for Case 2 are logically possible and, indeed, highly plausible as we could easily imagine cases that have similar ideological cut points as earlier cases and cases that have different cut points than earlier cases. For measurement, however, deciding

which scenario is most accurate is crucial. Looking at the vote tally on Case 2 of two conservatives versus one liberal provides no guidance, and we must rely on some external assumption or information. Martin and Quinn essentially assume that only scenario 1 is true.²

It is possible that the Martin and Quinn approach could mistake a shift in case cut points for a shift in justice ideal points. Figure 2 shows Martin and Quinn estimates for the Court median over time and, for reference, the percentage of nonunanimous Supreme Court decisions that were conservative. Higher values indicate a more conservative Court median. The Martin and Quinn estimates imply, for example, that the Court median was at one of its historically conservative peaks in 1973. This is hard to square with the fact that during this period the Court was generally considered rather liberal and produced two of its most liberal decisions ever: in *Roe v. Wade* (1973) the Court said that there is a constitutional right to abortion and in *Furman v. Georgia* (1972) the Court imposed a nationwide moratorium on the death penalty. The Martin and Quinn scores also have a dramatic move to the left from 1973 to 1981, something inconsistent with conventional views on the Burger Court and other scaling efforts (Bailey 2007; Grofman and Brazil 2002).

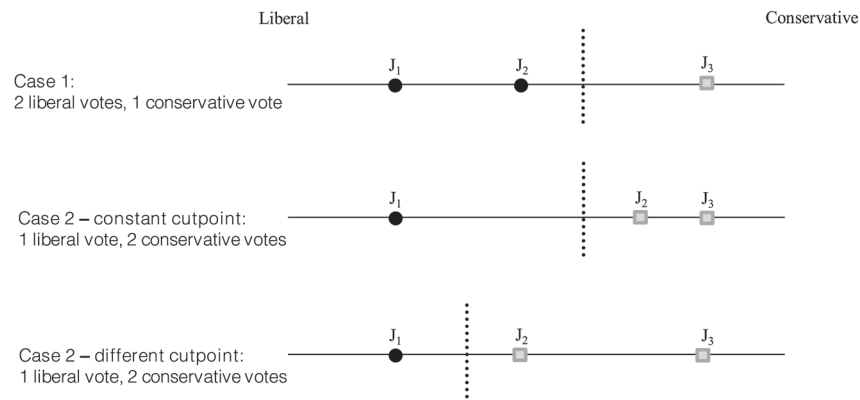
Issues of intertemporal comparability become particularly pressing when Martin and Quinn scores show plausible yet highly debatable shifts such as the sharp turn to the right evident in the figure after 2008, a shift completely absent in the percent conservative data. I discuss this period in detail in the third section.

Concerns about intertemporal comparability lead Ho and Quinn to say in "How Not to Lie with Judicial Votes" that "inferring meaning into cardinal values is misguided" (2010, 846) and to recommend that the Martin and Quinn measures should be used as ordinal measures. If accepted, this claim is highly consequential for the statistical literature on the Supreme Court as the scores (and their "judicial common space" offshoots) have become a mainstay of empirical analysis.³

²More precisely, their approach assumes that the distribution of votes is the same across terms.

³Martin and Quinn (2002) has been cited 454 times according to Google Scholar as of January 2012. Judicial common space scores (Epstein et al. 2007) are based on a mapping of Martin and Quinn scores onto the first-dimension Poole and Rosenthal common-space scores. Although using the first dimension of Poole and Rosenthal scores seems natural, it is actually quite controversial. Segregationist southern senators such as James Eastland (D, MS) are moderates on Poole and Rosenthal's first dimension. It seems more apt to use Poole and Rosenthal's second dimension (on which these members of Congress were arch-conservative) to characterize political preferences relative to the Court (Bailey 2007).

FIGURE 1 Difficulty in Identifying Preference Change or Cut-point Change (color online)



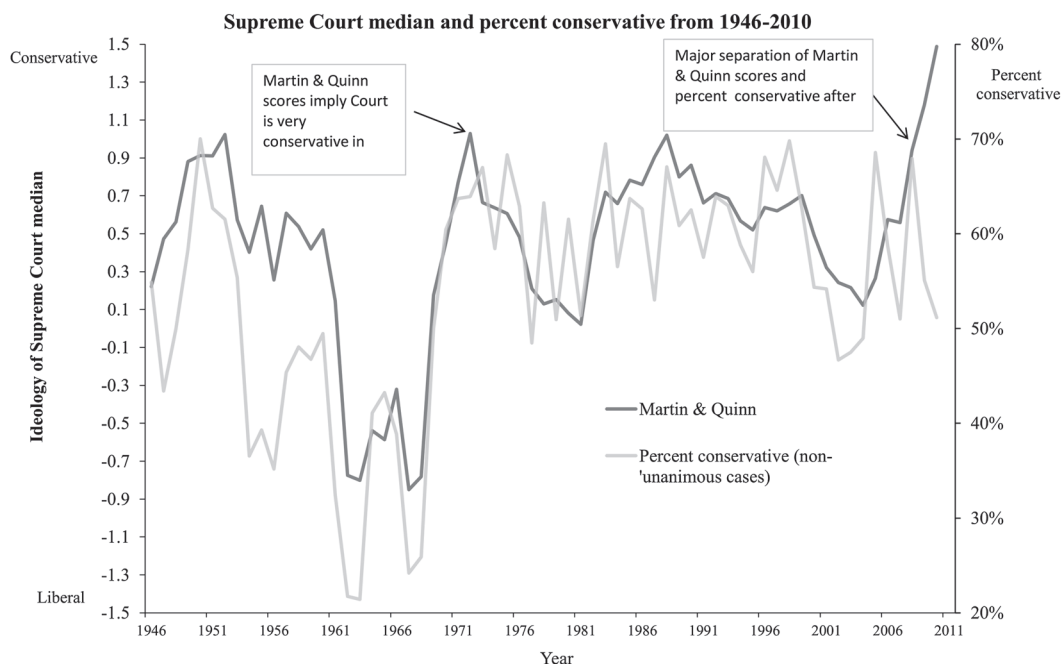
Note: The figure shows hypothetical justices and votes on two cases. The top row shows ideal points of three justices on “Case 1,” a case on which two of the justices voted liberally and one voted conservatively. The bottom two rows show how different cut points and ideal points could generate a shift in voting.

There are two directions one can go in response to Ho and Quinn’s critique. One is to accept it and to use Martin and Quinn scores only as ordinal measures. The question then becomes whether the additional complexity associated with calculating and interpreting the scores is worth using them over simple measures of the percent of time a justice votes conservatively, which generally control for agenda effects and provide

ordinally accurate rankings of justices within terms or natural Courts (Chiou and Imai 2008, 6).

Although ordinal Martin and Quinn scores would be less useful than currently distributed cardinal scores, they do offer four benefits over percent conservative scores. First, as with any item-response-theory (IRT) model, ordinal Martin and Quinn scores do not weight all votes equally. Those votes that do not divide justices

FIGURE 2 Martin and Quinn Estimates of Supreme Court Median over Time (color online)

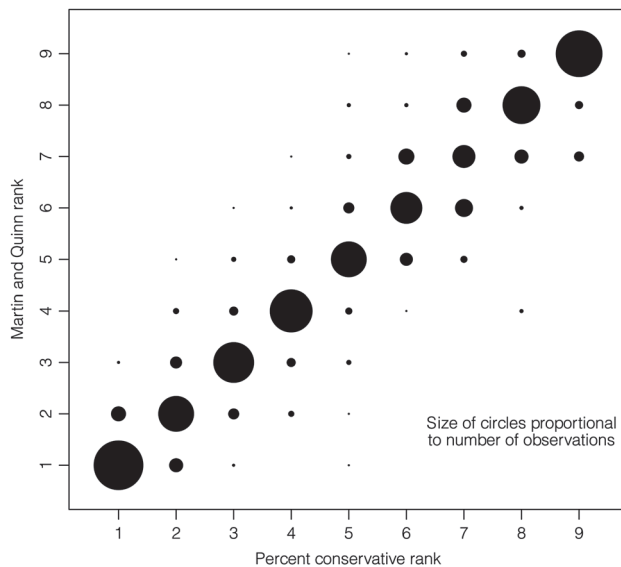


Note: The darker line is the median of the Supreme Court as estimated by Martin and Quinn. The lighter line is percent of times the Supreme Court ruled in a conservative direction based on Spaeth data.

according to “ideology” get a discrimination parameter near zero which decreases their influence in the calculation of preferences. Second, the model incorporates information from previous periods. Practically, this means that if two justices vote conservatively a similar percentage of time in a given term, then the one who was more conservative in the previous term will be estimated as more conservative. Third, by controlling for case characteristics, the Martin and Quinn ordinal scores deal with situations in which justices do not vote on the same cases in a given term. Finally, Martin and Quinn scores come with measures of uncertainty and percent conservative scores typically do not.

Despite these advantages, *ordinal* Martin and Quinn scores do not differ dramatically from ordinal rankings based on percent conservative scores. The correlation of the ordinal ranking from Martin and Quinn and percent conservative for the period 1950 to 2011 is 0.93. Figure 3 shows a scatterplot of ranks produced by the two approaches. The size of the circles reflects how many observations are at a given

FIGURE 3 Ordinal Martin and Quinn Scores Versus Percent Conservative



Note: The figure plots the relative frequency of two ordinal rankings of justice ideal points. The rankings on the horizontal axis are based on percent of times each justice voted conservatively (using Spaeth data). The rankings on the vertical axis are based on Martin and Quinn scores. The size of the circles reflects how many observations are at a given point. For example, of the justices who were ranked as more liberal by conservative voting, most were also ranked most liberal in Martin and Quinn, while a smaller portion were ranked second most liberal in Martin and Quinn and a very small number were ranked third most liberal in Martin and Quinn. None of these justices were ranked any more conservative in Martin and Quinn.

point, and the results illustrate how similar the ordinal rankings are.

This article goes in another direction in response to Ho and Quinn’s critique. The second section focuses on what can be done to generate intertemporally comparable scores and presents a bridging approach to account for agenda change.

Polarity. Another challenge in estimating judicial preferences is determining the ideological polarity of cases (Fischman and Law 2009, 154; Ho and Quinn 2010, 836). That is, how do we know what are liberal and conservative votes? Like pornography, we may know a liberal vote when we see it, but like pornography it is difficult to devise clear criteria. Just as *Lady Chatterley’s Lover* wove a risqué story into literature, *Gonzales v. Raich* (2005) wove profederal government principles into a decision to overturn a California law allowing medical marijuana. What are we to do? Do we deem *Raich* liberal because it favored a federal law over a state law, something liberals typically favor? Or is it a conservative decision because it ruled against medical marijuana? There may be no cut-and-dry answer.

There are three approaches to coding the ideological polarity of cases. Spaeth (2011) uses rule-based coding, deeming a decision liberal if the decision favors an accused criminal, a civil rights claimant (typically minorities, homosexuals, poor people, or some other “underdog”), the government in most-takings cases, a free speech or privacy claim, and so on. Usually this is unproblematic, but there are cases where the process strays from common political understanding. On campaign finance cases, for example, Spaeth codes voting to strike the 2004 Bipartisan Campaign Reform Act as a liberal vote because it favors speech rights even though this was the politically conservative position.⁴

The second approach is automated coding used by Martin and Quinn and standard IRT models; Fischman and Law (2009) call this the “agnostic” approach. In these models, the polarity of the case is estimated from the data. If liberals are in the majority on a case, and they are opposed by conservatives, the case polarity parameter will indicate the decision was liberal. If conservatives are in the majority opposed by liberals, the opposite will be true. If a mix of liberals and conservatives vote together, then the polarity parameter will be near zero, indicating that the case was not ideological. Simplifying only a little, automated coding says a decision is liberal if liberals voted for it and conservatives voted against.

⁴It is also possible that subjectivity enters as coders may code according to what they think they directionality “should” be based on how the vote turned out (Harvey and Woodruff 2011).

The agnostic approach risks circularity. In *U.S. v. Comstock* (2009), seven justices voted to uphold a federal law on sex offenders. Justices Scalia and Thomas voted to strike the law on the grounds that it was not effectuating an enumerated power. An agnostic approach would comfortably code this as a liberal decision as the more liberal justices lined up against the most conservative justices. In fact, though, the polarity was the opposite, as Scalia and Thomas were effectively voting to strike a law that increased penalties on sex offenders. Important implications of such behavior, especially with regard to the effect of policy attitudes versus law on voting, can get lost once the majority's vote is classified as liberal.⁵

A third approach is to use non-Court actors to identify the ideological polarity of Court cases. Harvey and Woodruff (2011) code polarity based on the ideological polarity implied by congressional votes on statutes that made their way to the Court. Hansford (2012) uses amicus filings to identify ideological polarity. Even if the ideological polarity of a case can be argued either way philosophically, the politics of the underlying statute sheds light on the practical ideological implications of a case. If liberal amici or liberals in Congress like something and conservatives don't, it's a liberal decision. If the liberals and conservatives don't disagree, the case is not ideological.

Functional form of preference dynamics. Most Court scholars believe at least some judicial preferences evolve over time. There are several different ways to model the preference change. One is to estimate preference change in terms of a polynomial function (Bailey 2007; Poole and Rosenthal 1997). This approach fits a broad range of nonlinear and nonmonotonic preference change and enables preferences from the previous period to provide information about preferences in a subsequent period. The approach does not, however, handle discrete shocks to preferences very well. Suppose, for example, that justices shift left or right in response to an election outcome. The polynomial approach will smooth out this shift, something that could be detrimental to many research questions.⁶

⁵In *Oregon v. Ice* (2009) Justices Ginsburg, Stevens, Kennedy, Breyer, and Alito argued that the Sixth Amendment allows judges rather than juries to do the fact finding necessary for imposing consecutive sentences on criminal defendants. Justices Scalia, Souter, Thomas, and Roberts argued for a more lenient outcome.

⁶For this reason Bailey and Maltzman (2011, chap. 6) estimate preferences separately in the period just before and just after an election in order to assess separation-of-powers shifts related shifts in Court behavior.

When Martin and Quinn estimate preferences separately for each term for each justice, they incorporate information about preferences in the previous term. They do this via Bayesian priors that "smooth" preferences over time. In contrast to the polynomial approach, preference estimates can still shift discretely, although not as much as they would without priors. One of the challenges is "tuning" the model by selecting a smoothing parameter. This is more of an art, than a science (Martin and Quinn 2002, 147).

Heterogeneity. In addition, questions about dimensionality of judicial preferences lurk in any measurement discussion. Changing dimensionality could induce appearances of preference change where they may be none. Farnsworth provides an example:

Suppose, to take a simplified example, that Justice Kennedy tends to vote for the government in cases involving criminal procedure, but against the government in cases involving free speech, while Justice Rehnquist—a less libertarian sort of conservative—tends to vote for the government in both situations. (Both assumptions happen to be accurate.) Imagine that in term T , there are many criminal procedure cases (where the two Justices vote the same way) and few speech cases (where they don't). Then in term $T + 1$, there are lots of free speech cases. Kennedy's preferences may appear to drift to the left relative to Rehnquist's when they haven't really changed at all. (2007, 1896)

If justices' preferences differ across substantive issue areas, a one-dimensional model could conflate preference change and agenda change. Martin and Quinn (2002, 146) present evidence inconsistent with such a worry, as the "first" dimension in their models seems to explain reasonably well across multiple issue areas. Lauderdale and Clark (2012), however, find that when each case is defined in relation to a substantive area and the cases it most cites, there are definite signs of multidimensionality as the identity of the Court median varies substantially across specific cases, even within a term.

Another way in which there could be multiple dimensionality is that justices could value legal principles, and these principles could ebb and flow in cases across issues. For example, if it is indeed the case that Kennedy is distinctively protective of free speech claims, this could not only have the measurement implications Farnsworth highlights, but it could also have implications for how ideology and the law matter on the Court. Bailey and Maltzman (2008, 2011) identify a large number of justices who systematically deviate from their one-dimensional preferences when precedent, congressional deference, and free speech are involved.

A Bridging Approach to Preference Measurement

This article addresses each of the major challenges described above. It uses a bridging approach to deal with agenda change. It uses position taking by non-Court actors to identify the ideological polarity of cases. It uses a prior-based flexible approach to model preference dynamics, and it is expandable to include covariates that capture an important subset of factors that may not be accounted for by the standard left-right ideological dimension. I discuss each of these in turn in this section.

Controlling for agenda with bridge observations.

As discussed above, the impediment to using Martin and Quinn scores as cardinal preference measures for the Court is Martin and Quinn's assumption that the Court agenda does not change over time.⁷ If, however, we can control for agenda change, we may be able to create measures that are comparable over time. We do not need the agenda to be identical in each period; we simply need to ensure that the agenda change is identifiable.

Therefore, this article builds on previous bridging estimates (Bailey 2007; Bailey and Maltzman 2008, 2011) to identify agenda change by incorporating bridging information that pins down preference changes over time. The analogy is to test standardization. A test score based on one set of questions is hard to compare to another score based on a different set of questions. If we gave identical tests, we could compare scores, of course, but that is infeasible for standardized tests that cannot simply give identical tests year in and year out. Some questions can overlap, though, and this is enough to identify relative test performance as the overlapping questions provide information about relative ability across tests. For the Court case, we will identify certain "questions" (cases) that have been "asked" (voted on) over time. We will also be able to identify cases that are to the left or right of previous cases, information that also helps pin down agenda and preference change over time.

We use two sources of bridging information. One is the positions taken by justices on cases decided by earlier courts. It is relatively common for justices to state a clear position about an earlier decision. For example, in *FEC v. Wisconsin Right to Life* (2007), Scalia stated that *Austin v. Michigan Chamber of Commerce* was "wrongly decided"; in *Allegheny v.*

ACLU (1989), Justice Kennedy wrote "I accept and indeed approve both the holding and the reasoning of Chief Justice Burger's opinion in *Lynch* [*Lynch v. Donnelly*] (1984)." We have collected 1,611 observations of justices taking positions on previous cases by reading all opinions from 1950 onward and looking for instances in which a justice took a clear position on a previous case.⁸

The second source of bridging information is about the relative position of cases over time. It is not unusual for case law to evolve in an ideologically understandable way. For example, the Court held in *Miller v. Alabama* (2011) that juveniles convicted of murder cannot be subject to mandatory life sentences. The cut point of this case was clearly to the left of *Graham v. Florida* (2010) in which the Court ruled against mandatory life sentences for juveniles for crimes other than murder. To see this, consider the vote of a liberal on *Graham*. Does that liberal vote imply support for the liberal position on *Miller*? It does, as voting against mandatory life sentences for murder implies opposition to mandatory life sentences for lesser crimes.

We have identified 327 instances in which case cut points can be linked in this manner. Abortion-related cases provide many examples as the Court has often considered restrictions short of banning it. If a justice were liberal on such a case (whether or not the Court opinion was), this implies that he or she believed *Roe* extends not only to the right to abortion, but also, for example, to the right to abortion in a nonhospital setting. One cannot get to such a holding without also upholding *Roe*, which implies the cut point is to the left of *Roe*. A similar logic applied for many civil rights cases. For example, if a justice voted liberally on busing, he or she implicitly supported the liberal side on *Brown v. Board of Education* (1954) because opposition to segregation is a necessary (but not sufficient) precursor for support for mandatory busing. In our terms, this means the cut point on busing cases and votes was to the left of the cut point in *Brown*. A list of cut-point constraints is available in the supplemental material.

This information is incorporated in the model by imposing a restriction that the estimated cut points reflect the substantive relations identified in the

⁷This critique also applies to Poole and Rosenthal's NOMINATE measures for Congress (Bailey 2007, 438).

⁸We also include 34 observations of Robert Bork taking positions on cases or roll calls at the time of his nomination or before. We estimate a single ideal point for Bork. No other failed nominee had sufficient number of positions that we could find. Given the small number of observations and the ideological extremity of Bork, it is unlikely that inclusion of these observations materially affects other estimates.

historical record. Recall, as in Figure 1, that vote data alone is not enough to identify preference change, but if we know something about movement of cut points, we could, for example, incorporate that information to distinguish between various scenarios in the figure. The appendix provides details on how this information is incorporated in the statistical model.

Some might object that use of bridging information to identify preference change over time yields unrealistic counterfactuals. Are we trying to say what Justice Holmes would make of GPS surveillance? Or what Justice Taney would think of Twitter? That is not what we are doing. There is no time traveling. Our bridge observation of Justice Thomas writing in 1992 that *Roe* was “wrongly decided” does not tell us what Thomas thought about the case in 1973 when it was originally decided. What it does tell us is what Clarence Thomas in 1992 thought of the case, and this has implications for the ideological location of the Court (or at least this one member of it) in 1992 relative to 1973.

The reliance of this approach on external data is both a strength and a weakness. It is a strength in the sense that if anyone has specific reasons to justify thinking that the Court has moved to the right or left, this can be incorporated into the estimation. That is, if someone believes the Court has become more antiaccused and can identify either a case that is clearly to the right of a previous case (based on the substance of a ruling) or finds instances of justices critiquing earlier liberal opinions, then this information can be incorporated and used to help pin down relative movement over time. The use of this data also creates challenges, including not only the effort of identifying such external information, but also the possibility for subjectivity to enter in the collection or coding of data. That the data is available in the supplemental material is one check on this concern.

Case polarity. This article addresses the polarity questions by using the policy stands taken by elected officials to identify the liberal-conservative valence for Court cases for which the valence is debatable. This follows the insight in Harvey and Woodruff (2011) that external actors provide a useful benchmark for coding Court cases because they allow us to identify ideological polarity even if nine justices do not provide clear evidence of it, either due to nonideological factors or by chance. In practice, codings typically follow Spaeth codings. The major exceptions relate to campaign finance and government-funded marketing. More details are in the supplemental material.

Preference dynamics. For each preference estimate for a justice in a given year, we implement a Bayesian

prior based on the ideal point in the previous period. The variance of this prior determines how much smoothing occurs. If it is set at a very large value, then almost no smoothing occurs; if it is set at a very small value, then preferences change very little from one period to the next. As discussed in Martin and Quinn (2002, 147), specific estimates can be sensitive to the setting of such a parameter, and there is no consensus way to determine its value; as with Martin and Quinn, I set this value at a point at which the estimates do indeed move from period to period but not too dramatically. Users of these scores (and Martin and Quinn scores) should note, though, the role this smoothing parameter plays.

Heterogeneity. This article addresses possible multiple dimensionality in two ways. First, it reports estimates based on models that include covariates that account for at least some important determinants of voting that may not reflect standard left-right ideological conflict. For example, suppose that some justices do indeed respect precedent and vote in favor of precedent against their ideological predisposition. These votes will appear moderate in left-right terms, but they do not necessarily reflect moderation in conventional policy-ideological terms.

I therefore also estimate models that control for five nonpolicy variables. Three were included in Bailey and Maltzman (2008, 2011): precedent, congressional deference, and free speech. Two are new: deference to the executive and to dealing with Sixth Amendment challenges. In all these areas, there are reasons to suspect that at least some justices deviate from conventional left-right ideology. The manner in which these variables are included is described in the statistical model below. These estimates are offered more in the spirit of a robustness check than a final word for two reasons. First, it is impossible to come up with a set of covariates that comprehensively captures nonideological factors. Certainly the set here is not comprehensive. Second, the meaning of the “preference measures” is nuanced, and whether one wants a measure net of legal factors depends on context. For example, suppose again that justices are indeed influenced by precedent, and this produces moderate-looking votes. At one level, we could describe these justices as, say, ideological liberals who respect precedent. We could equally validly say that a liberal who respects precedent is effectively a moderate as he or she does not vote as consistently liberal as a liberal justice who ignores precedent would.

The second measure taken to deal with potential multidimensionality is to limit the data set. I include only Court votes on the social-policy dimension that

has dominated the Court docket in the postwar area. This dimension covers crime, civil rights, free speech, religion, abortion, and privacy. Focusing on these issues allows us to focus on the most relevant areas of political-judicial exchange and to minimize chances that our results are affected by behavior on secondary issues that did not necessarily have the same structure of preferences. I also begin in 1950 as not to conflate the ideological splits of the New Deal era that revolved around economic legislation that was quite different than those of the post-1950 Court.

I include a large number of observations from members of Congress as well, including both bridge observations of members taking positions on Court cases and congressional votes on social and legal issues relevant to the Court such as abortion, crime, and race. These observations perform several useful functions. The bridge observations provide more data on the political implications of Court cases, providing information that makes the estimation of judicial preferences more precise. The congressional votes are generally not linked directly to Court cases, but they provide information about the preferences of members of Congress relative to each other, information that, in turn, helps pin down member preferences in a way that makes their positions taken on Court cases more informative.

Including congressional positions also helps identify the nonideological influences on Court voting (or, more precisely, the influence of factors that cannot be explained with the dominant left-right ideological dimension). For example, if we observe seven justices voting liberally against two voting conservatively, it could simply be that the ideological cut point explains the division. However, if members of Congress were overwhelmingly taking conservative positions on the case, then it is possible that the divergence of Court and congressional voting may be explained by one of the additional variables such as free speech and precedent.

Statistical model. This bridging information is incorporated into a standard item-response-theory model of ideal points. The model builds on the canonical formulation of latent preferences in the ideal-point estimation literature (see, e.g., Bailey 2001). The derivation and further details are in the appendix. The core model is

$$\text{Prob}(y_{itv} = 1) = \Phi(\alpha_v(\theta_{it} - \kappa_v)),$$

where y_{itv} is 1 if justice i votes for the conservative position in term t on case v , α_v is the vote “discrimination” parameter (see appendix for more details), θ_{it} is the ideal point of the justice at the time of proposal (the higher the value, the more conservative the justice), and κ_v is the vote cut point. For cases and

votes for which we have information on the relative locations of the cut points, we constrain the cut points to satisfy the inequality constraint implied by the information.

The model that includes covariates is

$$\begin{aligned} \text{Pr}(y_{itv} = 1) = & \Phi(\alpha_v(\theta_{it} - \kappa_v) + \pi_i \text{Precedent}_v \\ & + \delta_{1i} \text{DefCongress}_v + \delta_{2i} \text{DefExec}_v \\ & + \sigma_{1i} \text{Speech}_v + \sigma_{2i} \text{Sixth}_v). \end{aligned} \quad (1)$$

π_i , δ_{1i} , δ_{2i} , σ_{1i} and σ_{2i} are the weights justice i places on precedent, deference (to Congress and the executive branch, respectively), and certain constitutional claims (speech and Sixth Amendment, respectively). Precedent_v , DefCongress_v , DefExec_v , Speech_v and Sixth_v are the precedent, deference to Congress, deference to the executive, speech, and Sixth Amendment variables, coded as described in the supplemental material.

The model is estimated with Markov Chain Monte Carlo methods using Matlab code available in the supplemental materials. A modified Gibbs Sampler algorithm is used to repeatedly sample from the posterior density of the parameter distribution. The mean and standard error of the distribution of the parameters is estimated by the mean and standard error of the sampled observations.

The court voting data comes from Spaeth (2012). There are 32,318 observations of votes by individual justices on 3,701 cases from 1950 to 2011. In order to ensure that the data corresponds to the assumption of unidimensionality, I limit the sample to cases that are conventionally associated with the standard left-right splits on the Court.⁹ Only cases that are relatively salient are selected as well.¹⁰

⁹I use the Spaeth (2011) database and limit cases to those $\text{issueArea} < 6$ (criminal procedure, civil rights, First Amendment, due process and privacy) and not related to Indian affairs (issue equals 20150 or 20160). I exclude cases on unions, economic activity, judicial power, and federal taxation. Citations are the unit of analysis ($\text{ANALU} = 0$ in Spaeth’s data set), and I add split-vote decisions ($\text{ANALU} = 4$) when there are bridging observations. *Bakke* is a prominent example of a case with a split vote and many members of Congress taking positions on one or the other (or both) of the main holdings. I do not include memorandum cases and decrees ($\text{DEC TYPE} = 3$ or 4).

¹⁰Cases are selected if 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 member of Congress or noncontemporaneous justice took a position on the case, the case has clear cut-point relation to another case, the case implicates precedent, deference, or speech as coded.

The interinstitutional elements of the estimation use votes and bridge observations for members of Congress. There are 529,272 individual-level observations of congressional votes on 1,797 roll-call votes in Congress on social and court issues (covering the dimension of issues in the Supreme Court data used). There are also 24,728 bridge observations of members of Congress taking positions on Supreme Court cases. There are 721 individual-level bridge observations of presidential positions on congressional roll-call votes and 641 bridge observations of presidents taking positions on Supreme Court cases. More details on the votes are in the supplemental material.

Ideal points

Justice ideal points. The ideal points are generally what one would expect: Justice Douglas is on the left, and Justice Thomas is on the right, and so forth. Full results are in the supplemental material, and Figure 4 displays the estimated preference estimates for selected justices; figures and numerical results for all justices are available in the supplemental material. A 90% confidence interval is indicated with gray lines. There are some clear nonlinearities: Justice Black was more or less stable until around 1967 when he shot to the right. Frankfurter was liberal until his decisive move to the right from 1958. Justice White exhibited a secular rightward trend punctuated by a move left in 1978.

Supreme Court median over time. Figure 5 plots the Court median for the core model and the most recent available Martin and Quinn scores. To facilitate comparison, the average median has been subtracted from each set of measures.¹¹

If we accept a strong form of the Ho and Quinn critique, we would only use Martin and Quinn scores ordinally. Others, however, do use the Martin and Quinn scores cardinally (e.g., Epstein et al. 2007; Martin, Quinn, and Epstein 2005; Silver 2012), and it is with regard to these applications that the following discussion pertains.

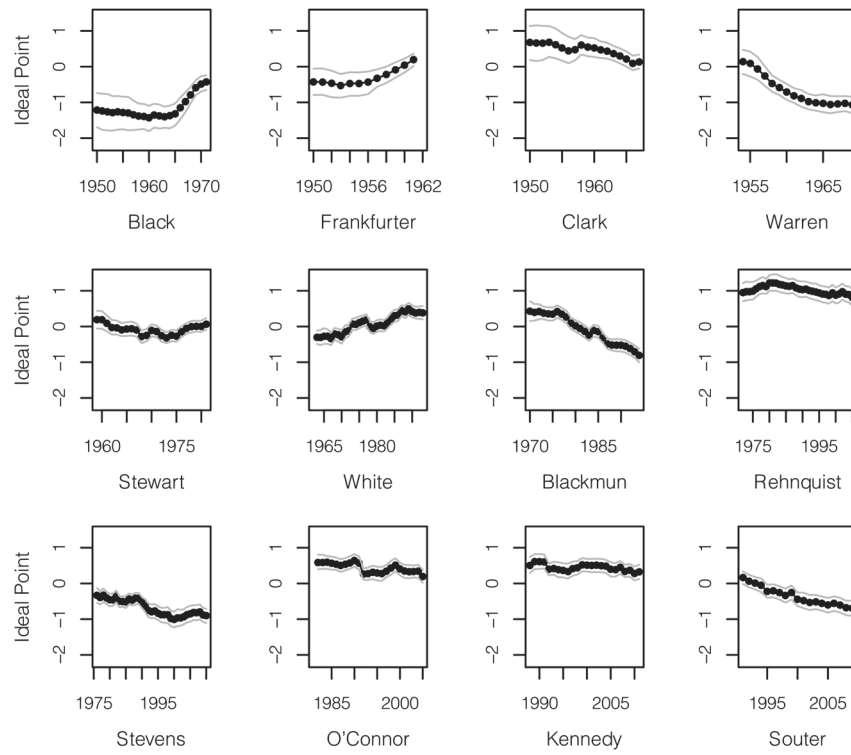
¹¹The variance of medians is similar across measures; the variance of individual ideal points is not the same as Martin and Quinn scores have a broader range. If the medians are normalized by the standard deviation of the individual ideal points, then the Martin and Quinn scores are compressed around their mean without affecting movement relative to itself. Martin and Quinn date cases by the Supreme Court term of the decision. This is not the same as calendar year as Supreme Court terms run from October to September. The 2010 term, for example, covers cases decided from October 2010 to July 2011. The bridging scores use calendar years so as to facilitate bridging between Congress and the Court.

The prime disjunction between the bridging and Martin and Quinn scores is the marked conservatism in the Martin and Quinn scores during the *Roe*-era Court and the substantial move to the left from 1973 to 1981. This pattern is not at all apparent in the bridge estimates which show a strong conservative trend in the early 1970s, but one that leaves the Court liberal compared to later Courts. This seems more consistent with general consensus that the Burger Court, while more conservative than the Warren Court, was not historically conservative (see, e.g., Friedman 2009, chap. 9). A second major disjunction is that the Martin and Quinn scores indicate that the contemporary Court is more conservative than any other time since 1937, the earliest date of the Martin and Quinn scores; we return to this later.

Model with covariates. Figure 5 also includes the medians as estimated from the model that includes the covariates described earlier. These medians are quite similar to the model without the covariates (even as the covariates themselves are statistically significant for many justices). The correlation of medians for the models with and without covariates is 0.97. The ideal points in the model with the covariates are very similar, with a correlation of 0.99. The differences that do exist typically occur when a justice's unexpected votes are explained by a particular covariate. For example, Justices Thomas and Scalia voted liberally on a number of Sixth Amendment cases (even as liberals voted conservatively). In the model with covariates, these votes do not make these justices look as "liberal", and hence their ideal points are shifted to the right for most years. Frankfurter is another interesting case, as his ideal points in the covariate model are more liberal, implying that some portion of his conservative votes are explained by the legal covariates.

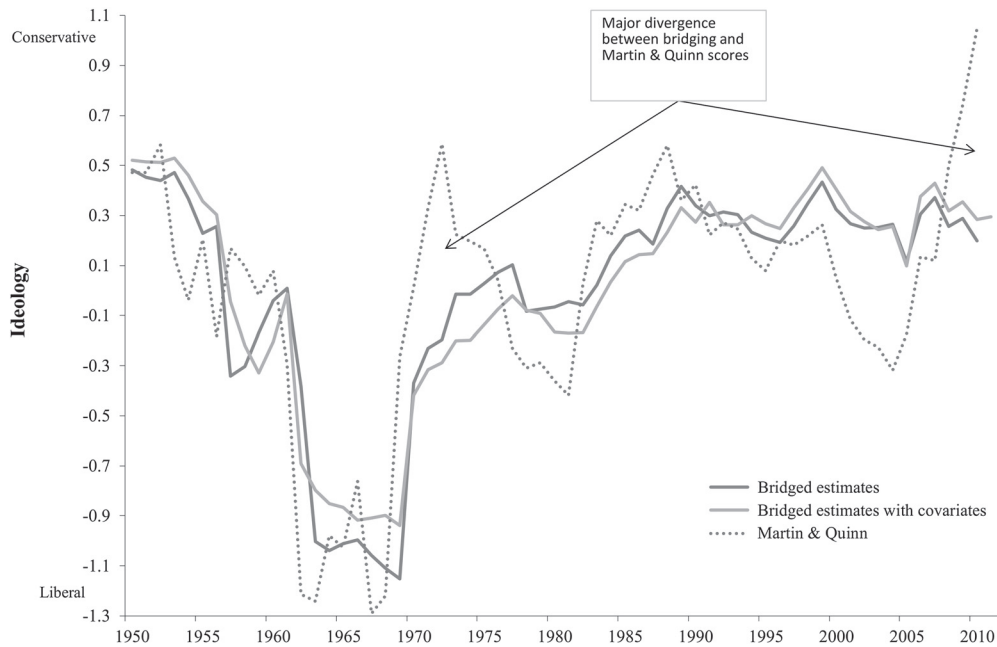
There is no clear way to determine whether the ideal points based on the model with covariates are "better" as it depends on the concept one is trying to assess. If one is interested, for example, in relations between the Court and Congress, one is probably better off with a model without covariates as this measures the differences between Congress and justices, regardless of the source of the voting by the Court. In standard separation-of-powers models, for example, it does not matter why justices vote against the policy preferences of Congress, only that they do. On the other hand, if one is trying to identify the "ideology" of justices in a way that is independent of agenda or case characteristics, one may prefer the estimates based on the model with covariates. The judicial preferences can then be adjusted based on the other estimated parameters

FIGURE 4 Estimated Ideal Points of Selected Justices



Note: The figure plots the ideal points estimated via the bridging methods described in this article. A full set of ideal-point data and plots is available in the online materials.

FIGURE 5 Comparison with Martin and Quinn Estimates of Supreme Court Median over Time (color online)



Note: The darker lines plot the median of the Supreme Court as estimated using the bridging methods described in this article, with and without covariates. For reference, the Martin and Quinn estimated Supreme Court medians are included as a dashed line.

depending on whether overturning precedent or speech and so forth are invoked by a given case or line of cases.

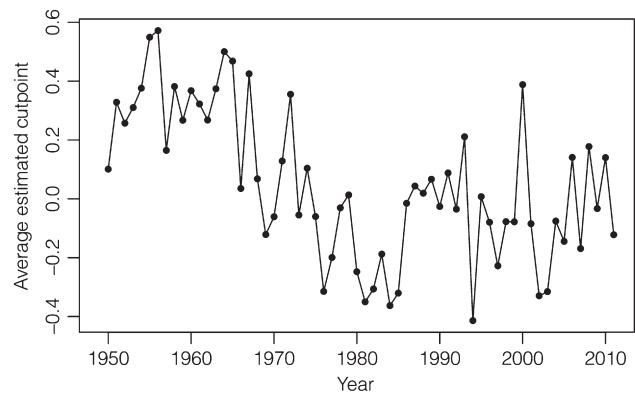
Cut points over time. Figure 6 shows the average case cut points for Supreme Court cases over time. Two points stand out. First, there is substantial variation in case cut-point tendencies. There are several spikes and dips in the data, indicating a somewhat voluble agenda. Second, the agenda moved in important ways in certain time periods. In the late sixties and early seventies, there was a marked (but not necessarily consistent) shift to more liberal case cut points. When such a shift occurs, a moderate justice would vote conservatively more often as the cut points would shift from being to his or her right (and inducing liberal votes) to being on his or her left (and inducing conservative votes). This potentially helps explain the shift to very conservative estimates for the Court median in the Martin and Quinn scores in the early seventies. Their method does not attempt to identify cut-point changes over time and may therefore treat conservative votes in the seventies as evidence of a move to the right when such votes are, it would appear, a function of a leftward shift in the agenda.

How Conservative is Justice Kennedy? Examining instances when the bridging and Martin and Quinn scores diverge in more detail may shed light on potential problems and possible solutions. Bailey (2007) discusses the early 1970s. In this section, we consider the divergence of the measures for the current Court.

The Martin and Quinn median estimate moves strongly left from 1995 to 2005 and then shifts sharply to the right after 2005 (see Figure 5). The initial turn to the right in 2005 occurred when Alito replaced O'Connor, shifting the Court median from O'Connor (who had a Martin and Quinn ideal point of 0.03 in 2005) to Kennedy (who had a Martin and Quinn ideal point of 0.52 in 2005). The Martin and Quinn Court median keeps moving, however, as Kennedy shifts from 0.52 in 2005 to 1.49 in 2010, taking Kennedy from being near the average scores of Justices Stewart and White (around 0.5) to being near the Martin and Quinn score for Justice Rehnquist at the end of Rehnquist's career (around 1.5).¹² The left-side of Figure 7 shows Kennedy's shift, with the Martin and Quinn scores of Justices O'Connor, Scalia, and Rehnquist as reference.

After 2008, the Martin and Quinn measures diverge not only from the bridging estimates but also from the percent conservative data in Figure 2. The

FIGURE 6 Average Case Cut Points over Time



Note: The figure plots the average of the estimated cut points across years.

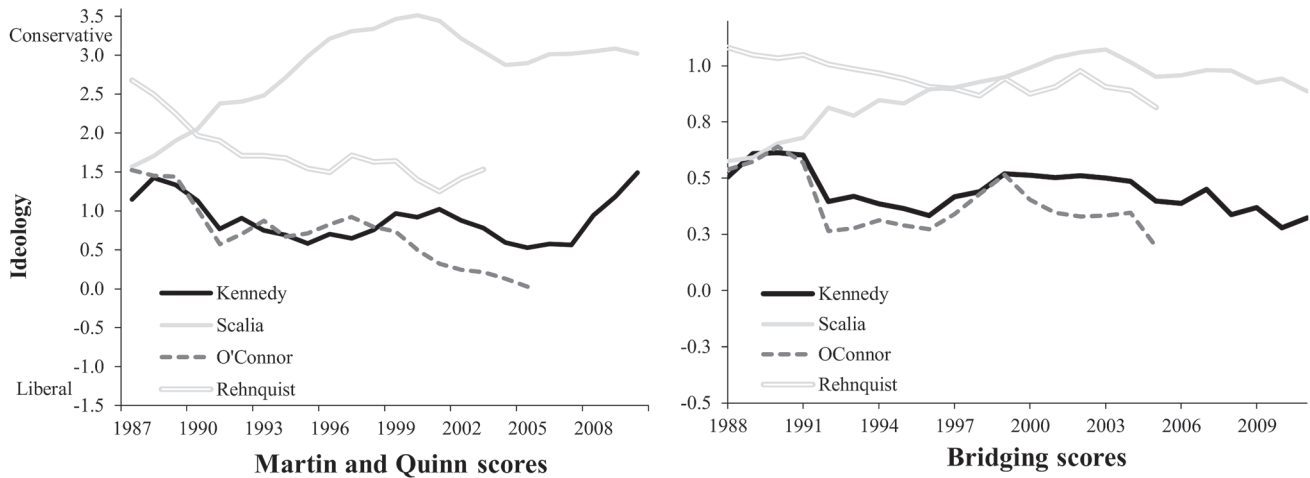
divergence from percent conservative is puzzling given that Martin and Quinn usually tracks the percent conservative fairly closely. The bridging scores, on the other hand, do not diverge from percent conservative in this time period.

The magnitude of the Martin and Quinn shift in median from 2008 to 2010 is historically significant. It is roughly 50% larger in magnitude than the movement to the right from 1980 to 1983 and about two-thirds the size of the largest three-year shift in the postwar which occurred from 1967 to 1970. In contrast, the bridging scores indicate Kennedy has been more or less the same since 1995. These scores put him in 2010 near O'Connor's ideal point at the end of her career and substantially to the left of Rehnquist, as seen in the right-side of Figure 7.

Because Kennedy's shift to the right after 2007 drives much of the recent divergence in preference scores, it is useful to consider Kennedy's votes and ask whether we can qualitatively discern a movement to the left or right. Note first that prominent conservative rulings supported by Kennedy in the 2006 or 2007 terms cannot explain the shift after 2007 (or do so only indirectly via the smoothing parameter). These include *Parents Involved v. Seattle School District* (holding that Seattle schools could not use race as a factor in assigning kids to schools), *Hein v. Freedom From Religion Foundation* (holding that taxpayers had no standing to object to challenge faith-based federal policies), *Morse v. Frederick* (holding that a school could punish a student for holding a sign that said "bong hits for Jesus"), *Ledbetter v. Goodyear* (holding that statute of limitations disallowed challenge to discrimination even if discrimination were discovered later), *Bowles v. Russell* (refusing to hear habeas appeals that were filed late even if a district court granted

¹²Figure 5 subtracts 0.44 from this median score, the average of the Court medians for the Martin and Quinn scores from 1950 onward.

FIGURE 7 Martin and Quinn Scores and Bridging Estimates for Selected Justices, 1987–2011 (color online)



Note: The figures plot the ideal points of selected members of the Supreme Court from 1987 to 2011. Martin and Quinn ideal points are on the left and bridging-based ideal points are on the right.

additional time), *Medellin v. Texas* (holding that International Court of Justice decisions are not binding in The United States) and *Gonzalez v. Carhart* (upholding Partial Birth Abortion Act of 2003).

It is from 2008 to 2010 that Martin and Quinn estimate a decisive shift to the right for Kennedy. Several of Kennedy's conservative votes in that time period do not imply a shift to the right. *McDonald v. Chicago* held that the Second Amendment covered a private right to gun ownership and was supported by more than 300 members of Congress as amici including many moderate members of Congress such as Senators Baucus, Feingold, Snowe, and Webb and Representatives DeFazio, Dingell, Giffords, and Oberstar. Only 56 members of Congress signed an amicus brief for the other side.¹³ The high-profile *Citizens United* case was indeed quite conservative but was a natural continuation of Kennedy's jurisprudence rather than a shift to the right; indeed Kennedy voted against the 1990 *Austin* decision that was overturned by *Citizens United* and joined Scalia's opinion in *FEC v. Wisconsin Right to Life* in 2007 that deemed *Austin* "wrongly decided."

There were also a number of cases on Fourth and Sixth Amendment cases in which Kennedy voted in a conservative direction, often against erstwhile conservatives Scalia and Thomas. They include *Bullcoming v.*

New Mexico and *Melendez-Diaz v. Massachusetts* (in both, Kennedy was in the minority, arguing that confrontation clause did not require testimony of forensic experts), *Arizona v. Gant* (in which Kennedy was in the minority arguing that police could search a vehicle after an arrest even if there was no clear threat to their safety or clear need to preserve evidence) and *Oregon v. Ice* (in which Kennedy joined Ginsburg's opinion stating that a judge, not a jury, could find certain sentencing facts). On every one of these cases, Kennedy voted with Breyer (and sometimes other liberals). These votes seem to reflect less a shift to the right than the "Farnsworthian" emergence of a distinctive cleavage on the Court.

The percent of times Kennedy voted conservatively bounced from 65% to 53% to 64% in the 2008, 2009, and 2010 terms. In the 2006 and 2007 terms (when Martin and Quinn scores are much more liberal for Kennedy), he voted conservatively 62% and 50% of the time. The question is whether these changes reflected changes in Kennedy's ideology or normal variation in the cut points of cases on the Court's docket.

This question is difficult to answer definitively. Of higher-profile cases after 2007, Kennedy supported conservative outcomes on *District Attorney v. Osborne* (finding no postconviction right to potentially exculpatory DNA evidence), *Connick v. Thompson* (ruling against holding prosecutor's office liable for civil rights violations that arose because of poor training), *Herring v. U.S.* (allowing certain good faith exceptions to exclusionary rule), *Salazar v. Buono* (allowing a cross on a public park), *Ricci v. New Haven* (disapproving of

¹³Kennedy's high-profile conservative vote on the Obama health care law occurred outside the range of the data for this article. He joined a very conservative opinion, but note that he was taking a position shared by every Republican in Congress, a majority position in the current House.

a New Haven decision to nullify a firefighter's exam on which whites did better than nonwhites), *14 Penn Plaza v. Pyett* (holding that contractual arbitration precluded judicial resolution of discrimination claims). During this period, Kennedy also wrote or joined liberal opinions on *Graham v. Florida* (holding that juveniles cannot be sentenced to life in jail for nonhomicide offenses), *Cone v. Bell* and *Harbison v. Bell* (providing rights for death-row inmates), *Brown v. Plata* (releasing California prisoners due to overcrowding), *Wyeth v. Levine* and *Altria v. Good* (ruling against claims that federal regulations preempt state laws) and *Caperton v. Massey* (holding that a West Virginia judge had to recuse himself from a case involving a major campaign contributor).

Public opinion is consistent with the idea that the current Court is conservative but not extremely so. Gibson found that Americans who were dissatisfied with the Court in 2011 were equally likely to say the Court was too liberal or too conservative, leading him to conclude "there is no consensus in American politics today about the ideological location of the current Supreme Court" (2012, 9). Jessee and Malhotra (2012) polled Americans on specific Supreme Court cases, presenting them summaries of each position and asking which way the Court should have ruled. Of the nine decisions, they polled with clear ideological valences (*Comstock*, discussed earlier, was the tenth case they polled), eight were conservative decisions (*Citizens United*, *Heller*, *Salazar*, *Ricci*, *Crawford*, *Baze*, *Parents Involved*, *Gonzales v. Carhart*) and respondents agreed with the actual decision 71.6% of the time. It was the one liberal decision in the survey, *Hamdan*, that had the lowest level of popular support, at only 29.9% agreement. Given that the public has not shifted dramatically to the right (Stimson [1999], 2012), these results are in tension with Martin and Quinn's claim that the current Court is the most conservative Court in the postwar era.

Conclusion

Measuring Supreme Court preferences is important for empirical testing and, in turn, for intellectual development in the study of the Supreme Court. Many studies need to make cardinal comparisons, investigating preference change or "spatial distances" between actors or preference differences across institutional boundaries. Scholars frequently use Martin and Quinn scores for these purposes.

The problem is that it may be, as Ho and Quinn (2010, 846) argue, "misguided" to use the Martin

and Quinn measures in this way. Theoretically, cardinal use of the Martin and Quinn scores requires an unrealistic assumption about fixed agenda space. Practically, the Martin and Quinn scores produce highly debatable claims that the *Roe* and *Furman* courts were among the most conservative of the postwar era.

This article presents an alternative approach that directly engages with the issue of changing agendas by using bridging information over time to produce intertemporally comparable preference estimates for Supreme Court justices from 1950 to 2011. The estimates do not suffer from the face validity problems of Martin and Quinn and offer an alternative to going back to the dark ages of ordinal-only preference measurement.

Acknowledgments

I appreciate comments from Albert Yoon. Any errors are mine.

References

- Bailey, Michael A. 2001. "Ideal Point Estimation with a Small Number of Votes: A Random Effects Approach." *Political Analysis* 9 (3): 192–210.
- Bailey, Michael A. 2007. "Comparable Preference Estimates Across Time and Institutions for the Court, Congress and Presidency." *American Journal of Political Science* 51 (3): 433–48.
- Bailey, Michael A., and Forrest Maltzman. 2008. "Does Legal Doctrine Matter Unpacking Law and Policy Preferences on the U.S. Supreme Court." *American Political Science Review* 102 (3): 369–84.
- Bailey, Michael A., and Forrest Maltzman. 2011. *The Constrained Court: Law, Politics and the Decisions Justices Make*. Princeton, NJ: Princeton University Press.
- Bartels, Brandon L. 2009. "The Constraining Capacity of Legal Doctrine on the U.S. Supreme Court." *American Political Science Review* 103 (3): 474–95.
- Baum, Lawrence. 1988. "Measuring Policy Change in the U.S. Supreme Court." *American Political Science Review* 82 (3): 905–12.
- Bonneau, Chris W., Thomas H. Hammond, Forrest Maltzman, and Paul J. Wahlbeck. 2007. "Agenda Control, the Median Justice, and the Majority Opinion on the U.S. Supreme Court." *American Journal of Political Science* 51 (4): 890–905.
- Chiou, Fang-Yi, and Kosuke Imai. 2008. "Nonparametric and Semiparametric Identification Problems in the Statistical Analysis of Roll Call Data." Working Paper, Princeton University.
- Epstein, Lee, and Jeffrey A. Segal. 2006. "Trumping the First Amendment." *Journal of Law & Policy* 21: 81–121.
- Epstein, Lee, Andrew Martin, Jeffrey A. Segal, and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics and Organization* 23 (2): 303–25.

- Farnsworth, Ward. 2007. "The Use and Limits of Martin-Quinn Scores to Assess Supreme Court Justices, with Special Attention to the Problem of Ideological Drift." *Northwestern University Law Review* 101 (4): 1891–1904.
- Fischman, Joshua B., and David S. Law. 2009. "What Is Judicial Ideology, and How Should We Measure It?" *Washington University Journal of Law and Policy* 29: 133–213.
- Friedman, Barry. 2009. *The Will of the People: How Public Opinion Has Influenced the Supreme Court and Shaped the Meaning of the Constitution*. New York: Farrar, Straus and Giroux.
- Gibson, James. 2012. "Public Reverence for the United States Supreme Court: Is the Court Invincible?" Washington University in St. Louis. Unpublished manuscript.
- Grofman, Bernard, and Timothy J. Brazil. 2002. "Identifying the Median Justice on the Supreme Court Through Multidimensional Scaling: Analysis of the Natural Courts 1953–1991." *Public Choice* 112: 55–79.
- Hansford, Thomas G. 2012. "Using the Amici Network to Measure the Ex Ante Ideological Loading of Supreme Court Cases." Presented at the annual meeting of the Midwest Political Science Association.
- Harvey, Anna, and Michael Woodruff. 2011. "Confirmation Bias in the United States Supreme Court Judicial Database." *Journal of Law, Economics and Organization*.
- Ho, Daniel E., and Kevin M. Quinn. 2010. "How Not to Lie with Judicial Votes." *California Law Review* 98: 813–76.
- Jessee, Stephen, and Neil Malhotra. 2012. "Ideological Proximity and Support for the Supreme Court." University of Texas at Austin. Unpublished manuscript.
- Johnson, Timothy R, Paul J Wahlbeck, and James F. Spriggs II. 2006. "The Influence of Oral Arguments on the U.S. Supreme Court." *American Political Science Review* 100 (1): 99–113.
- Krehbiel, Keith. 2007. "Supreme Court Appointments as a Move-the-Median Game." *American Journal of Political Science* 51 (2): 231–40.
- Lauderdale, Benjamin E. 2010. "Unpredictable Voters in Ideal Point Estimation." *Political Analysis* 18 (2): 151–71.
- Lauderdale, Benjamin E., and Tom Clark. 2012. "The Supreme Court's Many Median Justices." London School of Economics. Unpublished manuscript.
- Lindquist, Stefanie A., and David E. Klein. 2006. "The Influence of Jurisprudential Considerations on Supreme Court Decisionmaking: A Study of Conflict Cases." *Law & Society Review* 40 (1): 135–62.
- 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): 134–53.
- Martin, Andrew, and Kevin Quinn. 2011. "Martin-Quinn Scores." Available at <http://mgscores.wostl.edu>.
- Martin, Andrew, Kevin Quinn, and Lee Epstein. 2005. "The Median Justice on the U.S. Supreme Court." *North Carolina Law Review* 83 (5): 1275–1322.
- Martin, Andrew D., and Kevin M. Quinn. 2007. "Assessing Preference Change on the US Supreme Court." *Journal of Law, Economics, & Organization* 23 (2): 365–85.
- Poole, Keith, and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. Oxford: Oxford University Press.
- Sala, Brian R., and James F. Spriggs II. 2004. "Designing Tests of the Supreme Court and the Separation of Powers". *Political Research Quarterly* 57 (2): 197–208.
- Silver, Nate. 2012. "Supreme Court May Be Most Conservative in Modern History." blog post at FiveThirtyEight [Posted March 29 to <http://fivethirtyeight.blogs.nytimes.com/2012/03/29/supreme-court-may-be-most-conservative-in-modern-history/>].
- Spaeth, Harold. 2011. *United States Supreme Court Judicial Database*. from <http://scdb.wustl.edu/index.php>.
- Stimson, James A. ([1999] 2012). *Public Opinion in America: Moods, Cycles, and Swings*, 2nd Edition. Boulder, CO: Westview Press. Updated measures available at http://www.unc.edu/~cogginse/Policy_Mood.html.

Michael A. Bailey is the Colonel William J. Walsh Professor of American Government in the Department of Government and Public Policy Institute at Georgetown University, Washington, DC 20057.