

Did a Switch in Time Save Nine?*

Daniel E. Ho[†]

Kevin M. Quinn[‡]

October 29, 2009

Abstract

Franklin Delano Roosevelt’s court-packing plan of 1937 and the “switch in time that saved nine” animate central questions of law, politics, and history. Did Supreme Court Justice Roberts abruptly switch votes in 1937 to avert a showdown with Roosevelt? Scholars disagree vigorously about whether Roberts’s transformation was gradual and anticipated or abrupt and unexpected. Using newly collected data of votes from 1931-1940 terms, we contribute to the historical understanding of this episode by providing the first quantitative evidence of Roberts’s transformation. Applying modern measurement methods, we show that Roberts shifted sharply to the left in the 1936 term. The shift appears sudden and temporary. The duration of Roberts’s shift, however, is in many ways irrelevant, as the long-term transformation of the Court is overwhelmingly attributable to Roosevelt’s appointees.

*We thank David Abrams, Bruce Ackerman, Ian Ayres, Barry Cushman, John Donohue, Allie Dunworth, Lee Epstein, William Eskridge, David Fontana, Tonja Jacobi, Christine Jolls, Pam Karlan, Mark Kelman, Jim Lindgren, Yair Listokin, Anne Joseph O’Connell, Nick Parrillo, Roberta Romano, Susan Rose-Ackerman, Erica Ross, Jeff Strnad, Emerson Tiller, Elina Treyger, Neal Ubriani, Steve Van Winkle, Stephen F. Williams, and participants at the Searle Law and Political Economy Colloquium at Northwestern Law School, the Law, Economics, and Organization Workshop at Yale Law School, and the annual meeting of the Midwest Political Science Association for helpful comments and conversations, Sam Caporal, Olivia Puerta, and Neal Ubriani for terrific research assistance, Lee Epstein for generously sharing data, and Mark Ramseyer and an anonymous referee for valuable suggestions. The Institute for Quantitative Social Science at Harvard University, Stanford Law School, and the National Science Foundation (grants SES 03-50613, BCS 05-27513, and SES 07-51834) provided research support.

[†]Assistant Professor of Law & Robert E. Paradise Faculty Fellow for Excellence in Teaching and Research, Stanford Law School, 559 Nathan Abbott Way, Stanford, CA 94305; Tel: 650-723-9560; Fax: 650-725-0253; Email: dho@law.stanford.edu, URL: <http://dho.stanford.edu>

[‡]Professor of Law, UC Berkeley School of Law, 490 Simon #7200, Berkeley, CA 94720-7200. Email: kquinn@law.berkeley.edu

1 Background

If the 1936 election was a constitutional moment for law, it was a foundational moment for statistical measurement. Drawing an unprecedented survey sample of 2.4 million respondents from car registration and telephone books, the *Literary Digest* predicted Alf Landon would beat Roosevelt by a whopping 14% margin. Meanwhile, then little-known George Gallup, using a “mere” (quota) sample of 50,000, was ridiculed for predicting that Roosevelt would win. When results flowed in, Roosevelt won by a landslide of 63% of the voteshare, vindicating Gallup.

Until that time, political polls didn’t adhere to certain rules of measurement. There was no random sampling of a target population, little effort at accounting for nonresponse, and scant attention paid to how to account for nonvoters. The lack of principled measurement and the *Literary Digest*’s failure to forecast Roosevelt’s victory aren’t in that sense surprising, as the field of statistics was just maturing into a modern discipline. 1936 was a wake-up call for measurement.

At the same time, the election would lead to a showdown between Roosevelt and the Supreme Court that continues to animate central questions in constitutional history, law, and politics. The prevailing popular, but contested, account of the “switch in time that saved nine” begins with a Court of four stalwart conservatives who battled with three liberal “musketeers” for the survival of the New Deal. Holding the balance were the swing votes of Chief Justice Hughes and Justice Roberts, who in the 1934-35 terms sided with the four conservatives to successively demolish New Deal infrastructure. When Roosevelt unveiled the court-packing plan, the story has it, the Court — or more specifically Justice Roberts — caved. In *West Coast Hotel v. Parrish*,¹ Justice Roberts (arguably) reversed course from an earlier case,² voting with Hughes and the three musketeers to uphold Washington’s minimum wage law for women. Thus, in a somersault of constitutional history, the switch in time resurrected the New Deal and spared the Court packing.³

Despite the story’s stronghold on constitutional history, it remains controversial. While then-

¹300 U.S. 379 (1937).

²*Morehead v. New York ex rel. Tipaldo*, 298 U.S. 587 (1936).

³See Parrish (1982) (attributing “somersault” language to Felix Frankfurter).

Professor Felix Frankfurter charged that Roberts’s votes were irreconcilable, Justice Frankfurter later recanted the view on the Court, defending Roberts’s jurisprudence as consistent (Ariens, 1994; Frankfurter, 1955; Friedman, 1994a; Parrish, 1982). Virtually all historians agree that *Parrish* had been decided *prior* to Roosevelt’s unveiling of the court-packing plan (Cushman, 1998, p. 45; Leuchtenburg, 2005, p. 1081). Further, popular and congressional resistance to the court-packing plan in advance of *Parrish* was fierce, calling into question the seriousness of the packing threat (Cushman, 1998, p. 20).

Scholars have battled for decades — and with increasing intensity in the past years — over the role of Justice Roberts (Chambers, 1969; Farber, 2002; Ross, 2005). “Externalists,” who emphasize the role of external factors such as Roosevelt’s landslide election in 1936 (unanticipated in part due to the *Literary Digest*) and mounting legislative threats to the courts, argue that politics caused a dramatic reversal in Roberts’s jurisprudence, and, to a lesser degree, Chief Justice Hughes (Kalman, 1996, 1999; Leuchtenburg, 1995; Ackerman, 1998). “Internalists,” who emphasize internal doctrinal developments, as well as differences between pre and post-37 legislation, litigating strategies, and cases, argue that the doctrinal evolution was gradual and not marked by dramatic, abrupt change (see the exhaustive work by Cushman, 1994, 1997, 1998; Friedman, 1994b; White, 2000; cf. Ross, 2005; Schwartz, 1995). Just in 2005, the *American Historical Review* convened a symposium about the debate over the “constitutional revolution of 1937,” particularly the question of gradual versus abrupt change in the constitutional jurisprudence (Brinkley, 2005; Kalman, 2005; Leuchtenburg, 2005; White, 2005). Understanding the switch *vel non* “is about the central issue of legal thought during the last sixty years”⁴ and crucial to most accounts of law (Caldeira, 1987; Carson and Kleinerman, 2002; Corley, 2004; Farber, 2002; Friedman, 2000; Gely and Spiller, 1992; Griffin, 1999; Harrison, 1984; Lessig, 1995; McKenna, 2002; Nelson, 1988; Olken, 2002; Pepper, 1998; Purcell, 1994; White, 1996).

In this article, we contribute to existing work by conducting the first modern quantitative

⁴Kalman (1999, p. 2188) (quoting Morton Horwitz).

study of the evolution of the 1937 Court and the realignment of the Roosevelt Court. Using newly collected data on all nonunanimous cases decided in the 1931-1940 terms and modern measurement methods, we study whether the voting patterns can shed any insight into the switch. Doing so allows us to address one of the central questions in this debate: did a switch in the jurisprudence of Roberts occur, and, if so, when?

Our findings inform the debate with two chief findings. First, we demonstrate that Roberts shifted sharply (and statistically significantly) to the left in the 1936 term. Second, we show that the shift appears to have been temporary, but in the long run irrelevant. We statistically document the drastic realignment of the Court that quickly marginalized Roberts, who in three terms would be left as the single most conservative justice on the Court. Our analysis informs and helps to refine existing scholarship, thereby deepening our understanding of this critical episode of constitutional law.

2 Our Approach

Our approach is to study the voting blocs of the justices around the time of the *Parrish* case to determine whether Roberts shifted in a substantively and statistically detectable way. This approach is similar in spirit to methods for detecting credit card fraud, which assess whether specific purchases deviate systematically from the card holder's purchasing history. Similarly, modern measurement methods enable detection of systematic deviation in Roberts's merits votes. This "cliometric" approach by no means seeks to displace prevailing historical analysis, but rather to complement it (see Fogel, 1975; Morrison, 1977). Just as Gallup polls cannot capture all nuances of sentiments in the population, our study cannot capture all the nuances of jurisprudential developments (Ho and Quinn, 2009a). What our approach offers is to formalize claims about the nature of Roberts's development based on voting blocs.

First, while many scholars have examined voting alignments in key terms before and after the court-packing announcement (e.g., Corley, 2004, p. 37-39; Currie, 1990, p. 271-73; Cushman,

1997, p. 560, 586-602 nn.22-25; Cushman, 1994, p. 235-36, 244-48 n.254; Cushman, 1998, p. 33-36; Friedman, 1994b, p. 1909-10, 1935-74; McCloskey, 2000, p. 113; Pritchett, 1948, p. 240-63; Schubert, 1959, p. 193-210; White, 2000, p. 198-235), not a single quantitative study using *modern measurement* methods has examined the question of 1937.⁵ For example, internalists stress the fact that Roberts and Hughes already voted with liberals prior to the 1936 term in cases such as *Nebbia v. New York*⁶ and *Home Building & Loan Association v. Blaisdell*,⁷ and that voting patterns were not always consistent with perceived jurisprudential cleavages (see, e.g., White, 2000, p. 201-04). Of course, the fact that Roberts and Hughes sided with the liberals for several cases may only affirm that they were swing voters even before *Parrish*, raising the question of whether Roberts simply continued in this role in 1937. The key question then becomes, “[w]ere the[] decisions [of the 1936 term] startling departures or continuations of earlier trends” (Farber, 2002, p. 1004)? Our analysis formalizes inferences we can draw by comparing voting coalitions, allowing us to test for a sharp rupture in voting behavior.

Second, differences in extant scholarship stem at least in part from disagreement about the relevant cases. Internalists, particularly, have focused on contrasting positions of justices in comparable cases, but considerable disagreement exists over which cases are a priori relevant. Key contested questions, for example, remain about the substantive importance of *Nebbia* and *Blaisdell* (cf. Olken, 2002, p. 311-22; Pepper, 1998, p. 84-146). Our purpose here is not to resolve the murky criteria for case selection (although we offer suggestions on clarification in Section 6.3). By

⁵Cf. Schubert (1959); Lanier (2003); Corley (2004). In pioneering work, Schubert (1959) studies voting coalitions and finds that Hughes and Roberts are less likely to support the liberals prior to the announcement of the plan and less likely to support the conservatives after the plan, but doesn’t address uncertainty or the question of gradual evolution versus abrupt change. Lanier (2003) studies generally the correlates of liberalism for the Court from 1888-1997, estimating a time series (error correction) model using an indicator of the 1937 term or later. Lanier finds “Roosevelt’s proposal of a court-packing plan . . . is *not* . . . associated with an increase in the Court’s economic liberalism” (p. 216, emphasis added). Because outcomes are directionally coded, the model only uses aggregated (as opposed to justice-specific) liberalism for the Court, post-*Parrish* 1936 cases are not excluded, and the 1937 indicator serves as a control in a study not directed at the question of the switch, the implications for the Roberts shift are unclear. Corley (2004) calculates the percentage of a term’s cases favorable to the states for 732 cases (excluding cases pertaining to national legislation) from 1921-37. Corley concludes that “1937 is not the year of the revolution but rather the culmination of the evolution” that started when Taft and Sanford are replaced by Hughes and Roberts in 1930 (p. 52).

⁶291 U.S. 502 (1934) (Roberts, J.) (upholding state milk price control statute against due process challenge).

⁷290 U.S. 398 (1934) (Hughes, C.J.) (upholding state mortgage moratorium statute against contracts clause challenge).

including all nonunanimous cases, the vast majority of which are discussed in the literature, and using model-based weights, we provide one transparent method of case selection and weighting. Effectively, we engage in (Bayesian) learning about the *evolution of the view of the justices across all contested cases of this historical period*. The methods we offer can easily be extended to provide conditional effects for discrete areas of doctrine, although the small set of cases (some twenty per term pre-1937) would significantly reduce the statistical power to detect any shift.

To be clear, there are at least three discrete questions of constitutional history: (a) did an abrupt switch happen? (b) if so, why? (c) did the switch itself contribute to the demise of the court-packing plan? We limit our inquiry to the first.

Our analysis and evidence proceed as follows. Section 3 describes our newly-collected voting data on the New Deal Court. Section 4 provides the methodological intuition of our cliometric approach that formalizes inferences about the relative positions and evolution of the justices based on voting records. To illustrate, we establish the baseline validity of our model for justices serving from the 1931-40 terms, under the strong assumption that the justices' views don't evolve. (Appendix A provides formal details of statistical methodology.)

Section 5 presents our main results in two parts. First, using only the natural *Parrish* Court (i.e., the 1931-36 terms of the same justices who heard the *Parrish* case), we show that there's sharp evidence that Roberts changed in the 1936 term (posterior p -value ≈ 0.000). We illustrate the intuition of how Roberts's shift is identified – even allowing for general changes in case characteristics, the voting patterns are most consistent with a leftward shift of Roberts. Second, we model the 1931-40 terms to account for longer dynamic trends, showing that the evidence for Roberts's shift remains robust. Results suggest that Roberts's transformation is temporary. Evidence of temporariness, however, is somewhat weaker because of what we identify as *bridging sensitivity*: because Roosevelt appointed five new justices from 1937-40, Roberts's trajectory post-1937 has the potential to be more sensitive to our modeling assumptions. That is, given the rapid turnover, bridging the pre and post-1937 Courts is inherently model-dependent, a methodological challenge

we identify as relevant in much work that seeks to bridge actors across institutions and time (e.g., Bailey, 2007; Epstein et al., 2007). The permanence of Roberts’s shift, however, is in many ways irrelevant: regardless of whether Roberts becomes more conservative than in the 1936 term, as a relative matter he fast becomes the most conservative member of the Roosevelt Court, to the right even of McReynolds (see Lawrence, 2005).

Section 6 documents that the transformation most likely occurred at the beginning of the 1936 term, even before *Parrish* was issued and quickly after Roosevelt’s landslide election in 1936. We further conduct a series of sensitivity analysis, documenting that our findings are robust to a form of randomization inference, (prior) specification, and case selection. We also show suggestive (though weaker) evidence that Chief Justice Hughes simultaneously shifted. Section 7 concludes.

3 Data

To assess the Roberts shift, we study nonunanimous cases issued from the 1931 to 1940 terms. To clarify terminology, we refer to “terms” of the Supreme Court, which do not necessarily correspond to the date of decision — the *Parrish* case, for example, was part of the Court’s 1936 term, but was issued on March 29, 1937. The collected data span from 1921 to the present (Ho and Ross, 2010, App. A), but because identification of the Roberts shift stems largely from the period around the 1936 term and to minimize bridging sensitivity, we include cases beginning when Justice Cardozo was confirmed on February 24, 1932, through the end of Chief Justice Hughes’s service on June 20, 1941. We will also refer to the natural “*Parrish* Court” as the set of terms in which all nine justices for the *Parrish* case were active (i.e., the 1931-36 terms). Our results are insensitive to this choice of observation period, for reasons that will become apparent below. The chief observed outcomes are the votes on the judgment, measured as for the majority or the minority.

As the primary breakpoint of interest, we use March 29, 1937, the date when *Parrish* was issued. We refer to cases decided on that day or after as “post-*Parrish*” and cases decided before as “pre-

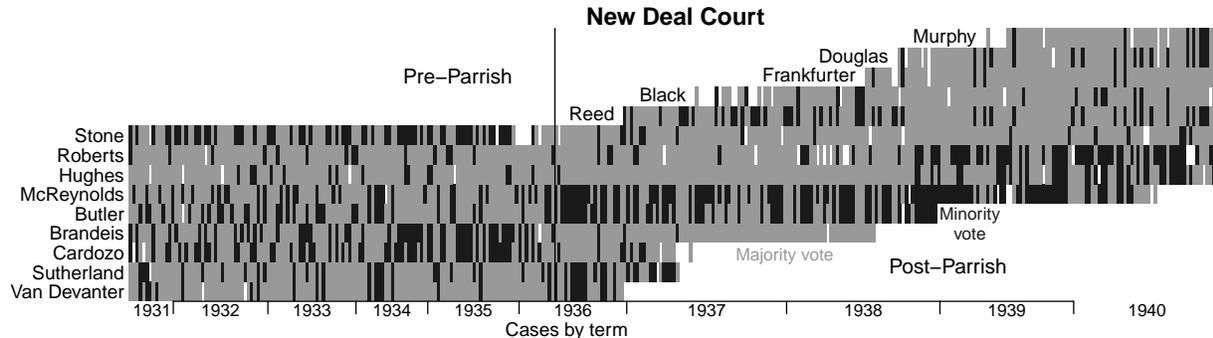


Figure 1: Votes in all nonunanimous cases for the 1931-1940 terms. Justices on the y -axis are sorted by years of service and cases are sorted by dates of decision. Grey (black) cells indicate a majority (minority) vote. The vertical line denotes the *Parrish* case, decided in the 1936 term, with a formal decision issued on March 29, 1937. This figure demonstrates a marked rupture in voting patterns around the time of the case.

Parrish.⁸ The conference vote in *Parrish* had been taken on December 19, 1936, resulting in a 4-4 split, with Stone’s vote expected upon his medical return. Roosevelt unveiled the court-packing plan on February 5, several days after Stone’s return, and the justices delayed the issuance of the opinion until March 29, so as to avert the appearance of retaliating directly against Roosevelt’s announcement. In Section 6.1, we investigate the nature of the timing, showing robustness to choice of breakpoint.

Figure 1 displays our data, with 334 cases sorted by date of decision on the x -axis⁹ and 14 justices sorted by years of service on the y -axis. Each of the 2,813 observations represents one vote cast by a justice for the majority or the minority on the judgment, colored grey or black, respectively. Cells are white when a justice is not serving or participating on a case. The vertical black line represents our breakpoint of interest, separating the pre and post-*Parrish* periods. The vertical cells immediately to the right of the line hence represent the votes in *Parrish* itself, with the four black cells denoting dissenting votes by Justices McReynolds, Butler, Sutherland, and Van Devanter.

Voting patterns shift considerably after *Parrish*. While Justice Stone often dissented pre-*Parrish*, as denoted by the black cells prior to the vertical line in the row corresponding to Justice

⁸Three other nonunanimous cases, *Dugas v. American Surety*, 300 U.S. 414 (1937); *Henneford v. Silas Mason*, 300 U.S. 577 (1937); and *Highland Farms Dairy v. Agnew*, 300 U.S. 608 (1937), were decided on the same day as *Parrish*.

⁹Where multiple cases are issued on the same day, we sort by page on the U.S. Reports.

Stone, he voted overwhelmingly in the majority post-*Parrish*. Figure 1 also illustrates, however, the crucial identification challenge in estimating Roberts’s views over time: beginning with Van Devanter’s retirement after the 1936 term, Roosevelt appointed five justices within three terms, making comparability from pre to post-1937 difficult. (While six justices retire during this observation periods, only five new appointees are observed because of the delay from McReynolds’s retirement on January 31, 1941 to Byrnes’s appointment on July 8, 1941.) Our approach leverages the timing between the decision of *Parrish* in the 1936 term and the new appointees in the subsequent term to capture the Roberts shift.

Two specific patterns emerge in Figure 1. First, the three justices who are widely described as liberals (Stone, Cardozo, and Brandeis) are significantly more likely to vote with the majority starting in the 1936 term. In 1934-35, all three were in the majority less than half the time, but that fraction jumps to 0.79, 0.84, and 0.88 for Cardozo, Brandeis, and Stone, respectively. The four conservatives (Van Devanter, Sutherland, Butler, and McReynolds) exhibit the reverse pattern, dropping in their overall majority rates from 0.78 from 1932-1935 to 0.52 in the 1936 term. Second, Hughes and Roberts retain their critical roles as swing voters in the 1936 and 37 terms, casting votes for the majority in roughly 90% of cases in 1936.

While suggestive, these raw statistics say little about the relative views or the evolution of the justices and the Court, nor do they account for chance variation in the votes. To do so, we turn to modern measurement methods.

4 Methodological Intuition

We now present the non-technical intuition behind our statistical measurement approach, which readers familiar with such methods may wish to skip.¹⁰

¹⁰Details of our statistical methodology are outlined in Appendix A. For additional background see Poole and Rosenthal (1991); Johnson and Albert (1999); Martin and Quinn (2002); Clinton et al. (2004), and Ho and Quinn (2009a).

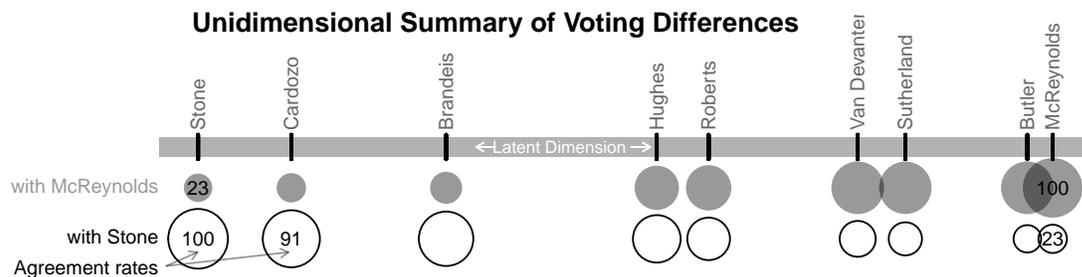


Figure 2: Illustration of location in unidimensional space as a summary of voting differences between justices. The x -axis represents a single latent dimension on which justices might be located, marked by black vertical bars. For example, Justices Stone, Cardozo, and Brandeis are each to the “left” of Chief Justice Hughes. The space between the justices is proportional to how much they agree with each other. The bottom dots plot example pairwise agreement rates between each of the justices and Justices McReynolds and Stone for the *Parrish* Court, with the area proportional to the agreement rate. Agreements with Justice McReynolds generally increase from left to right, while agreements with Justice Stone generally decrease.

4.1 Merits Votes and Spatial Positions

Our primary data consists of votes on the judgment for the majority or minority. While such dichotomous coding of votes abstracts away from much of what the Court does substantively, it provides one key piece of information: the extent to which justices tends to vote together (pairwise agreement rates).

Agreement rates provide useful information about the similarity of voting profiles of two justices. From these pairwise measures of similarity it is possible to summarize differences with a single measure of spatial position. For instance, suppose we simply wanted to compare how similar each of the justices are to Justice McReynolds. The grey dots in Figure 2 plot the pairwise agreement rates of the eight other justices for the *Parrish* Court, where the area is proportional to the agreement rate. Justice Stone votes together with McReynolds in 23% of cases, while Justice Roberts votes together with McReynolds in 55% of cases. The grey dots almost uniformly increase moving from left to right. Similarly, the hollow dots plot the agreement rates with Justice Stone. Justice Cardozo, for example, agrees with Justice Stone in roughly 91% of cases, while Justice Van Devanter agrees with Stone in only 36% of cases. Interestingly, the orderings are almost exact inverses of one another: as the agreement with McReynolds increases, the agreement with Stone

decreases. Indeed, if we are willing to place Justice Stone to the left of Justice McReynolds, such rates allow us to locate the justices proportional to their similarity in voting. This is what is plotted in the grey bar of Figure 2, the “latent” (i.e., unobserved) dimension. The dark vertical bars represent each justice’s location in this space, which roughly speaking summarize all 35 unique pairwise agreement rates. Figure 2 thereby illustrates the key intuition that differences in merits votes can be summarized as positions in a unidimensional space (the grey bar), which provides one formalization of differences between justices.

While the actual model builds in much more nuance, as we explain next, it leverages the key information contained in voting patterns to estimate the location of justices in a single dimension. Justices proximate in this underlying dimension tend to vote together, compared to justices distant from one another. In political science, these positions are often referred to as “ideal points,” meaning the *ideal position* in a single dimension that a decisionmaker prefers.

4.2 Accounting for Case Differences

How does our approach account for differences in cases? For each case, we estimate two parameters that account for (a) the degree of dissensus amongst the justices, and (b) the degree to which locations in the underlying dimension are associated with voting. For convenience we will refer to the former (a) as the “dissent” parameter, as it approximately models the the number of *dissents* in a case. Similarly, we will refer to the latter (b) as the “valence” parameter, as it approximately models how and to what degree the dissents are driven by the *valence* of the underlying dimension.¹¹ Votes are modeled probabilistically as a function of these case parameters and justice locations.

To illustrate, Figure 3 illustrates the model at the case level. Each panel in the plot depicts the observed votes of the justices, coded as 1 on the y -axis if for the majority and 0 if for the minority. In addition, each panel plots the model-based estimate of the probability of voting for the majority

¹¹In educational testing, these parameters are referred to as “item difficulty” and “item discrimination” parameters. From a factor analytic perspective, the valence parameter is the “loading factor.”

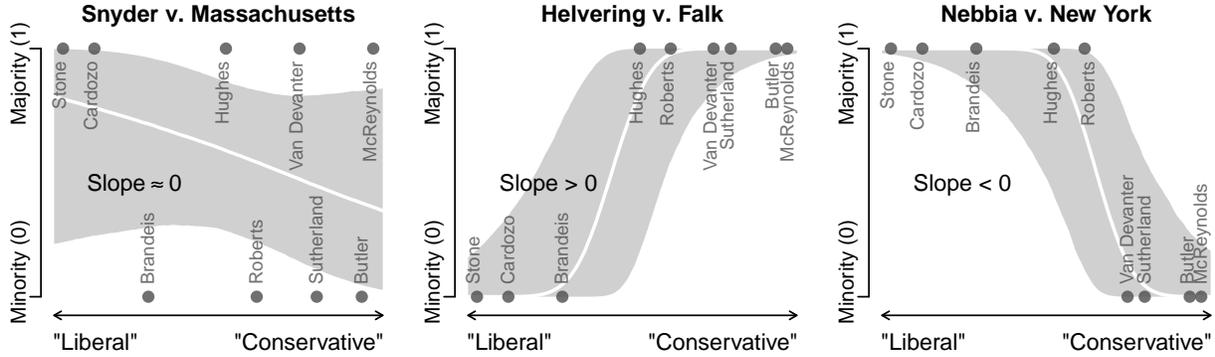


Figure 3: Modeling the probability of judicial votes as a function of a latent dimension. The x -axis represents the latent dimension. The y -axis represents the probability of a vote for the majority, ranging from 0 to 1. The grey dots are the observed votes by each justice. The white lines represent the estimated probability as a function of the latent dimension (with 95% credibility bands). A positive slope means that the majority is estimated to be “conservative” and a negative slope means that the majority is estimated to be “liberal.” The steepness of the slope reflects how much we learn about differences in the ideal points of the justices. Since slopes are not equal, cases contribute differentially to ideal points. *Snyder v. Massachusetts* in the left panel features an atypical majority coalition and thus has a slope that is not statistically different from 0. It therefore contributes little information about the ideal points. *Helvering v. Falk* on the other hand, has a sharp positive slope reflecting the larger amount of information this case contains for our estimation strategy. *Nebbia v. New York* contributes a similar amount of information, although here there is a liberal majority and hence a negative slope.

position as a function of a justice’s ideal point (with the grey bands representing uncertainty). For example, in the left panel, Justice Van Devanter is predicted to have a roughly 50% probability of voting with the majority, since the white curve roughly intersects 0.5 where Justice Van Devanter is located. This also provides us with a more formal understanding of the dissent and valence parameters. The dissent parameter affects the overall height of the curve – if it were a flat line, a dissent parameter of $\frac{4}{9}$ would indicate that four justices are predicted to dissent.¹² The valence parameter is akin to the slope of the curve – a flat line would indicate that dissents are entirely unrelated to the primary dimension characterizing differences in the justices, while a positive slope would indicate that justices to the right of the space are more likely to vote for the majority.

Knowledge of the relative frequency of voting coalitions allows one to discern which cases are

¹²As a purely technical matter, this last statement is not quite correct. A more accurate statement would be that if the curve was flat and the dissent parameter was equal to 0.1397 then the curve would predict that 4 justices would dissent. The reason for this is the such predictions are generated via the standard normal distribution function $\Phi(\cdot)$ and the expected fraction of dissents is given by $1 - \Phi(0.1397) \approx \frac{4}{9}$.

typical and therefore have steep valence parameters. Not all cases are created equal. Consider *Snyder v. Massachusetts*¹³, plotted in the left panel of Figure 3, in which a criminal defendant challenged his murder conviction on the grounds that the denial of his petition to be present at the jury’s view of the crime scene violated due process. *Snyder* involved atypical voting coalitions of the *Lochner* Court, with Justices Stone and Cardozo joining Justice McReynolds in the majority that affirmed the conviction, and Justice Brandeis joining Justices Sutherland and Butler in dissent. If votes provide us with information about the primary dimension driving differences between the justices in a single space, the flat slope shows that *Snyder* may not be very indicative of this dimension. Accordingly, because *Snyder* features an atypical majority coalition, the valence parameter (slope) is relatively flat (statistically indistinguishable from 0 as denoted by the grey bands). The valence parameter can thereby be interpreted as the model-based weight we give to the case in learning about primary differences between justices.

On the other hand, the remaining two cases in Figure 3 feature more typical voting cleavages. *Helvering v. Falk*¹⁴ features what would typically be considered a conservative majority (sustaining an income tax depletion allowance for trust beneficiaries) while *Nebbia* features a liberal majority (upholding price regulation).¹⁵ Our model captures this fact in the valence parameters—the positive slope in *Falk* and the negative slope in *Nebbia*—without manually judging the “liberal” or “conservative” direction of specific case outcomes. In *Falk*, for example, Justices Butler and McReynolds, who occupy positions at the right of the Court are predicted to join the majority with probabilities approaching 1, while Justices Stone and Cardozo, who occupy positions at the left of the Court are predicted to join the majority with probabilities approaching 0.

This approach towards accounting for case differences has three particular virtues. First, it allows for the possibility that votes in some cases, such as *Snyder*, may be unrelated to the primary

¹³291 U.S. 97 (1933).

¹⁴291 U.S. 183 (1934).

¹⁵While it may have been unexpected in *Nebbia* for Chief Justice Hughes and Justice Roberts to join Justices Cardozo, Stone, and Brandeis in upholding New York’s milk price controls, the division predictably pitted the liberal wing of the *Lochner* Court against the conservatives, with Chief Justice Hughes and Justice Roberts occupying the middle.

underlying dimension. Agreement on a representative case, such as *Nebbia*, producing steep slopes / high valence parameters, may provide strong evidence that justices voting together are proximate in space. On the other hand, agreement on a case with very unusual voting blocs, such as *Snyder*, producing a slope / valence parameter statistically indistinguishable from 0, may not provide much information about systematic differences in the justices. Second, because our approach employs a probability model, we obtain direct estimates of uncertainty for quantities of interest (e.g., relative positions of justices). This allows us to address crucial concerns of whether particular cases represent broader patterns or just chance variation. Third, no strong assumption about the direction of votes on any particular case (i.e., in a “liberal” or “conservative” direction) is assumed – the valence of the case is estimated based on the frequency of voting coalitions. .

We emphasize that the estimated spatial location of the justices interpretation does not warrant any strong interpretation of what drives judging. The dimension simply represents one method of characterizing the most salient difference in votes on the judgment between the justices, much like an SAT score best characterizes the propensity of students to answer questions correctly. The dimension is fixed by assuming two justices to be on opposite sides of the origin, but there is no inherent meaning to the cardinal scale. For instance, in our primary analysis we constrain Justice Brandeis, conventionally conceived of as one of the three liberal justices of the *Lochner* Court, to be on the opposite side of the scale as Justice Butler, conventionally conceived of as conservative. For convenience and in accordance with the literature, we use the labels “liberal” and “conservative” as shorthand to refer to directions of this latent scale (but cf. White, 2000). Use of such terminology, however, does not warrant or imply an “attitudinal” interpretation of judges as policymakers, as the scale may just as well represent jurisprudential differences (Ho and Quinn, 2009a).

4.3 Bayesian Learning

While Figure 3 provides some basic intuition about the model at the case-level, our statistical approach *jointly* estimates the case-specific and justice-specific parameters based on the observed

votes. If we were given the positions of the justices, it would be trivial to estimate the dissent and valence parameters using standard techniques (e.g., logit, probit). Joint estimation becomes more difficult, exacerbated by the large number of parameters relative to data ($334 \text{ cases} \times 2 \text{ parameters per case} + 14 \text{ justice locations} = 682 \text{ parameters}$ compared to 2,813 votes). The intuition, however, is straightforward: as each case is decided, the votes provide information about primary underlying differences in the justices. We hence update based on this data to draw inferences about the relative locations of the justices.

Figure 4 provides intuition as to how this form of learning (Bayesian updating) occurs. The top panel presents votes from all 27 nonunanimous cases in the 1933 term, with justices in rows and cases sorted chronologically in columns. As in Figure 1, light and dark grey cells denote that a justice voted for the majority or minority, respectively. For example, the first case is *Krauss Bros. Lumber v. Dimon*, 290 U.S. 117 (1933), in which Justices McReynolds, Sutherland, Butler, and Roberts dissented. The bottom panel presents our beliefs about the relative locations (of ideal points) of the justices after each case is issued. The shading in each cell corresponds to the relative rankings of each justice. Our “prior” before observing any cases is that the justices are identically located, corresponding to the uniform medium grey in the first column. The second column of the bottom panel imposes the “directional” prior that Justices Butler and Brandeis are on opposite ends of the dimension, corresponding to the dark grey for Justice Brandeis, the light grey for Justice Butler, and uniform medium shade for the rest. This directional prior allows us to interpret the dimension as “liberal” or “conservative,” although again we use such terms only as shorthand consistent with scholarship on the Hughes Court, with no implication for whether policy preferences or jurisprudence is driving the decision.

Starting with the third column, we update based on cases issued. After *Krauss* is decided, we essentially have two blocs of justices, although very imprecisely estimated: Justices McReynolds, Butler, Sutherland, and Roberts in light grey and Chief Justice Hughes, and Justices Van Devanter, Brandeis, Stone, and Cardozo in darker grey. With each additional case, our beliefs about the

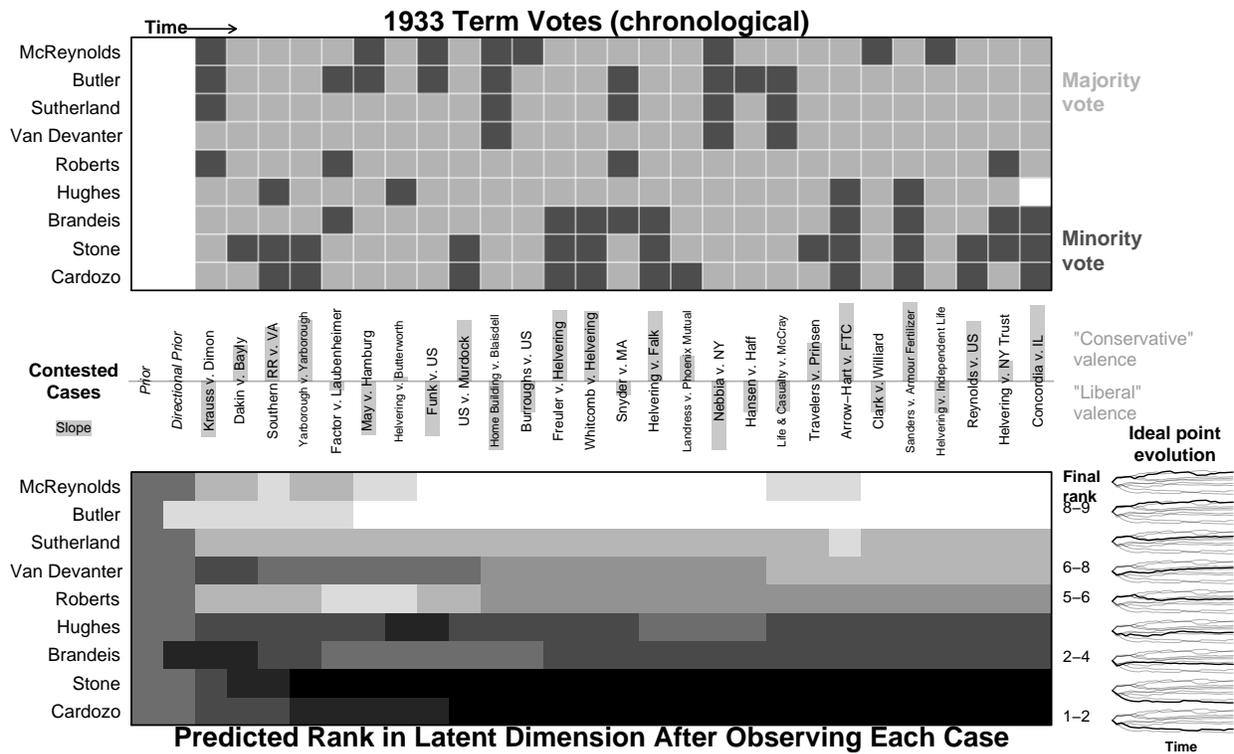


Figure 4: Illustration of “Bayesian learning” about ideal points of justices. This figure illustrates how we update beliefs about the ranks of the justices as nonunanimous cases from the 1933 term are issued chronologically. The top panel presents data on votes cast in each in the order issued. Each column represents one case, and the grey and black cells represent votes by each of the nine justices (in rows) for the majority or minority, respectively. The bottom panel represents the predicted rank of justices at each point in time, where Justice Brandeis is assumed to be on the opposite sides of the median rank from Justice Butler, solely for directional interpretation. The first column of uniform grey represents our “prior” of no differences between the justices, associated with no votes on the top panel. The second column represents a “directional prior,” assuming Brandeis and Butler to be ranked on opposite sides of the median rank. As each case is decided our belief is “updated.” The bars behind the names of cases represent how much weight is given to each case and the estimated direction (i.e., the slope in each case model). Cases with low weight are those that are unconventional according to this ranking (e.g., *Factor v. Laubheimer*, in which Justices Brandeis, Roberts, and Butler dissented). The bottom right presents the evolution of ideal points of each justice in the latent dimension, contrasted in each instance with the other justices.

relative positions tend to become more precise. For example, in the fourth divided case of the term, *Yarborough v. Yarborough*, 290 U.S. 202 (1933), Justices Stone and Cardozo dissented, pushing them toward the boundary of the ranks.

On the other hand, some cases present unconventional voting blocs. *Snyder*, for example, the fourteenth case of the term, receives very little weight. The middle panel of Figure 4 plots the slope estimated for each case, which represents the amount of weight placed on the case in inferring the latent position of the justices, as well as the estimated valence of the case. The longer the grey bars, the more weight. Bars pointing upwards are estimated to have a “conservative” valence, while bars pointing downwards are estimated to have a “liberal” valence. For *Snyder*, the slope is effectively zero, and we can see that the associated predicted rank does not change after observing the case. After observing the dissents of Justices Brandeis, Stone and Cardozo in the next case, *Falk*, however, we infer that Brandeis is closer to Justices Stone and Cardozo than Chief Justice Hughes. After observing all of the cases of the term, we are left with essentially five clusters of justices in the last column: ranks (1-2) Justices Stone and Cardozo; ranks (2-4) Chief Justice Hughes and Justice Brandeis; ranks (5-6) Justice Roberts; ranks (6-8) Justices Sutherland and Van Devanter; ranks (8-9) Justices McReynolds and Butler.

The bottom right panel plots an alternative way to visualize the evolution of justice ideal points over the 1933 term. Rather than presenting the estimated ranks, we plot the estimated position in the single dimension. The dark line presents the justice’s estimated position on the y -axis against time in the 1933 term on the x -axis, with grey lines representing the other justices. While all justices start out in the same position (due to the noninformative prior), the conservatives drift upwards, while the liberals drift downwards as cases are decided.

So far we’ve limited ourselves only to the 1933 term. Figure 5 therefore presents results from our basic model pooling all divided cases from the 1931-40 terms. Like the bottom right panel of Figure 4, the short vertical bars represent our best estimate as to the position, not the rank, with horizontal bands capturing uncertainty in the location (95% credible intervals). (Justices are sorted

Static Estimates, 1931–40

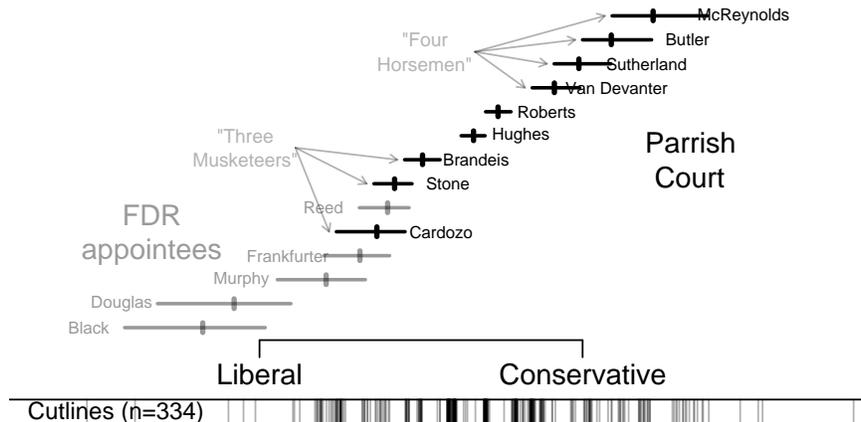


Figure 5: Static (median) ideal point estimates with 95% credible intervals for the 1931-40 terms. Justices are sorted from left to right by median ideal point. Ideal points for justices serving on or before the *Parrish* term are plotted in black, and others in grey. Bottom strip presents estimated splits between the majority in minority in 334 cases (κ from Appendix A). Stationary ideal points capture key conventional understandings of the jurisprudence of these justices – i.e., the major cleavages in the *Parrish* Court between the “Four Horsemen” and the “Three Musketeers”; the center of the Court occupied by Roberts and Hughes; and the liberal turn of the Roosevelt appointees.

from left to right.) Justices serving during the 1936 term are in black, and Roosevelt appointees are in grey. The results are facially reasonable, and capture the major cleavages of the *Parrish* Court, as well as the shift to the left with the Roosevelt appointees. The four conservatives (the “four horsemen”) are estimated to be on the right of the latent dimension, and the three liberals (the “three musketeers”) are estimated to be to the left of the other *Parrish* justices, with Hughes and Roberts occupying considerable space in the middle. Moreover, the Roosevelt appointees shift the Court significantly to the left, such that *Parrish* liberals appear to be considerably to the right of the likes of Black and Douglas. While we can’t infer much about Roberts other than that he appears to be the median justice of the *Parrish* Court, the consistency of these baseline estimates with qualitative assessments shows that formal voting patterns on the merits provide good leverage over assessing the viewpoints of the justices. Lastly, the bottom strip of Figure 5 provides a summary of the case parameters. The strip plots the “cutline,” namely the position estimated to split the majority from the minority in any given case. This corresponds to the location where the

white curve in Figure 3 would intersect with a horizontal line of a 0.5 probability of voting for the majority. These cutlines provide us a sense of where most of the action is: the thick cluster of cutlines between Justice Brandeis and Chief Justice Hughes displays the conventional pre-1937 6-3 cleavage that led to the court-packing plan.

We now turn to applying this general methodology to study the switch in time.

5 Results

We present our results in two steps. In Section 5.1, we limit our focus on the *Parrish* Court. This data reveals the Roberts shift in the difference between pre-*Parrish* and post-*Parrish* votes of the 1936 term. In Section 5.2, we relax strong assumptions of constant viewpoints over time. We show that both for the *Parrish* Court and the 1931-40 period the Roberts shift remains robust, but that it appears temporary.

5.1 Static Estimates

While demonstrating that our approach captures key differences on the *Lochner* Court, the baseline estimates in Figure 5 assume that the justice positions are constant over time. Our question of interest, however, entails the evolution of Roberts’s views over time. Because Roosevelt appointees complicate scaling the justices over time (a problem of bridging sensitivity we identify more fully in Section 5.2), we focus here on the key period of the *Parrish* Court spanning the 1931-36 terms and particularly Roberts pre and post-*Parrish*. To do so we fit the same model of Section 4.3, relaxing only the pooling assumption of Roberts by fitting separate ideal points for the pre and post-*Parrish* period.

Figure 6 presents results. In grey, we present the other justices, which exhibit the key cleavage of the *Parrish* Court between Stone, Cardozo, and Brandeis on the one hand and Van Devanter, Sutherland, Butler, and McReynolds on the other. The key evidence here is that Roberts’s ideal point shifts dramatically to the left during the post-*Parrish* period. While he is located between the jurisprudence of Chief Justice Hughes and Justice Van Devanter pre-*Parrish*, his votes become

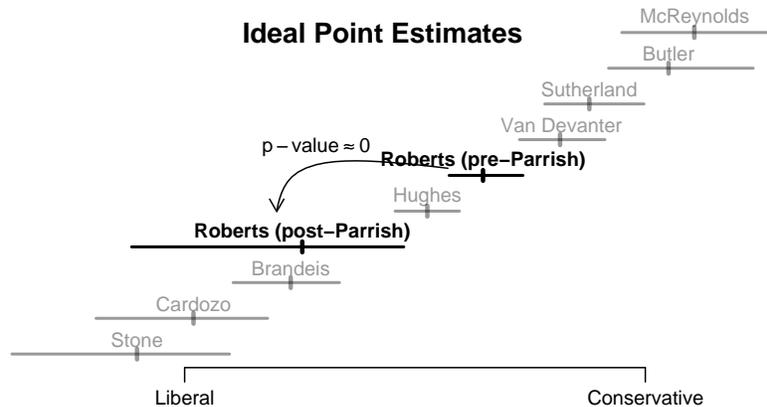


Figure 6: Static ideal point estimates for the *Parrish* Court, 1931-1936 terms, with separate ideal points for Roberts pre and post-*Parrish*. This figure shows the considerable shift of Roberts, from to the right of Hughes pre-*Parrish* to somewhere closer to Brandeis post-*Parrish*. The posterior (one-tailed) p -value that Roberts shifts is approximately 0.

statistically indistinguishable from Justice Brandeis post-*Parrish*. The intervals are wider, since there are only 21 nonunanimous cases in that period, compared to 131 pre-*Parrish* cases. The fact that our model-based method captures such variability due to sample size is a virtue of the approach. More importantly, the pre and post credible intervals don't overlap – the posterior probability that Roberts's post-*Parrish* ideal point is to the right of his pre-*Parrish* ideal point is effectively 0 (posterior p -value ≈ 0.000).¹⁶ This provides strong evidence of a Roberts 1936 transformation.

To illustrate how this effect is identified, we can refer back to Figure 1, focusing particularly on the trends before and after *Parrish*. (Because Stone was on medical leave the month before *Parrish* was decided, his vote is missing on a handful of pre-*Parrish* cases.) The pre-*Parrish* cases display considerable disagreement between Brandeis and Roberts, that disagreement vanishes post-*Parrish*. Aside from two solo dissents by Roberts, Brandeis and Roberts agree in every case post-*Parrish*, while they agree in just over half of all cases pre-*Parrish*. Similarly, the likelihood that Roberts votes with Van Devanter in close cases decreases significantly.

To summarize these raw patterns, Figure 7 plots the pairwise agreement rates between Justice

¹⁶The prior distributions for ideal points are standard normal distributions, so the pre and post ideal points are, if anything, attenuated to 0, leading our estimate of the difference to be biased towards 0.

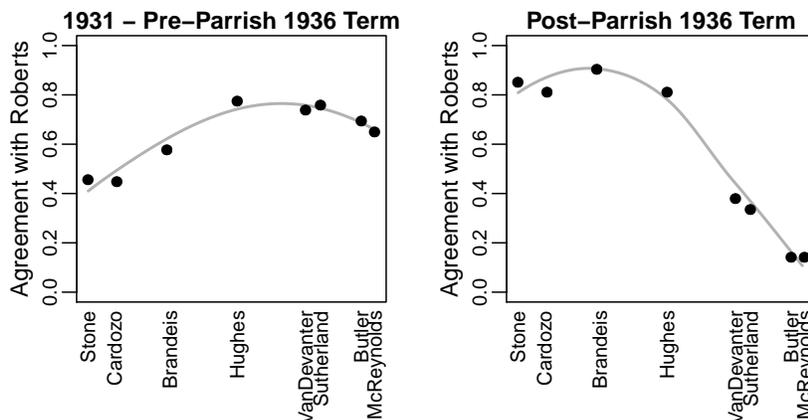


Figure 7: Voting Agreement with Justice Roberts in the *Parrish* Court. The left panel presents agreement rates based on cases in the 1931 term to the pre-*Parrish* 1936 term while the right panel presents agreement rates based on cases post-*Parrish* in the 1936 term. Grey curves present smoothed (loess) fit, depicting the sharp divergence between Roberts and the conservatives post-*Parrish*.

Roberts and the other eight justices. The left panel plots rates across cases from the 1931 term to the pre-*Parrish* 1936 term, while the right panel plots these for post-*Parrish* 1936 term cases. Justice Roberts is the pivotal voter, but agrees more often with the Horsemen prior to the 1936 term. For example, he agrees with Justice Van Devanter in roughly 74% of cases pre-*Parrish*, but that rate drops to slightly more than 40% post-*Parrish*. Justice Roberts’s agreement rate with all the conservatives similarly plummets after *Parrish*, while it increases from 45-58% to 81-90% with respect to the three liberals. The modes of the smoothed (loess) curves provide a rough sense of where a model is most likely to place Roberts for each period. The mode is in the middle pre-*Parrish* and moves towards the left of the Court post-*Parrish*.

Table 1 provides another way to understand the results, reporting Robert’s rate of agreement with unanimous musketeer or horsemen blocs across terms. For example, Roberts agrees with a unanimous musketeer bloc in 37% of of cases (10 of 27 cases) in the 1933 term, compared to 73% of cases in the 1936 term. The fourth and fifth columns report p -values from simple significance tests of the difference in agreement rates for adjoining terms. We detect no statistically significant ruptures from 1932 to 1935, but p -values are close to 0 comparing agreement between the 1935 and 1936 terms. In short, voting patterns for the *Parrish* Court alone strongly suggest a shift to the

Term	No. cases	Roberts unanimous with Musketeers		Roberts unanimous with Horsemen	
		Prop.	<i>p</i> -value	Prop.	<i>p</i> -value
1931	12	0.75		0.25	
1932	29	0.48	0.171	0.41	0.480
1933	27	0.37	0.430	0.52	0.592
1934	22	0.23	0.358	0.64	0.563
1935	28	0.18	0.732	0.68	0.773
1936	33	0.73	0.000	0.18	0.000

Table 1: Justice Roberts’s agreement with the “Three Musketeers” and the “Four Horsemen” over the course of the *Parrish* Court. The first column presents the term. The second column presents the number of nonunanimous cases per term in our observation period (after Justice Cardozo starts serving on the Court). The fourth column (Prop.) presents the proportion of times that Roberts and Justices Brandeis, Stone, and Cardozo are unanimous in a case. Note that when Justice Roberts is in agreement with one Musketeer but not the other, this table does not count this as an “agreement.” Missing votes of any justice (e.g., due to recusals or illness) are not counted as a disagreement with the bloc. The fifth column (*p*-value) presents the *p*-value with Fisher’s Exact Test comparing the agreement rate of that term to the preceding term. For example, the *p*-value of 0.171 tests the rate of 0.48 (=14/29) rate of Musketeer agreement in the 1932 term to the 0.75 rate (=9/12) of Musketeer agreement in the 1931 term (i.e., a 2×2 contingency table of whether Roberts agreed with the Musketeers in columns and the 1931 and 1932 terms in rows). The sixth and seventh column (Prop. & *p*-value) present analogous calculations with Justices McReynolds, Van Devanter, Sutherland, and Butler. Note that because some cases do not necessarily divide the Court exactly along Musketeer-Horsemen lines, these rates do not necessarily sum to one. The 1936 term proves to be a clear breakpoint in Justice Roberts’s agreement with the two main voting blocs on the Court. This analysis most closely tracks that of (Schubert, 1959).

left of Justice Roberts in the 1936 term.

5.2 Dynamic Estimates

While the static ideal point analysis above is parsimonious, it also assumes away a key quantity of interest, namely, the evolution of Roberts over the course of the 1930s. To address this, we make use of more sophisticated measurement models that allow the ideal points of justices to evolve over time.¹⁷ The advantage of such an approach is that it allows us to estimate not just a rupture in Roberts’s voting behavior in the 1936 term as the static model did, but also to obtain a sense of Roberts’s ideal point trajectory before and after *Parrish*. This modeling approach incorporates accounts of some “judicial movement . . . [as] a random walk” (Friedman, 1994b, p. 1896).

¹⁷We make use of an adaption of the model of Martin and Quinn (2002), sketched in more detail in Appendix A.2.

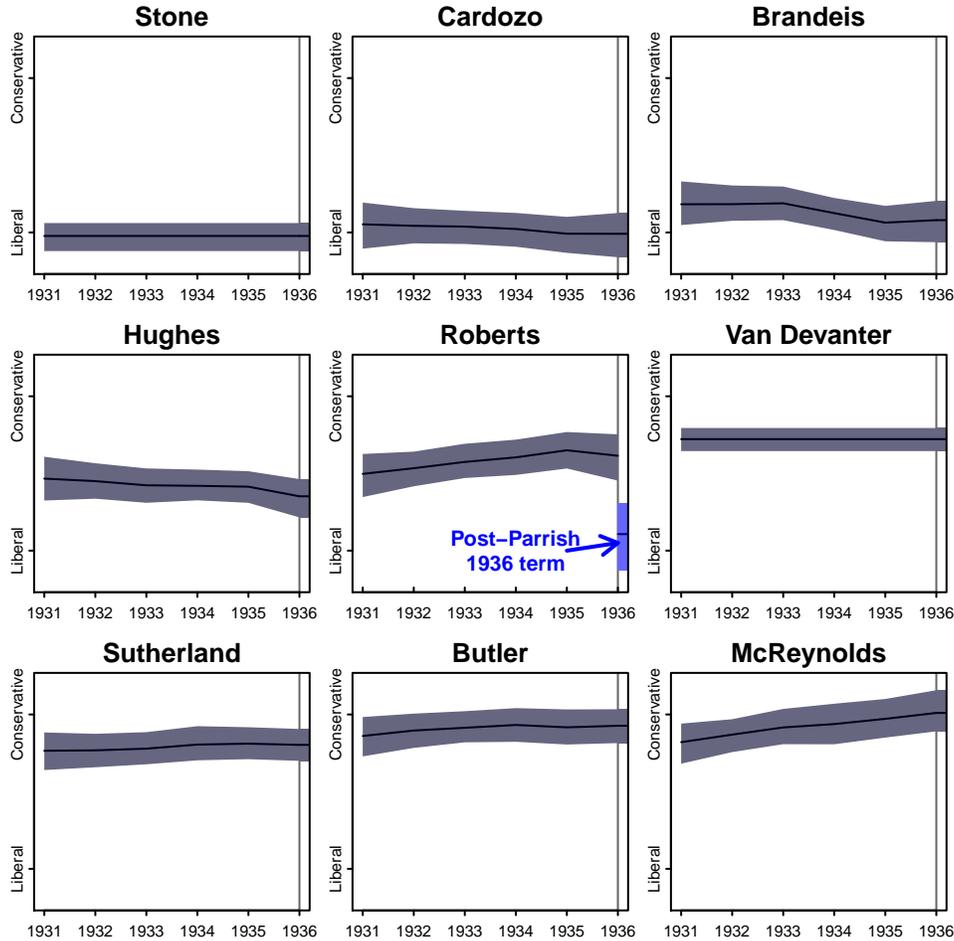


Figure 8: Dynamic ideal point estimates for the *Parrish* Court. Solid lines are posterior median and the shaded regions are the pointwise 95% posterior credible bands. Note the dramatic shift to a more liberal position by Justice Roberts after *Parrish*.

5.2.1 The *Parrish* Court

We begin by restricting ourselves to the *Parrish* Court. The advantage of looking at just these nine justices is that *relative* ideal point locations are easily estimated since they vote together on a large set of cases.¹⁸

Figure 8 displays the estimate (posterior median) of each justice’s ideal point along with uncertainty bands (central 95% pointwise credible bands). The model assumes that the ideal points of Stone and Van Devanter are constant over time and that the ideal points of McReynolds and Cardozo are above and below 0, respectively. The results square well with extant characterizations

¹⁸Positions are well-identified without overly strong modeling assumptions.

of the Court. The three most liberal members of the Court (Stone, Cardozo, and Brandeis) are clearly separated from the four most conservative members (Van Devanter, Sutherland, Butler, and McReynolds). More interestingly, we detect modest evidence that the Court was growing increasingly polarized during the period from the 1931 to 1935 terms with McReynolds, Butler, Sutherland, and Roberts all showing some evidence of drifting to the right while Brandeis and perhaps Cardozo and Hughes may have been drifting to the left.

Most interesting for our present purposes is the substantial change that we estimate in Roberts's ideal point after the *Parrish* decision. As in the static analysis above, the probability that Roberts's post-*Parrish* ideal point is to the left of his pre-*Parrish* ideal point is essentially 1. Substantively, the magnitude of the shift is striking as Roberts's ideal point becomes quite close to those of Stone, Cardozo and Brandeis.

5.2.2 1931-1940 Terms

While the analysis of the *Parrish* Court greatly informs our understanding of the break in Roberts's behavior in the 1936 term, it doesn't capture Roberts's evolution in later terms. To do so — and thus provide a more comprehensive view of Roberts as well as the Roosevelt realignment — we now expand our analysis to all nonunanimous cases between the time Justice Cardozo first joins the Court and the time Chief Justice Hughes leaves the Court (cases in the 1931-1940 terms).

Term-by-Term Analysis. We begin by fitting a sequence of independent static models to the data for each term. Because a separate model is fit to each term, the results are not directly comparable across terms because there is no common reference point across cases. One approach to dealing with this problem is to pick a reference point ex post and then adjust the location of the ideal point estimates across the various model fits to enforce constancy of this reference point.

Figure 9 displays the term-by-term static model results after enforcing two different normalizations. In the top panel it is assumed that Justice Stone's ideal point is constant over time. Again, the pre-1937 term results square well with all the results seen previously. In particular, we note

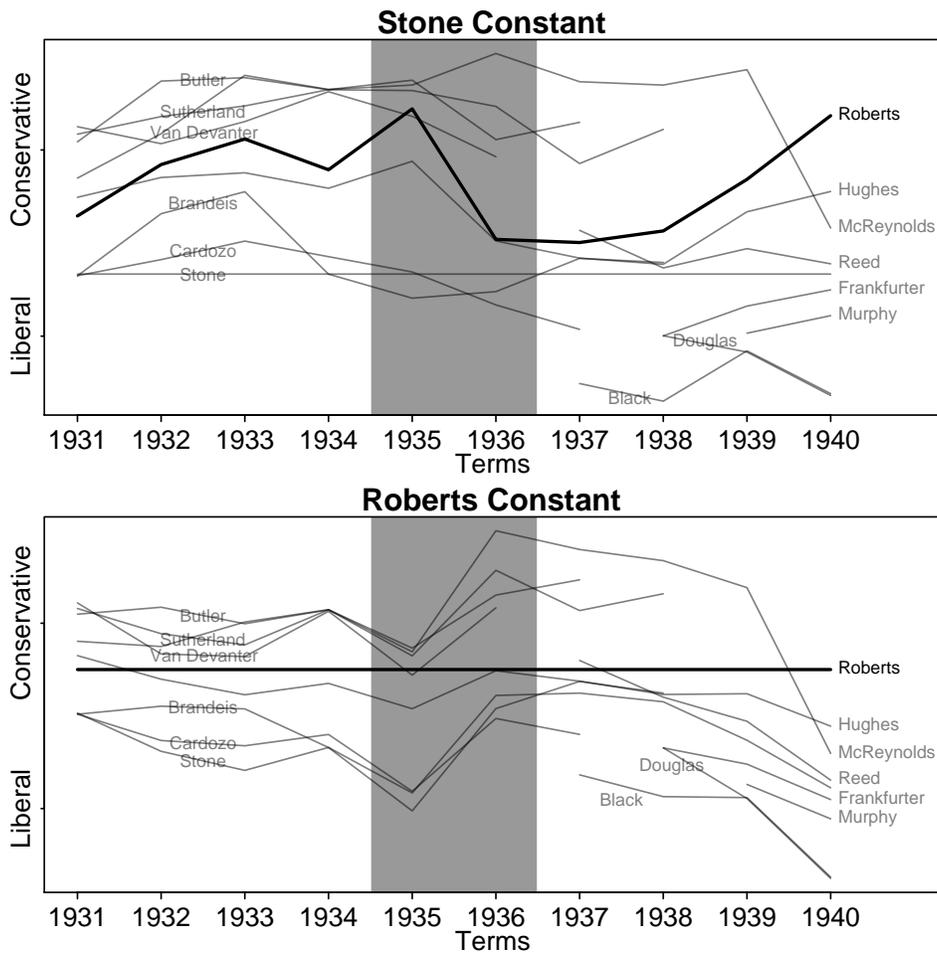


Figure 9: Term-by-term estimates that are anchored in different ways. Separate models are fit to the cases of each term. To anchor these unpooled estimates, the top panel assumes that Justice Stone’s position remained constant while the bottom panel assumes that Justice Roberts’s position is constant. Assuming Stone constant, the top panel depicts Roberts’s sharp shift leftward during in the 1936 term. Assuming Roberts constant, the bottom panel depicts sharp shifts upwards for every other justice during the 1936 term.

the sharp change to the left in Roberts's ideal point in the 1936 term. Under this normalization, it appears that Roberts maintains this liberal streak for two more terms before moving back to the right fairly rapidly in the 1939 and 1940 terms.

Another possible reference point is the ideal point of Justice Roberts. By normalizing his ideal point to be constant in time we can get a sense of the relative positions of the other justices to him over time. These results are presented in the bottom panel of Figure 9, depicting two important patterns. First, if Justice Roberts's voting behavior truly was constant during this time period then *all* of the other justices on the Court exhibited a very large and significant shift to to the right in the 1936 term. Substantively this makes little sense and thus provides more evidence that Roberts behavior changed during this period. Further, if Roberts was constant then all of the other justices were becoming more — sometimes dramatically more — liberal in the 1940 term. Again, this seems hard to reconcile with extant knowledge about the 1940 term.

What these results highlight is the degree to which inferences about movement of individual justices depend on assumptions about a baseline. In situations where the membership of the Court is constant over time it may be reasonable to assume that the ideal points of some justices remain constant over time, or at least that on average a justice is most likely to be at the same location from one period to the next. Such assumptions allow one to make inferences about individual change. In situations when only a single individual is replaced on the Court, the constancy (or assumed smoothness in viewpoint evolution) of the remaining eight justices provides a strong enough anchor to draw inter-temporal comparisons. However, when the membership of the Court changes and the only justices who serve in all terms are those for whom one is interested in estimating change over time, it becomes impossible to separate such individual-level change from the effects of membership replacement unless much stronger assumptions are made. We call this *bridging sensitivity*, as results are sensitive to assumptions imposed to bridge justices across time.

Because of the dramatic change in the Court's membership during this time period — when five leftward leaning Roosevelt appointees replaced more rightward leaning justices — it is difficult

to disentangle change in individuals from compositional change. Figure 9 points out, for example, that if we posited Stone to shift to the right, McReynolds might stay more constant, but the Roosevelt appointees would move to the center.¹⁹ While the results from the term-by-term analysis of Roberts’s post-1937 evolution should not be viewed as definitive, a reasonable inference is that Roberts eventually moved to the right after 1937. Otherwise, all other justices moved to the left, which seems unlikely.

While the term-by-term analysis is instructive in that it relies on easily understood assumptions, clarifies the issue of bridging sensitivity, and provides reasonably strong evidence to suggest there was a break in Roberts’s behavior in 1936 and perhaps a drift to the right in the post 1937 period, it is not without some problems. Most notably, because the ideal points in each term are assumed to be independent of the ideal points in all other terms, the estimates have much more estimation uncertainty than may be reasonable.

Pooling Terms. To get a more accurate statistical summary of Roberts’s behavior, we fit the dynamic ideal point model discussed above to the 1931-1940 data. In doing so, we split Roberts’s votes into pre-*Parrish* and post-*Parrish* votes and treat the two series of votes as distinct. This allows us to test for a break in Roberts’s behavior at the time of the *Parrish* decision. This model partially pools information from votes across terms and thus provides a middle ground between either the rigidity of the constant ideal point model or the high variability of the estimates from the term-by-term analysis, both of which are special cases of this more general approach (see generally Gelman and Hill, 2007).

Figure 10 plots point estimates of the ideal points from the dynamic model fit to the 1931-1940 term data. The results square well with the conventional wisdom. The three musketeers (Brandeis, Cardozo, and Stone) and the Roosevelt appointees occupy positions to the left of the space, while the four horsemen (Butler, McReynolds, Sutherland, and Van Devanter) take positions to the right.

¹⁹See, e.g., Konefsky 1958, p. 1136 (“in the conflicts that developed [in the post-1937 Court], the erstwhile liberal dissenter [Stone] found himself at odds with both the growing conservatism of an Owen J. Roberts and the unorthodox tactics and attitudes of the Court’s new liberal faction.” He came to be “cast in the role of a carping conservative.”).

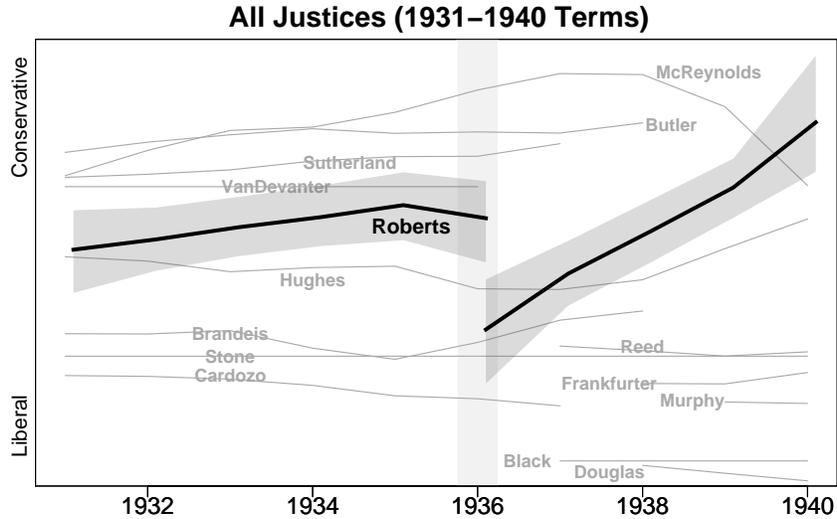


Figure 10: This panel plots median dynamic ideal points for all justices, with separate pre and post-*Parrish* trends for Roberts. The solid lines represent the pointwise posterior median ideal points. The grey bands provide pointwise 95% credible regions for Justice Roberts, who experiences a sharp shift to the left in the 1936 term.

Finally, Chief Justice Hughes and Justice Roberts generally take positions toward the middle of the space.

More importantly, consistent with earlier results, Roberts shifts leftward immediately after the *Parrish* decision. The substantive magnitude of Roberts’s shift again is large. Not only is he much more likely to vote with Brandeis, Cardozo, and Stone after the *Parrish* decision than before, but Roberts also moves to the left of Chief Justice Hughes (but cf. Section 6.4).

To get a sense of the statistical uncertainty associated with the change in Roberts’s behavior, Figure 10 overlays uncertainty bands for Justice Roberts in grey. Here we see that there is no overlap between the regions on either side of the *Parrish* break point. This is strong evidence of a rupture in Roberts’s voting behavior around this time.

Finally, Figure 10 suggests that Justice Roberts’s shift was temporary. After the strong shift leftward in the 1936 term, Roberts quickly moves back to the right, as seen by the upward movement of the dark line in Figure 10. Interestingly, Chief Justice Hughes also appears to be drifting to the right during the later 1930s. Of course, we must be cautious in drawing inferences about post-1937 movement because of bridging sensitivity. As a factual matter, we know that after the 1938 term

Chief Justice Hughes and Justices McReynolds and Roberts started voting together much more frequently. This could be the result of a move toward the center by McReynolds, a move to the right by Hughes and Roberts, a change in the nature of the cases coming before the Court, or some combination of all of these factors. The observed voting data alone do not allow us to determine the relative importance of these factors. Nonetheless, even allowing for McReynolds to move towards the center, our evidence provides credence to the claim of the “growing conservatism of . . . Roberts” (Konefsky, 1958, p. 1136).²⁰ Indeed, as we show in Section 6.2, the temporary nature of Roberts’s shift is robust to several different prior specifications. In short, while not conclusive, our best inference is that Roberts shift was sharp, but fleeting.

6 Robustness

In this section, we investigate robustness of the Roberts shift. In Section 6.1, we show that our results are invariant to different assumptions about the breakpoint. Examining all possible breakpoints sheds further light on the exact timing of Roberts’s transformation, showing that it occurred most likely at the beginning of the 1936 term (several cases before *Parrish* was issued) after Roosevelt’s reelection. In Section 6.2, we investigate sensitivity to prior beliefs. As far as we are aware, this paper is the first to point to bridging sensitivity of dynamic ideal point models when the majority of actors is replaced in a short period of time. Notwithstanding such bridging sensitivity, our findings suggest that Roberts’s shift to the left was followed by a period of moving to the right. In Section 6.3, we address concerns of case selection. Lastly, in Section 6.4, we examine the secondary claim often advanced in the literature that Chief Justice Hughes also shifted to the left.

6.1 Breakpoint Analysis

Throughout the paper, we have chosen *Parrish* as the breakpoint of interest, as that case specifically precipitated speculation and scholarship about Roberts’s transformation. Given histor-

²⁰See also Lanier (2003, p. 49) (noting that “Roberts became more conservative” after 1941).

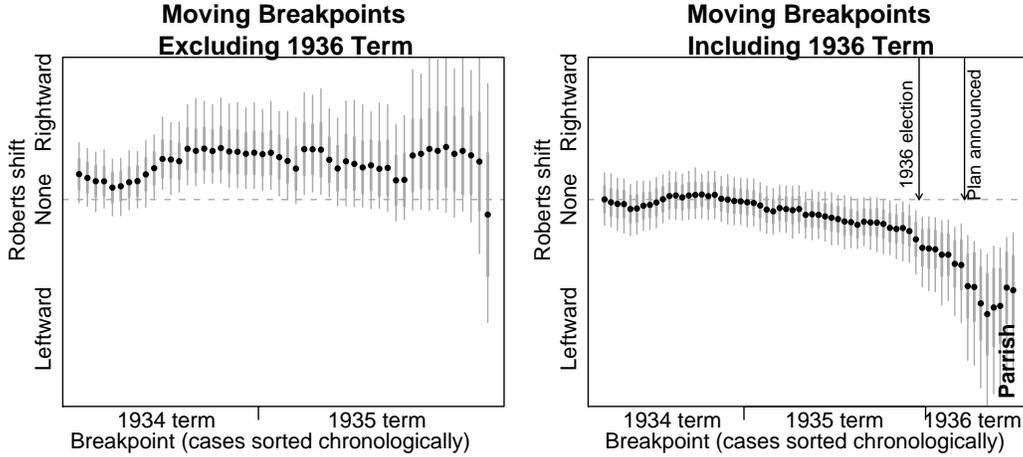


Figure 11: Breakpoint analysis and placebo test. The y -axis presents the estimated Roberts shift (the difference in the pre-breakpoint and post-breakpoint view) and the x -axis uses each case in chronological order as a breakpoint. Each interval represents the shift from a static model fit assuming the breakpoint to occur at any case from the 1934 to 1936 terms. Thick lines represent the median plus or minus one standard deviation and thin lines represent 95% credible intervals. This right panel shows the dramatic shift in the 1936 term. The left panel shows that excluding the 1936 term yields no comparable results.

ical evidence that the conference votes had been taken months before *Parrish* was actually issued, we investigate here the nature of and sensitivity to timing.

To do so, we use each of 61 cases from three years before *Parrish* was decided as the possible breakpoint. Such a sensitivity analysis will allow us to assess Type I error²¹ and when the sharpest drop may have occurred (see Donohue III and Ho, 2007). For example, if Roberts was continuously evolving in a liberal direction, we may falsely infer that Roberts’s shift was a response specifically to the circumstances in 1936. We therefore repeatedly fit the same model of Section 5.1, varying only the case for which Roberts is allowed to shift. We do this once including the 1936 term and once excluding the 1936 term. The latter provides a natural “placebo” test, as few scholars believe Roberts to have shifted prior to the 1936 term. As a discrepancy measure, we report the difference between the post and pre-*Parrish* ideal points for Roberts.

Figure 11 presents results from this breakpoint analysis. The y -axis represents post and pre-breakpoint differences: if positive, the model estimates a rightward shift of Roberts; if around the

²¹From hypothesis testing framework, we mean by Type I error a false rejection of the null hypothesis that Roberts stayed the same pre and post-*Parrish*.

dashed horizontal line at 0, the model estimates no shift; if negative, the model estimates a leftward shift for Roberts. The dots represent estimates (median differences) from each model, with vertical lines representing uncertainty. Estimates are sorted by case breakpoint chronologically on the x -axis. The left panel presents results from the analysis excluding the 1936 term, and the right panel presents the same analysis including the 1936 term with the last breakpoint as the *Parrish* case. The results overwhelmingly confirm that the breakpoint occurred in the 1936 term and that our inferences are not the result of Type I error.

In the left panel almost all estimates are above the origin, and most intervals contain the origin. In other words, excluding the 1936 cases, there's no evidence of any appreciable leftward transformation of Roberts. If anything, the late 1934 cases seem to suggest that Roberts may have become more conservative, which is consistent with the slight upward trend of our dynamic model (e.g., in Figure 10).

The trend in the right panel is drastically different: intervals contain the origin prior to 1936, but in the 1936 term there is a sharp drop in the post-pre difference. Note that the fact that the intervals are already slightly decreasing before the 1936 election does *not* mean that Roberts's transformation begins at that time period, since the post-period includes the 1936 term cases. Contrasting these two panels, the clear evidence is that Roberts's views did not begin to liberalize until the 1936 term and that the shift was sharp.

In addition to confirming that our findings are robust to choice of breakpoint, this analysis sheds considerable insight into the historical debate, much of which involves timing. See, e.g, Cushman 1998, p. 78-83 (positing breakpoint as *Nebbia*); Currie 1990, p. 236-37 (positing breakpoint was *Jones & Laughlin*); White 2000, p. 198-236 (emphasizing breakpoint of 1940s, but discussing general tendency towards shifts in constitutional law); Lanier 2003, p. 216 (finding breakpoint to be Great Depression, not court-packing plan); Ross 2005, p. 1157 (asserting breakpoint was during winter of 1936-37); Corley 2004, p. 51-52 (asserting breakpoint was 1930 based on statistical patterns). Some scholars have argued that Roberts's votes in later cases may be explained by the quality of

the Solicitor-General, better drafting of legislation in the “Second New Deal,” and general learning about constitutional law by the administration. The sharp (but temporary) shift in 1936 must somehow be reconciled with these claims. For example, Stanley Reed replaced James Crawford Biggs as Solicitor General in March 1935, making him counsel on a number of cases before the 1936 term, yet there’s no evidence of a shift for those cases. Part of that may of course be explained by the lingering influence earlier litigation choices. If the administration learned how to “pick off” Roberts during this term (e.g., by using better test cases, argumentation, or drafting), the puzzle becomes why Roberts evolves back to the right after the 1936 term. One might attribute this to recalibration due to Roosevelt appointees. We emphasize that our findings cannot lay to rest the externalist-internalist debate, but our findings of a sharp, temporary shift should help to refine any explanation.

6.2 Sensitivity to Priors

We now analyze how sensitive our results are to prior assumptions. Conceptually, the model parameters in the dynamic model can be grouped into case-specific parameters and justice-specific parameters (the ideal points). While more complicated priors for the case-specific parameters can certainly be constructed, we do not investigate these here. As long as the case parameters are assumed to be independent and identical draws from some distribution, the choice of that distribution primarily affects the scaling of the ideal points.

We focus instead on the priors governing the smoothness of each justice’s series of ideal points. Since these priors affect both the location and the variability of the ideal points, they may directly affect inferences about Roberts’s shift. As detailed in Appendix A.2, these priors assume that each justice’s ideal point series follows a “random walk” in time. The variability across time is governed by a parameter τ for each justice.²² Thus, when τ is close to 0, a justice’s location is essentially constant in time, i.e., equivalent to the estimates presented in Figure 5. As τ gets larger, the

²²Specifically, the increments of the random walk for justice j are mean 0 Gaussian random variables with variance τ_j^2 , assumed to follow an inverse gamma prior distribution with prior parameters.

justice’s location evolves more rapidly from term to term. Indeed, as τ approaches ∞ , the model is effectively the same as presented in Figure 9, with separate ideal points for each justice for each term. Our more general approach presented in Figure 10 strikes a balance between these two extremes, containing each as special cases.

In what follows we present results from this dynamic model under three different assumptions about the smoothness of views, ranging from essentially constant, to moderate variability, to high variability.²³ Each provides a different baseline against which to assess movement. For presentational clarity, we focus on six specifications that typify a wide range of scenarios (to which main results turn out to be robust). Varying these priors thereby allows us to assess sensitivity of inferences about Roberts.

Figure 12 represents results from this analysis. Under all six specifications the shift in behavior for Roberts around the time of the *Parrish* decision remains noticeable. As soon as viewpoints are allowed to evolve (i.e., when τ is moderately greater than 0 as in panel (b)), the sharpest movement occurs with Roberts’s shift in 1936 (with posterior probability of a shift to the left of 0.89). Moreover, under all specifications that allow for movement, we see evidence of a rightward drift by Roberts after 1937. While the structure of the data makes it difficult to untangle individual from compositional effects, our inferences remain robust to a wide range specifications.

6.3 Case Selection

A general objection to our analysis might be that it focuses on nonunanimous cases. The objection has perhaps two variants. One is that we may be ignoring significant implications of unanimously decided cases. Such omitted information, for example, may include opinions by the specific authors of opinions, agreement by Roberts and the liberals on cases with outcomes potentially hostile to the New Deal, or agreement by Roberts and the conservatives with outcomes hospitable to the New Deal. While we agree that much can be learned from such careful parsing of

²³Specifically, we allow the inverse gamma prior to range from a prior mean of each τ_j^2 of 0.00125 and prior variance of 5.2e-7 to a prior mean of τ_j^2 of 0.625 and prior variance of 0.13.

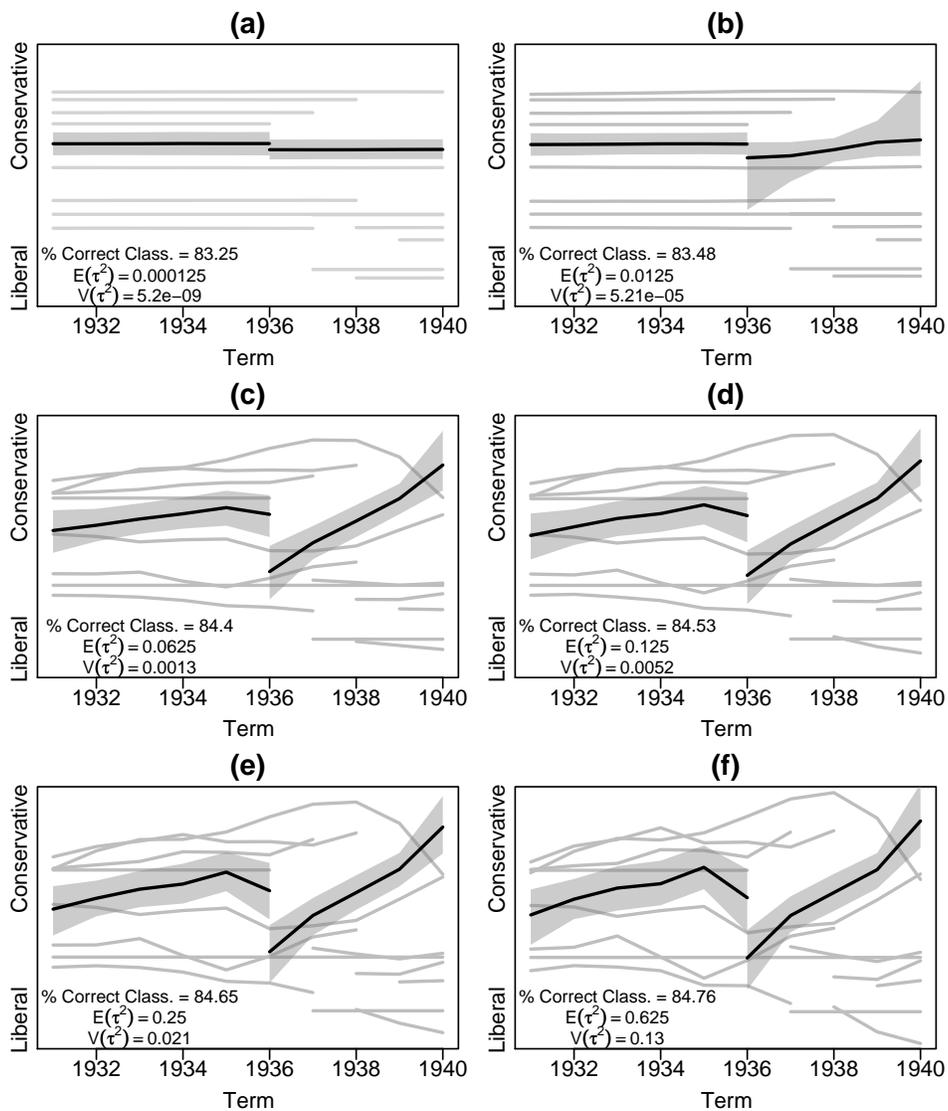


Figure 12: Ideal point estimates for the 1931-1941 period under different prior assumptions for how smoothly views change (τ^2) and different static reference points. In the lower left hand corner of each panel we report the in-sample percentage of correctly classified votes, the prior mean and variance of the smoothness parameter (τ^2). As the panels progress from (a) to (f) the prior mean and variance increase, meaning that views are allowed to change quickly across time. Results assume that Justices Black, Stone, and Van Devanter are constant. The prior specification depicted in panel (c) is that used in Section 5.2. Evidence of the Roberts shift remains even when views are assumed to be almost constant (i.e., when τ^2 is forced by prior assumption to take values very close to 0, equivalent to complete pooling of data across time). Roberts rightward drift after 1937 remains, to varying degrees, under all prior specifications shown in the figure.

opinions, we emphasize that opinions in which all nine justices agreed are likely to shed less insight into the specific evolution of Justice Roberts. Indeed, the model emphasizes this: inclusion would result in no “updating” on these votes.

Another variant is that our analysis is overinclusive, by focusing equally on *all* nonunanimous opinions. Several clarifications are worthwhile here.

First, all nonunanimous cases during this period do not in fact amount to that many cases. As the z -axis label of Figure 1 shows, far fewer nonunanimous cases were decided pre-37 than post-1937. Before the 1937 term, the number of nonunanimous cases range from 22 in the 1934 term to 33 in the 1936 term.

Second, nonunanimous cases include precisely the cases widely discussed in the literature. In the 1936 term alone, for example, our analysis includes the likes of *Parrish*, *NLRB v. Jones & Laughlin Steel Corp.*, 301 U.S. 1 (1937), *Associated Press v. NLRB*, 301 U.S. 103 (1937), *Carmichael v. Southern Coal & Coke Co.*, 301 U.S. 495 (1937), *Steward Machine Co. v. Davis*, 301 U.S. 548 (1937), and *Helvering v. Davis*, 301 U.S. 619 (1937), each widely analyzed in the literature. This shouldn’t be surprising: it wasn’t just the disagreement *between* Roosevelt and the Court, but also *within* the Court itself that elevated the Supreme Court to high politics in 1936-37 (Leuchtenburg, 1999, p. 2103-2108).

Third, our approach does not in fact weight cases equally. To the contrary, our approach formalizes how informative each case is about the differences in viewpoints about the justices, by estimating just how much weight each case should be given. As noted above, a case with unusual voting coalitions will be downweighted by the model and will thus provide less information about the relative locations of the justices. Moreover, the probabilistic nature of the model overcomes the pitfalls of deterministic inference by allowing for randomness of particular cases and votes.

That said, our analysis includes some cases that aren’t necessarily the primary focus of extant scholarship. Clear (exogenous) case selection criteria are, unfortunately, lacking. Extant scholarly disagreements might well then stem more from implicit assumptions or explicit claims about case

selection and weighting, rather than the correct historical understanding of the cases selected (see, e.g., Cushman, 1997; Friedman, 1994b; Pepper, 1998; Ross, 2005; White, 2000, 2005). Should cases like *Humphrey's Executor v. United States*, 295 U.S. 602 (1935), be included (see, e.g., Leuchtenburg, 1999, p. 2088 n.69)? What about foreign affairs and free speech cases (see, e.g., Olken, 2002, p. 285-90, 294-304)? How do we determine what constitute “political economy” decisions (White, 2000, p. 198)? Do cases have to involve New Deal agencies? Legislation? Must they be constitutional decisions? What if the choice between statutory and constitutional interpretation is itself the result of circumstances of interest?

To assess robustness to case inclusion criteria, we examine how five sources treat the 1936 term (Cushman, 1994; Friedman, 1994b; Olken, 2002; Pepper, 1998; Ross, 2005). We collect all citations to 1936 term cases from these sources. Of the 33 cases of the term, only 12 are not cited by any of these five sources. Rerunning the analysis of Section 5.1, all results remain the same, with a posterior p -value that Roberts’s post-*Parrish* ideal point is to the right of the pre-*Parrish* ideal point effectively 0. Our results do not appear to be driven by cases that are irrelevant to the New Deal.

While our results appear robust to case selection, this potential threat to validity highlights one of the major challenges of existing work: without clear articulation of case selection, it remains difficult to know whether inferences may stem solely from case selection (see Epstein and King, 2002). In that sense, our methods highlight a fruitful avenue of research, namely the clearer delineation of case selection, which may help to resolve disagreements.

6.4 Hughes Breakpoint

Since scholars have to varying degrees suggested a switch in Chief Justice Hughes’s behavior around the time of *Parrish* — in addition, of course, to Justice Roberts’s behavior — one might argue that the models that have been employed up to this point in the paper are misspecified in that they do not allow for a dramatic break in Hughes’s ideal point series (see Kalman, 1996, 1999;

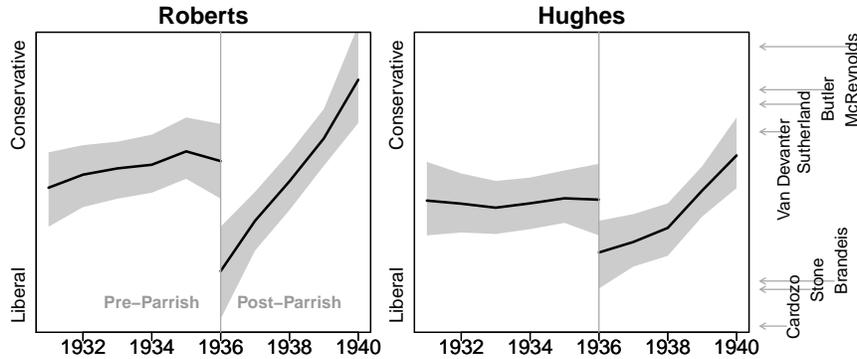


Figure 13: This figure plots dynamic ideal points with separate pre and post-*Parrish* trends for Justice Roberts and Chief Justice Hughes. The solid lines represent the pointwise posterior median ideal points and the grey bands provide pointwise 95% credible regions. Justice Roberts and, to a lesser extent, Chief Justice Hughes experience sharp shifts to the left in the 1936 term. For reference, arrows on the right provide posterior median ideal points of justices other than Roberts and Hughes on the Court in the 1936 term.

Leuchtenburg, 1995; Ackerman, 1998).

To investigate this possibility, we fit another model that is equivalent to the dynamic model used generate Figure 10 in all ways except that Hughes’s votes (in addition to Roberts’s) are split pre and post-*Parrish*. Figure 13 presents the ideal point estimates for Roberts and Hughes under this new specification along with uncertainty (95% credible) bands. There is some evidence for a shift to the left by Hughes, although the shift appears less pronounced than for Roberts. Importantly, the inferences regarding Roberts are effectively unchanged.

7 Conclusion

In this paper, we have provided the first systematic quantitative evidence for the switch – or rather shift – in time. Just as Alf Landon precipitated the scientific study of public opinion, Justice Roberts may be the impetus for incorporating statistical insights into historical scholarship of the Court. We find strong evidence that Roberts shifted sharply and temporarily to the left in the 1936 term.

Our evidence as to the timing also helps to refine historical accounts. For internalists, the explanation as to differences in cases and litigating strategies must correspond to the abrupt tem-

porary shift we identify. Unless the cases in the 1936 term themselves are sharply different, they cannot be reconciled with this evidence. Moreover, such differences must exist outside of the ways in which we have modeled cases. For externalists, our account seems most consistent with the focus on the 1936 landslide election, thereby rebutting naive accounts that Roberts's vote in *Parrish* was a direct result of the court-packing plan.²⁴

Methodologically, we have provided approaches to more systematically study votes and highlighted a key challenge of bridging sensitivity. Paying attention to bridging assumptions remains crucial in cliometric investigations of voting blocs where the length of time under study can be large and the overlap between decision-makers limited.

Our research also clarifies implicit differences in the literature: some argue that Republican gains in 1938 combined with the Court's continued support for New Deal legislation after 1938 contradict that the *Court* was influenced by elections. The argument may be correct as to the Court (but of course the Court changed fundamentally by 1938), but misguided as to Roberts, who did in fact trend back to the right after the 1938 elections. More generally, we hope our research highlights productive synergies between quantitative and qualitative research. Focusing on key identification assumptions can help crystallize questions for further qualitative research. Similar to matching methods, which can clarify assumptions and pinpoint observations that merit further study (Ho et al., 2007), measurement methods point scholars to the key cases and assumptions of interest (Ho and Quinn, 2009a). Like Gallup to polls, cliometric applications have broad potential to reinvigorate longstanding questions of constitutional history.

²⁴One might also speculate whether an implicit agreement between Justice Roberts and the liberal justices might explain the shift. Two key challenges to such an explanation are (a) why it would have rational for Roberts to engage in such an agreement, and (b) how the colleagues might have enforced such an agreement, given perceptions of impending changes to the Court. For related evidence of potentially strategic doctrinal collusion, and ambiguity of its rationale, see Ho and Ross (2010).

A Statistical Methodology

The statistical measurement models that are used in this paper are two-parameter item response theory (IRT) models with probit link. Such models have a long history in psychometrics and educational testing as well as in political science. In psychometric and educational applications the parameters that we have referred to as “ideal points” above are thought of as latent measures of intelligence or ability and the goal is to learn how responses to a battery of test questions relate to intelligence or ability. In political science, the data are typically votes (often legislative roll calls) and the goal is to infer the latent ideology of legislators from these votes. For instance, Clinton et al. (2004) applies an IRT model to congressional roll call data and show how it can be derived from a model of spatial voting and how it related to earlier estimation strategies (Heckman and Snyder, 1997; Poole and Rosenthal, 1985, 1991, 1997). The model used in this paper is the extension of the Clinton et al. (2004) model proposed by Martin and Quinn (2002) for Supreme Court merits votes (with some additional modifications) (see also Ho and Quinn, 2009b; Ho and Ross, 2010).

In what follows, we briefly sketch the measurement models used in this paper. Readers who are interested in a more detailed treatment of these models should see Clinton et al. (2004), Martin and Quinn (2002) and references therein.

A.1 Static Ideal Point Model

The simplest model we employ is what we term the “static ideal point model.” This model takes the following form.

To clarify notation, let J denote the set of justices under study and K the set of cases. The observed data \mathbf{Y} have typical element

$$y_{jk} = \begin{cases} 0 & \text{if } j \text{ is in the minority on case } k \\ 1 & \text{if } j \text{ is in the majority on case } k \\ \text{missing} & \text{if } j \text{ did not vote on case } k \end{cases}$$

It will be convenient to define the set $J_k = \{j \in J : y_{jk} \neq \text{missing}\}$ for all $k \in K$. In words, J_k is just the set of justices who took part in the decision on case k .

A.1.1 The Likelihood

The starting point for the model is the assumption that

$$y_{jk} \stackrel{\text{ind.}}{\sim} \text{Bernoulli}(\pi_{jk}) \quad j \in J, \quad k \in K$$

where the probability π_{jk} of a majority vote is parameterized as

$$\pi_{jk} = \Phi(-\alpha_k + \beta_k \theta_j)$$

and $\Phi(\cdot)$ is the standard normal distribution function.

The key parameters of this model are α , β , and θ . The elements of α and β are specific to cases and thus capture case characteristics: α collects what we refer to as the dissent parameter and β collects what we refer to as the valence parameter in Section 4. The elements of θ are unique to each justice and thus capture justice-specific attributes that are predictive of voting behavior. It is common to think of θ_j as representing the most preferred policy location of a decision maker along a one dimensional continuum (Clinton et al., 2004; Martin and Quinn, 2002). This is a convenient interpretation in the present context although to the extent that we are only interested in documenting and summarizing changes in behavior rather than inferring the reasons for such behavior we only need to interpret θ_j as the best one-dimensional summary of j ’s voting behavior.

Simple algebra reveals that the value of θ_j that would make it equally likely for justice j to join the majority or minority bloc on case k is $\kappa_k = \alpha_k/\beta_k$. If $\theta_j > \kappa_k$ and $\beta_k > 0$ then j has a better than 50% chance of joining the majority. If $\theta_j > \kappa_k$ and $\beta_k < 0$ then j has a better than 50% chance of joining the minority. Note that the sign of β_k determines the valence of values in θ space for a particular case. This is important in that we do not have to make any assumptions about what a “liberal” or “conservative” vote is on a particular case. Such determinations are made by the data in conjunction with some identifying assumptions.

Given the information above it follows that the sampling density for this model is

$$p(\mathbf{Y}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}) \propto \prod_{k \in K} \prod_{j \in J_k} \Phi(-\alpha_k + \beta_k \theta_j)^{y_{jk}} [1 - \Phi(-\alpha_k + \beta_k \theta_j)]^{(1-y_{jk})}$$

A.1.2 The Prior

We adopt a Bayesian approach to inference (see generally Strnad, 2007), and thus we must specify a prior distribution for $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta})$. As is common in the literature, we assume that the elements of $\boldsymbol{\theta}$ are a priori mutually independent of each other and each α_k and β_k . Our prior for $\boldsymbol{\theta}$ is that

$$\theta_j \stackrel{iid}{\sim} \mathcal{N}(0, 1) \quad \text{for all } j \in J$$

Our prior for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is that each (α_k, β_k) is independently drawn from a uniform distribution on the region $\{\alpha_k, \beta_k : \alpha_k \in [-4, 4], \beta_k \in [-2, 2], \alpha_k/\beta_k \in [-2, 2]\}$

This implies that *marginally* each α_k follows a triangle distribution with density

$$p(\alpha_k) = \begin{cases} \frac{1}{4} + \frac{\alpha_k}{16} & \text{if } \alpha_k \in [-4, 0) \\ \frac{1}{4} - \frac{\alpha_k}{16} & \text{if } \alpha_k \in (0, 4] \\ 0 & \text{otherwise} \end{cases}$$

and each β_k *marginally* follows the distribution with density

$$p(\beta_k) = \begin{cases} -\frac{\beta_k}{4} & \text{if } \beta_k \in [-2, 0) \\ \frac{\beta_k}{4} & \text{if } \beta_k \in (0, 2] \\ 0 & \text{otherwise} \end{cases}$$

It is fairly straightforward to show that this joint prior for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ induces a uniform prior from -2 to 2 on each of the cutting lines in $\boldsymbol{\kappa}$. This prior is thus a reasonable operationalization of prior ignorance about the case parameters.

A.1.3 Identification and Model Fitting

As written out above, the model is not identified because of the ability to change the sign of each element of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ and achieve the same posterior density value. Numerous possibilities exist to eliminate this invariance. Here we choose to constrain one justice, Butler, to have a positive θ value and one justice, Brandeis, to have a negative θ value.

We fit the model via Markov chain Monte Carlo (MCMC). The algorithm is essentially identical to that in Clinton et al. (2004) with the exception of the sampling of (α_k, β_k) which has to be slightly modified to account for the non-standard prior on $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. With 70,000 draws from the joint posterior, standard diagnostics suggest convergence.

A.2 Dynamic Ideal Point Model

The static ideal point model provides a parsimonious summary of voting behavior on the Court in that each case is represented by two parameters and every justice is represented by a single scalar parameter. While the the static model provides a reasonable representation of behavior in a narrow time window, over a longer time span it may be desirable to allow the ideal points of the justices to evolve over time. In these situations, we use a generalization of the model, which we call the “dynamic ideal point model.” This model is very similar to the static model except that now each justice’s ideal point is allowed to vary from term to term.

Specifically, we parameterize the probability π_{jk} of a majority vote as

$$\pi_{jk} = \Phi(-\alpha_k + \beta_k \theta_{jt})$$

where t represents the term in which case k was decided. With π_{jk} reparameterized, the likelihood for the dynamic models is essentially equivalent to the likelihood for the static model. The difference arises in the prior for θ .

A.2.1 The Prior

Here we assume that a priori θ_{jt} follows a Gaussian random walk in time. Formally

$$\theta_{jt} = \theta_{j(t-1)} + \epsilon_{jt}$$

We further assume that

$$\begin{aligned} \epsilon_{jt} &\stackrel{ind}{\sim} \mathcal{N}(0, \tau_j^2) \\ \tau_j^2 &\stackrel{ind}{\sim} \mathcal{IG}(c_j/2, d_j/2), \end{aligned}$$

where c_j and d_j determine the prior mean and variance of the inverse gamma distribution governing τ_j .

We let θ_{j0} denote j ’s ideal point in the term immediately before his service began. We assume

$$\theta_{j0} \sim \mathcal{N}(0, 1).$$

The special case of $\tau_j = 0 \forall j$ is the static ideal point model. When τ_j approaches ∞ , ideal points for justices are independent for each term, and the model is equivalent to the term-by-term estimates presented in Figure 9 (see generally Gelman and Hill, 2007). The priors for α and β are as described for the static ideal point model.

A.2.2 Identification and Model Fitting

This model has identification problems that are very similar to those of the static ideal point model. In the current data set where there is a dramatic replacement of justices in the late 1930s the problems are severe because of a lack of overlap in the voting records of the justices. We refer to this identification challenge as *bridging sensitivity*. Throughout the paper we fix the polarity of the model by assuming that Justice McReynolds’s ideal point is to the right of 0 and Justice Douglas’s ideal point is to the left of 0. In addition, in order to get a baseline against which to measure relative change in voting behavior we assume that the ideal points of Justices Black, Van Devanter, and Stone are constant over time.

Model fitting is accomplished via MCMC. The algorithm is essentially the same as that in Martin and Quinn (2002) with the exception of a slight modification to accommodate the non-standard prior for α and β . With 70,000 draws from the joint posterior, standard diagnostics suggest convergence.

References

- Ackerman, Bruce. 1998. *We the People: Transformations*. Cambridge, MA: Harvard.
- Ariens, Michael. 1994. "A Thrice-Told Tale, or Felix the Cat." *Harvard Law Review* 107(3):620–676.
- Bailey, Michael A. 2007. "Comparable Preference Estimates across Time and Institutions for the Court, Congress and Presidency." *American Journal of Political Science* 51(3).
- Brinkley, Alan. 2005. "The Debate over the Constitutional Revolution of 1937." *American Historical Review* 110:1046–1050.
- Caldeira, Gregory A. 1987. "Public Opinion and The U.S. Supreme Court: FDR's Court-Packing Plan." *The American Political Science Review* 81(4):1139–1153.
- Carson, Jamie L., and Benjamin A. Kleinerman. 2002. "A Switch in Time Saves Nine: Institutions, Strategic Actors, and FDR's Court-Packing Plan." *Public Choice* 113:301–324.
- Chambers, John W. 1969. "The Big Switch: Justice Roberts and the Minimum-Wage Cases." *Labor History* 10:44–73.
- Clinton, Joshua D., Simon Jackman, and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98:355–370.
- Corley, Roger W. 2004. "Was there a Constitutional Revolution in 1937?" In *Franklin D. Roosevelt and the Transformation of the Supreme Court* (Stephen K. Shaw, William D. Pederson, and Frank J. Williams, editors), Armonk, NY: M.E. Sharpe.
- Currie, David P. 1990. *The Constitution in the Supreme Court: The Second Century, 1888-1986*. Chicago, IL: University of Chicago Press.
- Cushman, Barry. 1994. "Rethinking the New Deal Court." *Virginia Law Review* 80(1):201–261.
- Cushman, Barry. 1997. "The Secret Lives of the Four Horsemen." *Virginia Law Review* 83(3):559–645.
- Cushman, Barry. 1998. *Rethinking the New Deal Court*. New York, NY: Oxford University Press.
- Donohue III, John J., and Daniel E. Ho. 2007. "The Impact of Damage Caps on Malpractice Claims: Randomization Inference with Difference-in-Differences." *Journal of Empirical Legal Studies* 4:69–102.
- Epstein, Lee, and Gary King. 2002. "The Rules of Inference." *The University of Chicago Law Review* 69(1):1–133.
- Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, and Organization* 23(2):303–325.
- Farber, Daniel A. 2002. "Who Killed Lochner?" *Georgetown Law Journal* 90:985–1007.
- Fogel, Robert William. 1975. "The Limits of Quantitative Methods in History." *The American Historical Review* 80(2):329–350.
- Frankfurter, Felix. 1955. "Mr. Justice Roberts." *University of Pennsylvania Law Review* 104(3):311–317.
- Friedman, Barry. 2000. "The History of the Countermajoritarian Difficulty, Part Four: Law's Politics." *University of Pennsylvania Law Review* 148(4):971–1064.

- Friedman, Richard D. 1994a. "A Reaffirmation: The Authenticity of the Roberts Memorandum, or Felix the Non-Forgers." *University of Pennsylvania Law Review* 142(6):1985–1995.
- Friedman, Richard D. 1994b. "Switching Time and Other thought Experiments: The Hughes Court and Constitutional Transformation." *University of Pennsylvania Law Review* 142(6):1891–1984.
- Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. New York, NY: Cambridge.
- Gely, Rafael, and Pablo T. Spiller. 1992. "The Political Economy of Supreme Court Constitutional Decisions: The Case of Roosevelt's Court-Packing Plan." *International Review of Law and Economics* 12:45–67.
- Griffin, Stephen M. 1999. "Constitutional Theory Transformed." *The Yale Law Journal* 108(8):2115–2163.
- Harrison, Robert. 1984. "The Breakup of the Supreme Court: The Contribution of History and Biography." *Law and History Review* 2(2):165–221.
- Heckman, James J., and James M. Snyder. 1997. "Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators." *RAND Journal of Economics* 28:S142–S189.
- Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15:199–236.
- Ho, Daniel E., and Kevin M. Quinn. 2009a. "How Not To Lie with Judicial Votes: Misconceptions, Measurement, and Models." *California Law Review (forthcoming)* 98.
- Ho, Daniel E., and Kevin M. Quinn. 2009b. "Viewpoint Diversity and Media Consolidation: An Empirical Study." *Stanford Law Review* 61:781–868.
- Ho, Daniel E., and Erica L. Ross. 2010. "Did Liberal Justices Invent the Standing Doctrine? An Empirical Study of the Evolution of Standing, 1921-2006." *Stanford Law Review (forthcoming)* 62.
- Johnson, Valen E., and James H. Albert. 1999. *Ordinal Data Modeling*. New York: Springer.
- Kalman, Laura. 1996. *The Strange Career of Legal Liberalism*. New Haven, CT: Yale.
- Kalman, Laura. 1999. "Law, Politics, and the New Deal(s)." *The Yale Law Journal* 108(8):2165–2213.
- Kalman, Laura. 2005. "The Constitution, the Supreme Court, and the New Deal." *American Historical Review* 110:1052–1079.
- Konefsky, Samuel J. 1958. "Review of Harlan Fiske Stone: Pillar of the Law." *The American Political Science Review* 52(4):1135–1137.
- Lanier, Drew Noble. 2003. *Of Time and Judicial Behavior: United States Supreme Court Agenda-setting and Decision-making, 1888-1997*. Selinsgrove, PA: Susquehanna University Press.
- Lawrence, Albert. 2005. "Biased Justice: James C. McReynolds of the Supreme Court of the United States." *Journal of Supreme Court History* 30:244–270.

- Lessig, Lawrence. 1995. "Understanding Changed Readings: Fidelity and Theory." *Stanford Law Review* 47(3):395–472.
- Leuchtenburg, William E. 1995. *The Supreme Court Reborn: The Constitutional Revolution in the Age of Roosevelt*. New York, NY: Oxford.
- Leuchtenburg, William E. 1999. "When the People Spoke, What Did They Say?: The Election of 1936 and the Ackerman Thesis." *Yale Law Journal* 108:2077–2114.
- Leuchtenburg, William E. 2005. "Comment on Laura Kalman's Article." *American Historical Review* 110:1081–1092.
- Martin, Andrew D., and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10:134–153.
- McCloskey, Robert G. 2000. *The American Supreme Court*. Chicago, IL: Chicago University Press.
- McKenna, Marian C. 2002. *Franklin Roosevelt and the Great Constitutional War*. New York, NY: Fordham.
- Morrison, Rodney J. 1977. "Franklin D. Roosevelt and the Supreme Court: An Example of the Use of Probability Theory in Political History." *History and Theory* 16(2):137–146.
- Nelson, Michael. 1988. "The President and the Court: Reinterpreting the Court-packing Episode of 1937." *Political Science Quarterly* 103(2):267–293.
- Olken, Samuel R. 2002. "Historical Revisionism and Constitutional Change: Understanding the New Deal Court." *Virginia Law Review* 88(1):265–326.
- Parrish, Michael E. 1982. *Felix Frankfurter and His Times*. New York: Collier Macmillan.
- Pepper, David A. 1998. "Against Legalism: Rebutting an Anachronistic Account of 1937." *Marquette Law Review* 82:63–154.
- Poole, Keith T., and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29:357–384.
- Poole, Keith T., and Howard Rosenthal. 1991. "Patterns of Congressional Voting." *American Journal of Political Science* 35(February):228–278.
- Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York, NY: Oxford University Press.
- Pritchett, C. Herman. 1948. *The Roosevelt Court*. New York, NY: The Macmillan Co.
- Purcell, Edward A. 1994. "Rethinking Constitutional Change." *Virginia Law Review* 80(1):277–290.
- Ross, William G. 2005. "When Did the "Switch in Time" Actually Occur?: Re-discovering the Supreme Court's "Forgotten" Decisions of 1936-1937." *Arizona State Law Journal* 37:1153–1220.
- Schubert, Glendon A. 1959. *Quantitative Analysis of Judicial Behavior*. Glencoe, IL: Free Press.
- Schwartz, Bernard. 1995. *A History of the Supreme Court*. New York, NY: Oxford.
- Strnad, Jeff. 2007. "Should Legal Empiricists Go Bayesian?" *American Law and Economics Review* 9:195–303.

White, G. Edward. 1996. "Cabining the Constitutional History of the New Deal in Time." *Michigan Law Review* 94(6):1392–1421.

White, G. Edward. 2000. *The Constitution and the New Deal*. Cambridge, MA: Harvard.

White, G. Edward. 2005. "Constitutional Change and the New Deal: The Internalist/Externalist Debate." *American Historical Review* 110:1094–1115.