JOURNAL OF
EMPIRICAL LEGAL STUDIES

*Journal of Empirical Legal Studies*
Volume 6, Issue 3, 485–511, September 2009

# How Will They Vote? Predicting the Future Behavior of Supreme Court Nominees, 1937–2006

*Charles M. Cameron and Jee-Kwang Park**

Previous research suggests that the future behavior of nominees to the U.S. Supreme Court is relatively unpredictable, except for civil liberties cases. We devise a new measure of nominees' political ideology that more efficiently uses preconfirmation information about the nominees. The measure employs Segal-Cover scores (based on content analysis of contemporary newspaper editorials) as well as DW-NOMINATE indicators, and is scaled into the DW-NOMINATE space. The measure predicts confirmed nominees' overall immediate, short-term, and longer-term voting behavior, as well as voting in issue-specific domains, much better than do previous measures. It is particularly successful for nominees confirmed after 1957.

## I. Introduction

Is it possible to predict the probable future behavior of nominees to the U.S. Supreme Court? Commonsense and casual observation suggest a nominee's political ideology is usually a good indicator of his or her future behavior as a justice. Indeed, presidents, senators, interest groups, and the media all seem to employ this commonsense judgment, since perceived nominee ideology is so central to confirmation politics (Epstein & Segal 2005; Cameron et al. 1990).

Political science scholarship offers a substantial caveat, however. Considerable effort and ingenuity have been devoted to measuring

*Address correspondence to Charles Cameron, Robertson Hall, Princeton University, Princeton, NJ 08540; email: ccameron@princeton.edu. Cameron is Professor of Politics and Public Affairs, Princeton University, and Visiting Professor of Law, New York University School of Law; Park is Post-Doctoral Research Associate at the Center for the Study of Democratic Politics, Woodrow Wilson School, Princeton University.

preconfirmation perceptions of nominee ideology. The most widely employed measure, the Segal-Cover scores, are derived from content analysis of newspaper editorials at the time of the nomination, essentially measures the percentage of statements indicating whether the nominee is liberal, conservative, or moderate (Segal & Cover 1989; Segal et al. 1995). The Segal-Cover scores pass normal standards for content analytic measures. Unfortunately, statistical studies suggest that the Segal-Cover scores are a relatively poor predictor of future voting behavior in issue areas other than civil liberties (Epstein & Mershon 1996; Martin & Quinn 2005; but see Epstein & Segal 2005:124–27). In fact, Epstein and Mershon conclude: "Students of the judicial process who seek to explore phenomena other than aggregated individual-level voting in civil liberties cases ought to give serious thought to devising new surrogates for judicial preferences" (1996:261).

In this article, we devise a new measure of nominees' political ideology that more efficiently incorporates preconfirmation information about the nominees. The measure employs the Segal-Cover scores, as well as what we call the "best available NOMINATE score" for each nominee, for example, his or her DW-NOMINATE score as a congressperson. We view both the newspaper score and the best available NOMINATE score as noisy measures of the likely ideology of a nominee—neither indicator is perfect, but each supplies worthwhile information. Accordingly, we use principal component analysis to extract a common factor for each nominee. Because one of the indicators is the best available NOMINATE score, it is straightforward to project the factor back into the DW-NOMINATE scale (we employ the Senate space). This yields a NOMINATE-scaled perception (NSP) score for each nominee. We calculate NSP scores for all nominees between 1937 and 2006.

We then examine the ability of the NSP score to predict justices' short- and longer-term overall voting behavior and their voting behavior in domain-specific issue areas. To measure overall voting behavior, we employ as a dependent variable the widely used Martin-Quinn (MQ) scores from 1937–2005 (Martin & Quinn 2002). More specifically, we employ the justice's first-year MQ score, first-five-year average MQ score, first-10-year average MQ score, and life-time average MQ score. To examine issue-specific voting, we employ three different measures. Two are based on the "liberal outcome" code in the U.S. Supreme Court Judicial Database (Spaeth 2006). The first examines "percent liberal" voting in six subareas, replicating and updating Epstein and Mershon's 1996 analysis, for 1953–2005. The second employs "percent liberal" voting scores in broader social and economic cases, as

derived by Segal et al. (2000); we update this measure so it is available from 1937–2005. Because the "liberal outcome" coding has been criticized as somewhat arbitrary (Harvey 2006), we supplement the two previous measures with new item-response theoretic scores for the six policy-specific domains for 1953–2005. These are based on coalitions in nonunanimous votes, rather than on Spaeth's "liberal outcome" coding.

We find that the NSP score affords substantially improved predictions of justices' immediate, short-term, and long-term overall voting. The NSP score outperforms Segal-Cover scores, the nominating president's DW-NOMINATE score, or the two of these used in tandem. In fact, NSP scores predict 10-year average voting behavior only slightly less well than Segal-Cover scores predict first-year voting behavior. In the important area of civil liberties cases, the NSP score continues to perform as well as Segal-Cover scores. However, NSP scores provide consistently improved predictions for justices' voting behavior on most non-civil-liberties issues (neither they nor the other measures do well in predicting future behavior in tax cases).

We also uncover an important change in the predictability of nominees' future behavior: the future behavior of nominees confirmed after about 1957 became much more predictable. For example, the $R^2$ of a simple regression of NSP score on first-five-year average Martin-Quinn score in the post-1958 period is 0.77. The predictability of pre-1958 nominees was much lower. This change suggests that modern presidents may have changed the way they select nominees in order to advance individuals with more predictable and reliable behavior to the high court.

We conclude that predicting the future behavior of nominees to the U.S. Supreme Court is less problematic than had been supposed, especially in the modern era. Consequently, the centrality of perceived nominee ideology in Supreme Court confirmation politics appears entirely understandable.

The article is organized as follows. In the following section, we derive the NSP score. We briefly discuss the resulting scores, which display considerable facial validity. The next two sections examine the ability of the NSP score and other measures to predict justices' future voting behavior. Section III considers the justices' overall voting tendencies in the short term and longer terms. We also examine changes in the predictability of nominees over time. Section IV provides a similar analysis, but focuses on voting in issue-specific domains. Section V provides out-of-sample predictions of the future voting behavior for Justices Roberts and Alito. Section VI discusses the results and concludes. Tables include the NSP score for each nominee from 1937 to 2006.

## II. Deriving NOMINATE-Scaled Perception Scores

A Supreme Court nominee presents himself or herself trailing markers of political allegiance. Depending on his or her career, these indicators may include opinions authored as a judge, votes cast as a member of Congress, actions taken as an executive branch official, scholarly articles written as an academic, and public speeches and private memos. Presidential aides, interest groups, media pundits, and academics read these tea leaves carefully, trying to discern a nominee's true ideological proclivities. Unfortunately, it is difficult, if not impossible, to subject all these sources to a systematic quantitative analysis. Nonetheless, it is possible to draw on them in a systematic way.

First, using content analysis of newspaper editorials, it is possible to score media perceptions of nominees (Segal & Cover 1989; Segal et al. 1995; Epstein & Mershon 1996). As discussed in the literature, these media-content scores meet conventional standards of reliability for content analysis. Appendix A indicates newspaper content scores for every nominee from Hugo Black (1937) to Samuel Alito (2006). These scores are on the 0–1 scale suggested by Epstein and Mershon, where 1 denotes a nominee seen by editorial writers as extremely liberal and 0 denotes a nominee seen as extremely conservative.[1] Using the same methodology employed in published studies, Segal recently calculated scores for John Roberts, Harriet Miers, and Samuel Alito; we have calculated scores for Douglas Ginsberg and Homer Thornberry, who were omitted in previous published work.

Although the media-content scores have proven useful to scholars, they are hardly perfect. For example, editorial coverage of some nominees was slight, suggesting that their media-content scores may be quite noisy.[2] Moreover, for studies of confirmation politics, it would be extremely useful to have a DW-NOMINATE-based indicator for nominees, as DW-NOMINATE scores are available for senators and presidents.

To address the latter problem, Epstein et al. (2006) proceed as follows. If the president's party controlled the Senate, they assume the nominee had

---

[1]Segal and Cover's original scores are scaled from −1 to 1. The Epstein-Mershon scale is simply (1 + Segal-Cover score)/2.

[2]Szmer and Songer (2005) show that media scores based on fewer editorials predict subsequent behavior of justices less well than scores based on more editorials.

on average the same ideology as the president. So, for these nominees, they regress the nominees' Segal-Cover scores (on a 0–1 scale) onto the president's DW-NOMINATE score and apply the estimated linear transformation to all the Segal-Cover scores (on the 0–1 scale). This modified Segal-Cover (MSC) Score is shown in Column 2 of the Appendix.

We employ a different approach that attempts to efficiently employ the DW-NOMINATE-based "tracks" left by nominees in the 1937–2006 period. In particular, the following measures are available.

- Five nominees served in Congress and thus have DW-NOMINATE scores as legislators.[3]
- Three nominees for Chief Justice (Stone, Fortas, and Rehnquist) served on the Supreme Court and thus have "common space" Martin-Quinn behavioral voting scores, very comparable to DW-NOMINATE scores (Epstein et al. 2007a).
- Nineteen nominees served in the U.S. Courts of Appeal and thus have Giles-Hettinger-Pepper inferential scores based on presidential and senatorial DW-NOMINATE scores (Giles et al. 2001).[4]
- All nominees were nominated by a U.S. president, whose ideology has been measured in DW-NOMINATE space based on his announced positions on roll-call votes (McCarty & Poole 1995).[5]

Clearly, each measure is less than perfect. Some legislators' voting behavior may deviate systematically from their personal ideology; Martin-Quinn behavioral voting scores probably reflect influences other than the justices' personal ideology; inferential measures for appellate judges are at best noisy proxies for their true ideology; and presidents occasionally nominate individuals who stand at some distance ideologically. For example, Roosevelt nominated Republican Justice Harlan Fiske Stone as Chief Justice, Truman nominated Republican Senator Harold Burton, and Eisenhower nominated

---

[3]Black, Byrnes, Burton, Vinson, and Minton.

[4]Harlan, Whittaker, Stewart, Marshall, Thornberry, Burger, Haynsworth, Carswell, Blackmun, Stevens, Scalia, Bork, D. Ginsburg, Kennedy, Souter, Thomas, R.B. Ginsburg, Breyer, Roberts, and Alito. However, Roberts's score is not coded in the original data. We used George W. Bush's DW-NOMINATE score instead.

[5]We employ McCarty's version of the presidential NOMINATE scores but little changes if one employs the version available from Poole at ⟨www.voteview.com⟩. We thank Nolan McCarty for sharing these scores.

Democrat Judge William Brennan.[6] Nonetheless, legislator DW-NOMINATE scores are usually excellent indicators of legislator ideology, justices tend to vote in accord with their underlying political proclivities, inferential DW-NOMINATE scores for judges of the U.S. Courts of Appeal are good predictors of their subsequent behavior in those courts, and presidents tend to pick Supreme Court nominees who reflect their political values. Accordingly, we utilize these five indicators to create a "best available NOMINATE score" at the time of nomination for each nominee from 1937–2006. The best available measure is a direct measure of behavior if available (i.e., a legislative DW-NOMINATE score or a Martin-Quinn judicial common space score) and an inferential measure if a direct behavior measure is not available (i.e., a Giles-Hettinger-Pepper score if available and a presidential DW-NOMINATE score if not). In each case, we translate the best available measure into the Senate DW-NOMINATE space.[7] The best available DW-NOMINATE score in the Senate space for each nominee from 1937 to 2006 is shown in the Appendix.

If we view both the media score and the best available DW-NOMINATE score as observable but noisy indicators that tap into an underlying unobserved latent variable (ideology), a natural procedure is to factor analyze the two observable indicators and recover the underlying latent variable. The fourth column of the Appendix indicates the value of the first principal component for each nominee from a factor analysis of the two indicators. The first principal component accounts for 85 percent of the variation in the two indicators.[8]

Although the first principal component provides an arguably plausible measure of each nominee's political ideology, we wish to scale the measure into DW-NOMINATE space. Because one of the indicators used to construct the measure is the best available DW-NOMINATE score, projecting the factor into (Senate) DW-NOMINATE space is straightforward. This may be done using the loadings and component scores from the factor analysis or, equivalently, the coefficients from a simple OLS regression of the factor on

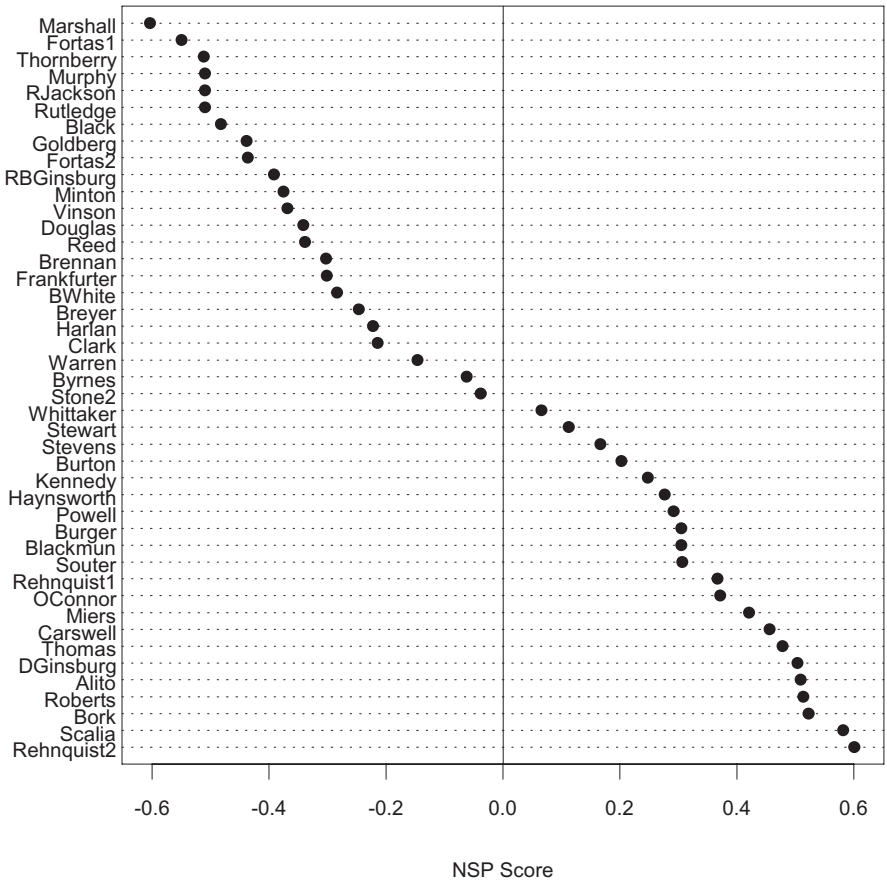
---

[6]The nominations of Stone and Burton both occurred during unified party government.

[7]We thank Nolan McCarty for sharing conversion factors for the House, common space, and Senate DW-NOMINATE scores. The conversion factors are based on comparisons of scores for individuals who served in multiple chambers.

[8]Additional descriptive statistics for the factor analysis are available from the authors on request.

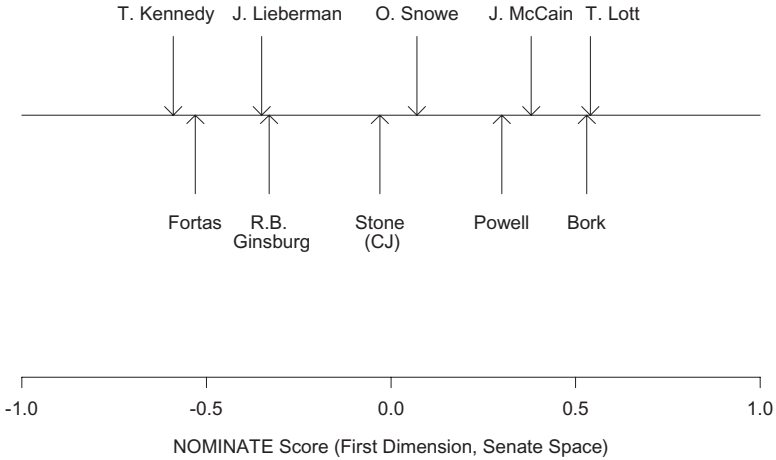
*Figure 1:* NOMINATE-scaled perceptions scores. Nominees with large negative scores were perceived as liberals (e.g., Fortas, Marshall, Murphy, and Jackson), those with high positive scores were perceived as conservatives (Rehnquist, Scalia, Douglas Ginsburg, Bork, Roberts, Alito, Thomas, and Carswell).




the best available DW-NOMINATE score. The resulting NOMINATE-scaled perceptions (NSP) score for each nominee (in the Senate space) is shown in Column 5 of the Appendix. The NSP score may be thought of as the best available DW-NOMINATE score for each nominee, adjusted for the contemporary perceptions of the nominee's ideology.

Figure 1 displays each nominee's NSP score. The "liberal" end of the scale contains the scores of Fortas, Marshall, Thornberry, Murphy, Rutledge,

*Figure 2:* Illustrative nominees and senators in the DW-NOMINATE scale (Senate space). Shown are the DW-NOMINATE scores for some well-known senators in the 109th Congress and the NSP scores (Senate space) of some notable Supreme Court nominees.




Black, and Goldberg. The scores of Rehnquist, Scalia, Douglas Ginsburg, Bork, Roberts, Alito, and Thomas lie at the conservative end of the scale.

To provide a helpful comparison, Figure 2 displays the NSP scores of some notable nominees relative to the DW-NOMINATE scores of well-known senators in the 109th Congress. As shown, the score of arch-liberal Abe Fortas appears somewhat comparable to that of Senator Edward Kennedy (D-MA), while that of moderate liberal Ruth Bader Ginsburg appears comparable to that of maverick former Democrat Senator Joseph Lieberman (I-CT).The score of liberal Republican Harlan Fiske Stone appears similar to that of Senator Olympia Snowe (R-ME), while moderate conservative Lewis Powell has a score similar to that of Senator John McCain (R-AZ). The scores of arch-conservatives Robert Bork and Senator Trent Lott (R-MS) suggest that the two are ideological soul mates.

NSP scores and modified Segal-Cover scores are often quite similar (see the Appendix). In fact, the mean difference between the two scores (0.01) is small and the two measures are highly correlated. Nonetheless, the two measures often differ, sometimes considerably.

As an illustrative example of differences in the scores, consider James F. Byrnes, an FDR appointee. The contemporary editorialists portrayed the

South Carolina Democrat as a conservative-leaning moderate, giving him an MSC score of 0.15. However, his NSP score of −0.08 indicates a liberal-leaning moderate. The discrepancy arises because Byrnes's NSP score partially reflects his rather liberal voting record in Congress (his (Senate-scaled) DW-NOMINATE score was −0.32). In fact, Byrnes's first-year voting score on the Supreme Court was −0.18, confirming his position as a liberal-leaning moderate.

Sandra Day O'Connor also receives rather different scores under the two measures. The editorialists scored O'Connor at 0.07, suggesting a liberal Republican in the mold of Maine Senator Olympia Snowe. This score probably reflected hints that O'Connor was an abortion moderate during her service in the Arizona state legislature. However, her NSP score of 0.39 suggests instead a moderate conservative, rather similar to her fellow Arizona Republican, Senator John McCain. This score partially reflects the fact that O'Connor was nominated by stalwart conservative Ronald Reagan. In her first year on the Court, O'Connor's voting score was 0.29 (in the Senate scale), confirming her position as a moderate conservative.

## III. Predicting Overall Voting Tendency

If the fundamental tenant of much judicial politics research is correct—justices tend to vote in accord with their values—then an accurate perceptions measure ought successfully to predict Supreme Court justices' subsequent voting behavior. Indeed, the ability to predict subsequent voting behavior is something of an acid test for preconfirmation measures of nominee ideology. After all, if a perceptions measure predicts subsequent behavior only poorly, how could nomination politics hinge critically on the values indicated by the measure?

Several caveats are in order. No perceptions measure will predict later voting behavior perfectly, especially over the longer term. In the first case, perceptions of nominee ideology can sometimes be mistaken. Justices John Paul Stevens and David Souter are often cited as "mavericks" whose subsequent behavior simply confounded initial expectations. Second, some justices show substantial changes in voting behavior as the Court's docket evolves. For example, FDR appointee Stanley Reed was a reliable supporter of a strong, federal administrative state, and thus scores as a "liberal" in the late 1930s and prewar 1940s when economic regulation dominated the Court's agenda. But he also adopted an accommodating attitude toward the

executive in wartime and in national security matters during the Cold War, and so scores as a "conservative" from the wartime 1940s through the end of his tenures on the Court, when those matters loomed large (O'Brien 1992). No one-dimensional score can account for docket-related swings in justices with distinct multidimensional preferences. Third, some legal issues fail to carry a reliable ideological valence. No ideological score is likely to accurately predict variation in voting patterns on technical, nonideological issues. Finally, some justices may actually shift their ideological stance over time. Needless to say, politically sophisticated adults of mature years like Supreme Court justices rarely display large ideological changes, but it is not impossible (Epstein et al. forthcoming).

Despite these caveats, we undertake the "acid test" in this and the following section. In this section, the dependent variable we employ is Martin and Quinn's behavioral voting score for the justices, often taken as the best one-statistic summary of voting tendency (Segal 2007). More specifically, we examine the ability of the NSP scores to predict confirmed nominees' first-year Martin-Quinn voting score, first-five-year average voting score, first-10-year average voting score, and life-time voting score (if the justice served more than a decade).[9] We contrast the performance of the NSP score with other obvious predictors, namely, the modified Segal-Cover score, the nominating president's DW-NOMINATE score, and both the modified Segal-Cover score and nominating president's DW-NOMINATE score taken in tandem. Additionally, we examine the changing utility of the predictors over time.

Table 1 shows the results of the analysis using all the data from 1937–2005. The four panels of the table consider four horizons for the voting scores (first year, five-year average, 10-year average, life-time score if service was longer than a decade). Each panel reports the results of OLS regressions for the predictor variables. As shown, NSP scores outperform the other predictor or predictors in each of the four time horizons, often quite substantially. Adding the MSC or presidential DW-NOMINATE score to the NSP score, either singly or jointly, does not improve the ability to predict subsequent voting behavior (analysis not shown).

The left-hand panel of Figure 3 contrasts the predictive ability of the NSP score with MSC scores and presidential DW-NOMINATE scores over the time period 1937–2005. The time horizon for life-time scores is shown at 21

---

[9]For the sake of comparability, we have scaled the MQ common space scores into the Senate space. The MQ scores are available on the web at ⟨mqscores.wustl.edu⟩. We thank Andrew Martin and Dennis Quinn for making these scores publicly available.

Table 1: Predicting Overall Voting Behavior of Justices Over Four Time Horizons, 1937–2005

| | *First-Year Voting Score (MQ)* | | | | *First-Five-Year Average Voting Score (MQ)*[a] | | | |
|---|---|---|---|---|---|---|---|---|
| *Variables* | *1* | *2* | *3* | *4* | *1* | *2* | *3* | *4* |
| Intercept | 0.03 | 0.02 | 0.03 | 0.02 | 0.02 | 0.01 | 0.02 | 0.01 |
| | (0.03) | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) | (0.05) | (0.12) |
| NSP score | 0.80 | — | — | — | 0.74 | — | — | — |
| | (0.08)*** | | | | (0.11)*** | | | |
| Segal-Cover score | — | –0.41 | — | — | — | –0.39 | — | — |
| | | (0.06)*** | | | | (0.07)*** | | |
| Best NOMINATE | — | — | 0.70 | — | — | — | 0.63 | — |
| | | | (0.09)*** | | | | (0.12)*** | |
| Presidential NOMINATE | — | — | — | 0.62 | — | — | — | 0.53 |
| | | | | (0.09)*** | | | | (0.12)*** |
| Df | 36 | 36 | 36 | 36 | 30 | 30 | 30 | 30 |
| *Adjusted* $R^2$ | 0.72 | 0.59 | 0.61 | 0.55 | 0.57 | 0.49 | 0.44 | 0.36 |
| | *First-10-Year Average Voting Score (MQ)*[2] | | | | *Life-Time Average Voting Score (MQ)*[b] | | | |
| | *1* | *2* | *3* | *4* | *1* | *2* | *3* | *4* |
| Intercept | –0.02 | –0.01 | –0.02 | –0.03 | –0.05 | –0.06 | –0.06 | –0.06 |
| | (0.05) | (0.05) | (0.06) | (0.06) | (0.16) | (0.06) | (0.06) | (0.07) |
| NSP score | 0.75 | — | — | — | 0.70 | — | — | — |
| | (0.13)*** | | | | (0.16)*** | | | |
| Segal-Cover score | — | –0.37 | — | — | — | –0.36 | — | — |
| | | (0.08)*** | | | | (0.06)*** | | |
| Best NOMINATE | — | — | 0.66 | — | — | — | 0.59 | — |
| | | | (0.14)*** | | | | (0.16)*** | |
| Presidential NOMINATE | — | — | — | 0.53 | — | — | — | 0.45 |
| | | | | (0.14)*** | | | | (0.16)** |
| Df | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |
| *Adjusted* $R^2$ | 0.57 | 0.45 | 0.46 | 0.35 | 0.42 | 0.35 | 0.31 | 0.21 |

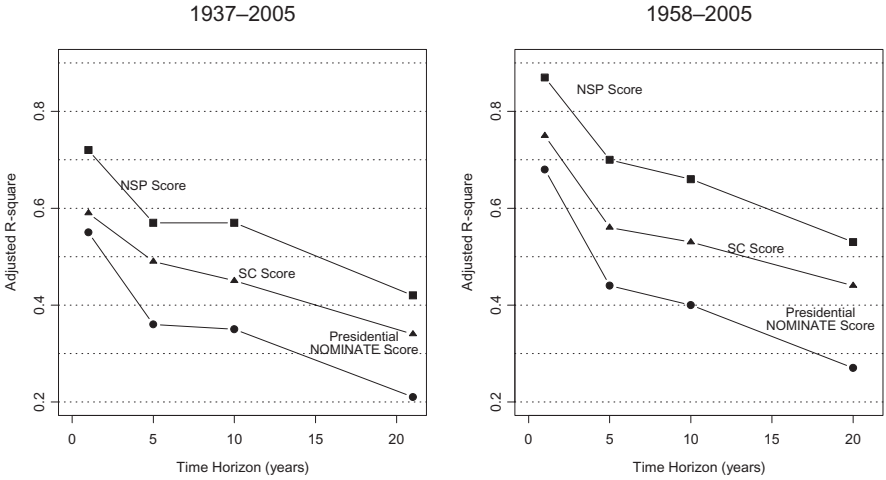
[a]For justices who served at least five years on the Supreme Court.

[b]For justices who served at least 10 years on the Supreme Court.

***$p < 0.01$; **$p < 0.05$; *$p < 0.10$. Standard errors are in parenthesis.

Dependent variable is indicated Martin-Quinn score, first converted into the judicial common space (Epstein et al. 2007a), and from there to the Senate DW-NOMINATE space.

*Figure 3:* Predicting future overall voting tendency on the Supreme Court, 1937–2005. Over the entire period (left panel), the NOMINATE-Scaled Perceptions (NSP) score substantially outperforms the Segal-Cover (SC) score and the appointed president's DW-NOMINATE score as a predictor of future aggregate voting behavior on the Supreme Court, as measured by Martin-Quinn voting scores. For justices appointed since 1957 (right panel), near-term behavior has become quite predictable. In this period, NSP scores remain the superior predictor.




years, the average tenure (with scores) for justices in this period who served longer than a decade. As shown, the NSP score does about as well predicting average five-year voting behavior as the modified Segal-Cover score does in predicting first-year voting behavior, and a better job predicting average 10-year behavior than the modified Segal-Cover score does predicting average five-year behavior. In fact, the NSP score predicts average behavior over a decade of service on the Court only slightly less well than MSC score predicts immediate behavior.

An important question concerns the behavior of the predictors over time: Have nominees become more predictable? A formal statistical test as well as examination of residuals suggests a structural break occurred in the late 1950s, after which the behavior of nominees became much more predictable.[10] Accordingly, Table 2 duplicates the analysis in Table 1, but splits

[10]A Zivot-Andrews test indicates a structural break at 1957 at the 0.01 level (Zivot & Andrews 1992).

Table 2: Predicting Overall Voting Behavior of Justices Over Four Time Horizons, 1937–1957 and 1958–2005

*Justices Nominated 1937–1957*

| | *First-Year Voting Score (MQ)* | | | | *First-Five-Year Average Voting Score(MQ)*[a] | | | |
|---|---|---|---|---|---|---|---|---|
| *Variables* | *1* | *2* | *3* | *4* | *1* | *2* | *3* | *4* |
| Intercept | 0.00 | −0.08 | −0.06 | −0.07 | 0.05 | −0.02 | −0.05 | −0.09 |
| | (0.01) | (0.08) | (0.10) | (0.13) | (0.12) | (0.10) | (0.12) | (0.15) |
| NSP score | 0.70 | — | — | — | 0.77 | — | — | — |
| | (0.30)** | | | | (0.35)** | | | |
| Segal-Cover score | — | −0.24 | — | — | — | −0.31 | — | — |
| | | (0.13)* | | | | (0.16)* | | |
| Best NOMINATE | — | — | 0.48 | — | — | — | 0.46 | — |
| | | | (0.30) | | | | (0.34) | |
| Presidential NOMINATE | — | — | — | 0.38 | — | — | — | 0.26 |
| | | | | (0.38) | | | | (0.44) |
| Df | 15 | 15 | 15 | 15 | 14 | 14 | 14 | 14 |
| *Adjusted* $R^2$ | 0.21 | 0.12 | 0.09 | 0.00 | 0.21 | 0.16 | 0.05 | −0.05 |
| | *First-10-Year Average Voting Score (MQ)*[b] | | | | *Life-Time Average Voting Score (MQ)*[b] | | | |
| | *1* | *2* | *3* | *4* | *1* | *2* | *3* | *4* |
| Intercept | 0.01 | −0.07 | −0.08 | −0.16 | 0.01 | −0.03 | −0.11 | −0.22 |
| | (0.16) | (0.15) | (0.14) | (0.18) | (0.18) | (0.16) | (0.16) | (0.20) |
| NSP score | 0.75 | — | — | — | 0.72 | — | — | — |
| | (0.47) | | | | (0.53) | | | |
| Segal-Cover score | — | −0.25 | — | — | — | −0.31 | — | — |
| | | (0.23) | | | | (0.24) | | |
| Best NOMINATE | — | — | 0.53 | — | — | — | 0.36 | — |
| | | | (0.42) | | | | (0.48) | |
| Presidential NOMINATE | — | — | — | 0.16 | — | — | — | −0.12 |
| | | | | (0.55) | | | | (0.60) |
| Df | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| *Adjusted* $R^2$ | 0.13 | 0.02 | 0.05 | −0.10 | 0.08 | 0.06 | −0.05 | −0.11 |

Table 2 *Continued*

*Justices Nominated 1958–2005*

| | *First-Year Voting Score (MQ)* | | | | *First-Five-Year Average Voting Score (MQ)*[a] | | | |
|---|---|---|---|---|---|---|---|---|
| | *1* | *2* | *3* | *4* | *1* | *2* | *3* | *4* |
| Intercept | 0.03 | 0.01 | 0.06 | 0.04 | –0.00 | –0.02 | 0.05 | 0.04 |
| | (0.03) | (0.04) | (0.09) | (0.05) | (0.05) | (0.06) | (0.06) | (0.15) |
| NSP score | 0.81 | — | — | — | 0.80 | — | — | — |
| | (0.07)*** | | | | (0.13)*** | | | |
| Segal-Cover score | — | –0.47 | — | — | — | –0.44 | — | — |
| | | (0.06)*** | | | | (0.10)*** | | |
| Best NOMINATE | — | — | 0.69 | — | — | — | 0.62 | — |
| | | | (0.09)*** | | | | (0.14)*** | |
| Presidential NOMINATE | — | — | — | 0.62 | — | — | — | 0.52 |
| | | | | (0.09)*** | | | | (0.15)*** |
| Df | 19 | 19 | 19 | 19 | 14 | 14 | 14 | 14 |
| *Adjusted* $R^2$ | 0.87 | 0.75 | 0.76 | 0.68 | 0.70 | 0.56 | 0.56 | 0.44 |
| | *First-10-Year Average Voting Score (MQ)*[b] | | | | *Life-Time Average Voting Score (MQ)*[b] | | | |
| | *1* | *2* | *3* | *4* | *1* | *2* | *3* | *4* |
| Intercept | –0.05 | –0.07 | 0.01 | –0.01 | –0.12 | –0.14 | –0.06 | –0.07 |
| | (0.06) | (0.07) | (0.07) | (0.08) | (0.08) | (0.09) | (0.08) | (0.09) |
| NSP score | 0.84 | — | — | — | 0.85 | — | — | — |
| | (0.15)*** | | | | (0.20)*** | | | |
| Segal-Cover score | — | –0.47 | — | — | — | –0.49 | — | — |
| | | (0.11)*** | | | | (0.14)*** | | |
| Best NOMINATE | — | — | 0.64 | — | — | — | 0.63 | — |
| | | | (0.16)*** | | | | (0.20)*** | |
| Presidential NOMINATE | — | — | — | 0.53 | — | — | — | 0.52 |
| | | | | (0.16)*** | | | | (0.20)** |
| Df | 14 | 14 | 14 | 14 | 14 | 14 | 14 | 14 |
| *Adjusted* $R^2$ | 0.66 | 0.53 | 0.51 | 0.40 | 0.53 | 0.44 | 0.38 | 0.27 |

[a]For justices who served at least five years on the Supreme Court.
[b]For justices who served at least 10 years on the Supreme Court.
***$p < 0.01$; **$p < 0.05$; *$p < 0.10$. Standard errors are in parentheses.

the data into two time periods, 1937–1957 and 1958–2005. As shown, none of the predictors in the earlier period are not terribly successful in predicting later behavior, though the NSP scores fare somewhat better than the other candidates. However, the behavior of justices confirmed after 1957 becomes dramatically more predicable.

The right-hand panel of Figure 3 contrasts the predictive ability of the NSP scores with MSC scores and presidential DW-NOMINATE scores during the more recent period.[11] As shown, the near-term behavior of justices has become extremely predictable, particularly with NSP scores. The NSP score does better predicting average five-year voting behavior than the modified Segal-Cover score does in predicting first-year voting behavior, and a better job predicting average 10-year behavior than the modified Segal-Cover score does predicting average five-year behavior.

## IV. Predicting Voting Tendency in Issue Areas

### *A. "Liberal" Votes by Issue Area*

The U.S. Supreme Court Judicial Database (Spaeth 2006) provides a "liberal direction" coding for each vote cast by a justice. The coding is based on the identity of litigants. For example, "liberal" votes support government regulation in economics cases, unions in unions cases, and nongovernmental actors in most civil liberties cases. Judicial politics scholars often employ Spaeth's "liberal direction" code as the basis for evaluating justices' ideological tendency (see, e.g., Epstein et al. 2007b).[12] The standard measure is simply the percentage of votes cast in the "liberal direction," either overall or in a particular class of cases. We refer to this measure as "litigant-liberalism" to distinguish it from other possible measures of liberal outcomes (Harvey 2006).

Segal and Cover (1989) reported considerable success in predicting litigant-liberalism in civil liberties and civil rights cases, using the Segal-Cover scores. For example, a simple OLS regression yielded an $R^2$ of 0.62.

---

[11]The time horizon for life-time scores is shown at 20 years, the average tenure (with scores) for justices in this period who served longer than a decade.

[12]An earlier literature applied factor analysis to votes to derive ideology measures by issue area (Schubert 1965). Martin-Quinn and similar item-response theoretic scores (Bafumi et al. 2005), as well as the domain-specific item-response theoretic scores analyzed in the following section, represent the modern continuation of this line of inquiry.

Epstein and Mershon 1996 updated and extended this analysis to cases in six issue areas: civil liberties, unions, economic activity, federalism, judicial power, and taxation. They found the Segal-Cover scores continued to predict civil liberties litigant-liberalism, though not as well as in the earlier collection of votes (adjusted $R^2 = 0.43$). They also found a degree of predictive success for union cases ($R^2 = 0.28$). However, they found disappointing predictive ability for economic activity cases ($R^2 = 0.18$), federalism cases ($R^2 = 0.12$), judicial power cases ($R^2 = 0.04$), and taxation cases ($R^2 = 0$). The Segal-Cover score was statistically insignificant at the 0.10 level for the latter three categories. These results led Epstein and Mershon to warn against the use of Segal-Cover scores outside of the area of civil liberties.

We replicate Epstein and Mershon's analysis using the U.S. Supreme Court Judicial Database, updated through 2005.[13] On average, a justice's litigant-liberalism in civil liberties cases is correlated with that in economics and judicial power, and to a lesser degree with unions and federalism. Litigant-liberalism in taxation cases appears uncorrelated with litigant-liberalism in any other category, except unions.

Table 3 indicates the results for OLS regressions of the Segal-Cover score and the NSP score on percent litigant-liberal votes in each of the six areas from 1953–2005. The ability of Segal-Cover scores to predict percent litigant-liberalism in civil liberties cases, union cases, and taxation remains similar to that reported by Epstein and Mershon. However, Segal-Cover scores do somewhat better than formerly in economics and federalism cases. The OLS coefficients on Segal-Cover scores are now statistically different from zero in all the categories except taxation.

Perhaps more striking, however, is the superior ability of NSP scores to predict litigant-liberalism in issue-specific domains. NSP scores are superior predictors in every issue area except civil liberties, where they perform equivalently to Segal-Cover scores. Neither predictor is at all successful predicting voting in taxation cases.

Table 4 further investigates the increased predictability of justices confirmed after 1957. Both measures display dramatically improved performance as predictors in civil liberties cases and economics cases, which are

---

[13]We include cases whose ANALU = 0 and DEC_TYPE is 1 (orally argued cases with signed opinion), 5 (cases with an equally divided vote), 6 (orally argued per curiam cases), or 7 (judgments of the Court). We include unanimous cases and only justices who participated in more than 10 decisions (following Epstein & Mershon 1996). The updated scores are available at ⟨http://www.princeton.edu/~ccameron⟩.

Table 3: Predicting the Justices' Percent Litigant-Liberal Votes in Six Topical Areas, 1937–2005

| | *Segal-Cover* | | *NSP* | | *Best NOMINATE* | |
|---|---|---|---|---|---|---|
| *Dependent Variable* | *Coefficient* | *Adj* $R^2$ | *Coefficient* | *Adj* $R^2$ | *Coefficient* | *Adj* $R^2$ |
| Civil liberties | 0.21 | 0.42 | –0.35 | 0.41 | –0.27 | 0.27 |
| | (0.04)*** | | (0.08)*** | | (0.08)*** | |
| Unions | 0.13 | 0.31 | –0.29 | 0.51 | –0.27 | 0.52 |
| | (0.04)*** | | (0.05)*** | | (0.05)*** | |
| Economics | 0.1 | 0.19 | –0.22 | 0.28 | –0.20 | 0.25 |
| | (0.04)*** | | (0.06)*** | | (0.06)*** | |
| Federalism | 0.08 | 0.28 | –0.17 | 0.40 | –0.16 | 0.37 |
| | (0.02)*** | | (0.04)*** | | (0.04)*** | |
| Judicial power | 0.07 | 0.28 | –0.12 | 0.28 | –0.09 | 0.16 |
| | (0.02)*** | | (0.04)*** | | (0.04)** | |
| Taxation | 0.02 | –0.03 | –0.07 | 0.01 | –0.09 | 0.06 |
| | (0.04) | | (0.06) | | (0.06) | |

***$p < 0.01$; **$p < 0.05$; *$p < 0.10$. Includes only justices who participated in more than 10 cases in a topical area.

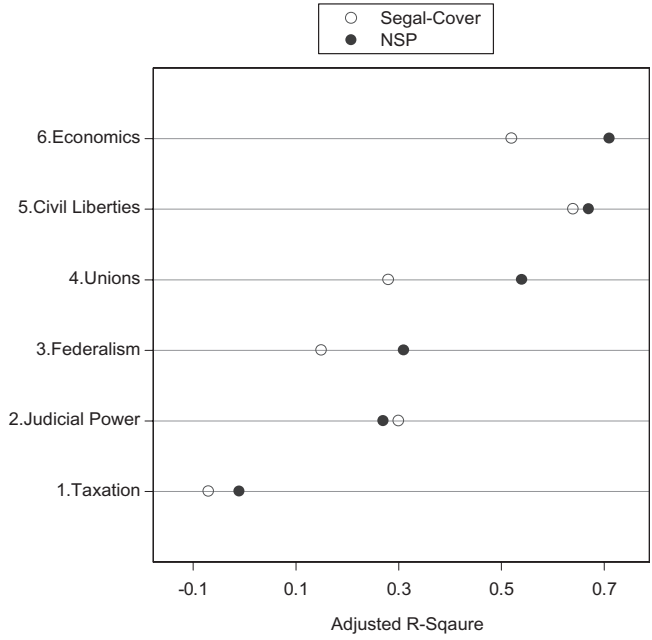
Table 4: Predicting the Justices' Percent Litigant-Liberal Votes in Six Topical Areas, 1958–2005

| | *Segal-Cover* | | *NSP* | | *Best NOMINATE* | |
|---|---|---|---|---|---|---|
| *Dependent Variable* | *Coefficient* | *Adj* $R^2$ | *Coefficient* | *Adj* $R^2$ | *Coefficient* | *Adj* $R^2$ |
| Civil liberties | 0.26 | 0.64 | –0.43 | 0.67 | –0.34 | 0.52 |
| | (0.05)*** | | (0.07)*** | | (0.08)*** | |
| Unions | 0.13 | 0.28 | –0.29 | 0.54 | –0.28 | 0.62 |
| | (0.05)** | | (0.06)*** | | (0.05)*** | |
| Economics | 0.11 | 0.52 | –0.20 | 0.71 | –0.17 | 0.61 |
| | (0.02)*** | | (0.03)*** | | (0.03)*** | |
| Federalism | 0.06 | 0.15 | –0.13 | 0.31 | –0.13 | 0.38 |
| | (0.03)* | | (0.05)*** | | (0.04)*** | |
| Judicial power | 0.07 | 0.30 | –0.10 | 0.27 | –0.07 | 0.13 |
| | (0.02)*** | | (0.04)*** | | (0.04)* | |
| Taxation | 0.00 | –0.07 | –0.07 | –0.01 | –0.10 | 0.08 |
| | (0.04) | | (0.07) | | (0.06) | |

***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

the largest categories of cases. However, they perform somewhat less well in the area of federalism (the Segal-Cover scores particularly so). As shown in Figure 4, the NSP score outperforms the Segal-Cover in predicting the justices' litigant-liberalism in economics, unions, and federalism cases. The two

*Figure 4:* Relative ability of Segal-Cover scores and NSP scores to predict voting behavior ("liberal votes") in six issue-specific domains, 1958–2005. The NSP score outperforms the Segal-Cover score in predicting the justices' litigant-liberalism in economics, federalism, and unions cases. In the areas of civil liberties and judicial power, the two measures perform comparably. Neither predicts litigant-liberalism in taxation cases at all well.




measures perform comparably in the areas of civil liberties and judicial power. Again, neither is successful in predicting voting in taxation cases.

Segal et al. (2000) employ litigant-liberalism to characterize the ideological tendency of justices over a longer time period (1937–1994) but in a more aggregated set of cases, "civil liberties policy" and "economic policy." Civil liberties policy combines First Amendment rights, criminal procedures, equal protection, due process, privacy, and attorneys in the original classification in the U.S. Supreme Court Judicial Database. The economic policy category merges all economics and unions cases in the database. We have updated these two scores through 2005.[14]

---

[14]The updated scores are available at ⟨http://www.princeton.edu/~ccameron⟩.

Table 5: Predicting Percent Liberal Votes in Two Broad Issue Areas, 1937–2005, 1937–1957, 1958–2005

| | *Civil Rights* | | | *Economics* | | |
|---|---|---|---|---|---|---|
| | *Model 1* | *Model 2* | *Model 3* | *Model 1* | *Model 2* | *Model 3* |
| *Predicting Voting Tendency 1937–2005* | | | | | | |
| Best NOM | — | — | −7.57 (8.38) | — | — | −11.82 (5.27)** |
| Segal-Cover | — | 20.68 (3.78)*** | 17.38 (5.27)*** | — | 12.83 (2.53)*** | 7.64 (3.31)** |
| NSP | −35.38 (6.66)*** | — | — | −24.18 (4.08)*** | — | — |
| Adj $R^2$ | 0.44 | 0.46 | 0.46 | 0.50 | 0.42 | 0.49 |
| Sample size | 35 | 35 | 35 | 35 | 35 | 35 |
| *Predicting Voting Tendency 1937–1957* | | | | | | |
| Best NOM | — | — | 12.06 (20.17)* | — | — | −18.92 (16.54) |
| Segal-Cover | — | 22.76 (10.46)** | 25.07 (11.41)** | — | 13.35 (8.90) | 9.72 (9.35) |
| NSP | −31.97 (23.53) | — | — | −34.34 (17.40)* | — | — |
| Adj $R^2$ | 0.06 | 0.21 | 0.17 | 0.17 | 0.08 | 0.10 |
| Sample size | 15 | 15 | 15 | 15 | 15 | 15 |
| *Predicting Voting Tendency 1958–2005* | | | | | | |
| Best NOM | — | — | −14.25 (7.80)* | — | — | −11.02 (2.88)*** |
| Segal-Cover | — | 27.01 (4.49)*** | 19.89 (5.75)*** | — | 11.10 (2.06)*** | 5.59 (2.12)** |
| NSP | −44.45 (6.83)*** | — | — | −20.21 (2.44)*** | — | — |
| Adj $R^2$ | 0.68 | 0.65 | 0.69 | 0.78 | 0.59 | 0.77 |
| Sample size | 20 | 20 | 20 | 20 | 20 | 20 |

***$p < 0.01$; **$p < 0.05$ *$p < 0.10$.

The results reported in Table 5 analyze the ability of Segal-Cover and NSP scores to predict justices' litigant-liberalism in these two categories for the longer time series, 1937–2005. Models 1 and 2 in the table reconfirm the previous finding: the NSP scores predict economic policy scores considerably better than the Segal-Cover scores, while performing equally well in predicting the civil liberties policy scores. The behavior of justices in these broad categories became much more predictable after the late 1950s. (We discuss Model 3 in Table 5 in Section VI.)

*B. Item-Response Theoretic Scores by Issue Area*

The Spaeth "liberal direction" coding has an intuitive appeal, but Harvey (2006) criticizes it as occasionally arbitrary or unnatural, and untethered by a sense of the policy status quo. An alternative eschews the "liberal direction" coding and instead uses item-response theory to scale nonunanimous votes, as in Martin and Quinn (2002) and Bafumi et al. (2005). Under rather strong assumptions (such as sincere voting and an exogenous and unchanging docket), the resulting scores can be interpreted as judicial ideal points (Jackman 2001; Clinton et al. 2004). Without committing to such a strong position, one may view item-response theoretic scores as an alternative summary measure of ideological voting tendency free from the alleged idiosyncrasies of the "liberal outcome" coding.

Using the same data we employed to update Epstein and Mershon's study but avoiding the "liberal direction" coding, we employ Markov chain Monte Carlo to generate item-response theoretic scores for the six domain-specific categories.[15] We employ only justices who participated in at least 10 cases in each of the six categories, from 1953 to 2005. Again due to the sparseness of data, we estimate life-time scores for each justice rather than term-varying scores. Thus, these scores are analogous to Martin and Quinn's or Bafumi et al.'s (2005) "constant ideal point" estimates of overall ideology, but domain specific.[16] Broadly speaking, excluding taxation, a given justice's scores across the issue areas tend to be similar. The item-response theoretic scores tend to be highly correlated with percent litigant-liberal votes in each issue area, as well as with Martin and Quinn's "constant ideal point" scores.

Table 6 indicates the results of OLS regressions of the Segal-Cover and NSP scores on the domain-specific item-response theoretic scores. The top panel in the table examines all the data from 1953–2005, while the lower panel focuses on 1958–2005. The analysis confirms that NSP scores offer superior performance in predicting voting behavior in non-civil-liberties cases and equivalent performance in civil liberties cases, and that the behavior of the justices became more predictable after 1957.

---

[15]We use Martin and Quinn's R package "MCMCpack" to derive the scores. Each estimate was acquired after 200,000 Gibbs scans following 20,000 burn-in scans. Standard diagnostics performed on the posterior samples suggest that the chains reached a steady state. Detailed results from MCMCpack are available on request.

[16]The scores are available at ⟨http://www.princeton.edu/~ccameron⟩.

Table 6: Predicting the Justices' Item-Response Theoretic Scores in Six Topical Areas, 1937–2005 and 1958–2005

| | *Segal-Cover* | | *NSP* | | *Best NOMINATE* | |
|---|---|---|---|---|---|---|
| *Dependent Variable* | *Coefficient* | *Adj* $R^2$ | *Coefficient* | *Adj* $R^2$ | *Coefficient* | *Adj* $R^2$ |
| *Predicting Voting Tendency, 1937–2005* | | | | | | |
| Civil liberties | 1.56 | 0.32 | –2.73 | 0.33 | –2.20 | 0.23 |
| | (0.40)*** | | (0.69)*** | | (0.69)*** | |
| Unions | 0.84 | 0.27 | –1.85 | 0.44 | –1.76 | 0.44 |
| | (0.26)*** | | (0.41)*** | | (0.385)*** | |
| Economics | 0.64 | 0.11 | –1.43 | 0.20 | –1.37 | 0.21 |
| | (0.30)** | | (0.51)*** | | (0.48)*** | |
| Federalism | 0.78 | 0.21 | –1.87 | 0.40 | –1.69 | 0.39 |
| | (0.30)** | | (0.46)*** | | (0.43)*** | |
| Judicial power | 0.76 | 0.21 | –1.44 | 0.25 | –1.18 | 0.19 |
| | (0.27)*** | | (0.46)*** | | (0.44)** | |
| Taxation | 0.11 | –0.03 | –0.46 | 0.00 | –0.55 | 0.03 |
| | (0.26) | | (0.45) | | (0.41) | |
| *Predicting Voting Tendency, 1958–2005* | | | | | | |
| Civil liberties | 1.94 | 0.64 | –3.23 | 0.68 | –2.64 | 0.56 |
| | (0.33)*** | | (0.51)*** | | (0.55)*** | |
| Unions | 0.79 | 0.35 | –1.71 | 0.62 | –1.67 | 0.69 |
| | (0.26)*** | | (0.34)*** | | (0.28)*** | |
| Economics | 0.60 | 0.29 | –1.26 | 0.50 | –1.16 | 0.52 |
| | (0.22)** | | (0.31)*** | | (0.27)*** | |
| Federalism | 0.72 | 0.13 | –1.85 | 0.43 | –1.68 | 0.48 |
| | (0.41)** | | (0.55)*** | | (0.45)*** | |
| Judicial power | 0.83 | 0.45 | –1.45 | 0.50 | –1.19 | 0.40 |
| | (0.22)*** | | (0.35)*** | | (0.35)* | |
| Taxation | 0.22 | 0.00 | –0.75 | 0.19 | –0.84 | 0.33 |
| | (0.23) | | (0.34)** | | (0.28)*** | |

***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

## V. Out-of-Sample Predictions

Because their tenure on the Supreme Court is so brief at the time we write, Justices John Roberts and Samuel Alito do not figure in the analysis in the preceding sections. However, we have calculated NSP scores for both. Accordingly, these justices afford an opportunity to test the models by making out-of-sample predictions about their future voting behavior. Of course, the accuracy of the predictions will not be known for some time.

Table 7: Out-of-Sample Predictions: Justices Roberts and Alito

| | *Roberts* | | *Alito* | |
|---|---|---|---|---|
| *Scores* | *Prediction* | *(95% CI)* | *Prediction* | *(95% CI)* |
| 5-year MQ | 1.92 | (1.36, 2.57) | 1.90 | (1.35, 2.54) |
| 10-year MQ | 1.91 | (1.12, 2.87) | 1.90 | (1.12, 2.84) |
| Civil liberties[a] | 0.06 | (–0.04, 0.16) | 0.06 | (–0.04, 0.15) |
| Unions[a] | 0.15 | (0.06, 0.24) | 0.15 | (0.06, 0.24) |
| Economics[a] | 0.17 | (0.13, 0.21) | 0.17 | (0.12, 0.21) |
| Federalism[a] | 0.24 | (0.18, 0.31) | 0.24 | (0.17, 0.30) |
| Judicial power[a] | 0.14 | (0.09, 0.19) | 0.14 | (0.09, 0.19) |
| Civil rights policy[b] | 0.05 | (–0.05, 0.15) | 0.05 | (–0.05, 0.15) |
| Economic policy[b] | 0.17 | (0.14, 0.21) | 0.17 | (0.14, 0.21) |

[a]Topical categories defined by Epstein and Mershon (1996) (see text).

[b]Broad categories defined in Segal et al. (2000) (see text).

For all percentage liberal votes, 0 denotes conservative, 1 indicates liberal.

Table 7 provides nine out-of-sample predictions about the future voting behavior of each of these justices, based on the earlier OLS regression models and the justices' NSP scores. Shown are predicted average five-year and average 10-year Martin-Quinn scores, predicted percent litigant-liberal voting in five topical areas (excluding taxation), and predicted percent litigant-liberal voting in two broad topical areas, civil liberties policy and economic policy.

The predicted Martin-Quinn scores for the two justices suggest an overall voting pattern somewhat similar to that of Chief Justice Warren Burger, or Justice Antonin Scalia at an equivalent point in his career. The predicted percentage of litigant-liberal votes in five topical areas and two broader areas are again similar to those of Chief Justice Warren Burger and Justice Scalia, and also to those of Chief Justice William Rehnquist.

## VI. Discussion and Conclusion

The preceding analysis raises two questions. First, why do the NSP scores outperform the Segal-Cover scores in the manner they do? And, second, why has the behavior of confirmed justices become so much more predictable?

Because the NSP scores incorporate additional information about nominee ideology, it may not be surprising that they outperform the Segal-Cover scores in predicting nominees' future overall behavior. Perhaps more

intriguing is the variegated ability of the NSP scores in the topical areas: putting aside tax cases (where no measure predicts future behavior), NSP scores consistently outperform Segal-Cover scores in most areas except civil liberties, where both perform relatively well but equivalently.

A plausible reason for this pattern is the following. As Segal and Cover note, in the post-1952 period, the newspaper editorialists tended to focus on nominees' positions on civil liberties and civil rights. Their positions on other issues received relatively short shrift. In contrast, the nominees' DW-NOMINATE "tracks" reveal information about their likely positions on issues closely associated with the first DW-NOMINATE dimension—typically matters involving economics and redistribution (Poole & Rosenthal 1997). Because the NSP scores incorporate the editorialists' information on civil rights and civil liberties, they perform nearly equivalently to the Segal-Cover scores in that area, but because they also incorporate information about the nominees' positions on first DW-NOMINATE issues, they perform better in issue areas related to that dimension, for example, economics.

Model 3 in Table 5 provides some support for this conjecture. This model employs both components of the NSP score *separately* to predict the justices' voting tendency in the two broad areas, civil liberties policy and economic policy. As shown, the best available NOMINATE score adds virtually nothing to the Segal-Cover score as a predictor of voting in civil liberties policy cases. However, in the economic policy cases, both the best available NOMINATE score and the Segal-Cover score are statistically significant predictors. The regression that incorporates both measures is a better predictor than a regression that employs only the Segal-Cover score. These results suggest that the Segal-Cover score is primarily a civil liberties predictor and only secondarily taps into first dimension DW-NOMINATE issues. In contrast, the best available NOMINATE score primarily taps into first dimension DW-NOMINATE issues. Combining both measures yields a single measure that works as well as the Segal-Cover scores in predicting civil liberties voting, and is superior in predicting voting in most other issues, particularly economics-related cases.

Why has the behavior of confirmed justices become so much more predictable? A careful investigation of this question lies outside the scope of this article. We note, however, that from 1937–1957, less than half (47 percent) of the nominees were professional jurists (i.e., judges). In contrast, from 1958–2005, over three-quarters (78 percent) were professional jurists, a two-thirds increase. In fact, from 1967–2005, 87 percent of the nominees were professional jurists. It is hardly a shock that the past behavior of a

professional jurist is an excellent predictor of his or her future behavior as a jurist. This may be particularly true when the jurist has had to deal with controversial, high-profile legal issues like abortion, capital punishment, gay rights, euthanasia, and so on. Conversely, the past behavior of a nonprofessional jurist, while supplying useful information, is not as good a predictor of his or her likely voting as a jurist.

From this perspective, it is not so surprising that the future behavior of nominees has become so much more predictable. What is more striking is how presidents have altered the type of person advanced to the high court. Whether by design or not, the result is nominees who are more likely to consistently and predictably pursue a given set of ideological goals in their future voting.

In conclusion, we find that predicting the future behavior of nominees to the U.S. Supreme Court is less problematic than has been supposed. This is particularly true in the modern period, after the late 1950s. In light of this fact, the centrality of perceived nominee ideology in Supreme Court confirmation politics is entirely understandable and, some might conclude, largely justified.

## References

Bafumi, Joseph, Andrew Gelman, & David Park (2005) "Issues in Ideal Point Estimation," 13 *Political Analysis* 171.

Cameron, Charles, Albert Cover, & Jeffrey Segal (1990) "Senate Voting on Supreme Court Nominees: A Neoinstitutional Model," 84(2) *American Political Science Rev.* 525.

Clinton, Joshua, Simon Jackman, & Douglas Rivers (2004) "The Statistical Analysis of Roll Call Voting: A Unified Approach," 98(2) *American Political Science Rev.* 355.

Epstein, Lee, Rene Lindstadt, Jeffrey Segal, & Chad Westerland (2006) "The Changing Dynamics of Senate Voting on Supreme Court Nominees," 68(2) *J. of Politics* 296.

Epstein, Lee, Andrew Martin, Kevin Quinn, & Jeffrey Segal (forthcoming) "Ideological Drift Among Supreme Court Justices: Who, When, and How Important?" *Northwestern Univ. Law Rev.*

Epstein, Lee, Andrew Martin, Jeffrey Segal, & Chad Westerland (2007a) "The Judicial Common Space," 23 *J. of Law, Economics, & Organization* 303.

Epstein, Lee, & Carol Mershon (1996) "Measuring Political Preferences," 40(1) *American J. of Political Science* 261.

Epstein, Lee, & Jeffrey Segal (2005) *Advice and Consent: The Politics of Judicial Appointments.* New York: Oxford Univ. Press.

Epstein, Lee, Jeffrey Segal, Harold Spaeth, & Thomas G. Walker (2007b) *The Supreme Court Compendium*, 4th ed. Washington DC: CQ Press.

Giles, Michael, Virginia Hettinger, & Todd Peppers (2001) "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas," 54 *Political Research Q.* 623.

Harvey, Anna (2006) "What Makes a Judgment 'Liberal'? Measurement Error in the United States Supreme Court Judicial Database," Working Paper, Department of Politics, New York Univ. Available at ⟨http://www.nyu.edu/gsas/dept/politics/faculty/harvey/harveymeasurementerror.pdf⟩.

Jackman, Simon (2001) "Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference and Model Checking," 9 *Political Analysis* 727.

Martin, Andrew, & Kevin M. Quinn (2002) "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for U.S. Supreme Court 1953–1999," 10 *Political Analysis* 134.

—— (2005) "Can Ideal Point Estimates be Used as Explanatory Variables?" Working Paper, Washington Univ. at St. Louis.

McCarty, Nolan, & Keith Poole (1995) "Veto Power and Legislation: An Empirical Analysis of Executive and Legislative Bargaining from 1961 to 1986," 11(2) *J. of Law, Economics, & Organization* 82.

O'Brien, David (1992) "Stanley Forman Reed," in Kermit Hall, ed., *The Oxford Companion to the Supreme Court*, pp. 712–13. New York: Oxford Univ. Press.

Poole, Keith, & Howard Rosenthal (1997) *Congress: A Political-Economic History of Roll-Call Voting*. New York: Oxford Univ. Press.

Schubert, Glendon (1965) *The Judicial Mind: The Attitudes and Ideologies of Supreme Court Justices 1946–1963*. Evanston, IL: Northwestern Univ. Press.

Segal, Jeffrey (2007) "Measuring Political Preferences," 17(2) *Law & Courts* 1.

Segal, Jeffrey, & Albert Cover (1989) "Ideological Values and the Votes of U.S. Supreme Court Justices," 83(2) *American Political Science Rev.* 557.

Segal, Jeffrey, Lee Epstein, Charles Cameron, & Harold Spaeth (1995) "Ideological Values the Votes of U.S. Supreme Court Justices Revisited," 57(3) *J. of Politics* 812.

Segal, Jeffrey, Richard J. Timpone, & Robert M. Howard (2000) "Buyer Beware? Presidential Success Through Supreme Court Appointments," 53(3) *Political Research Q.* 557.

Spaeth, Harold (2006) *The U.S. Supreme Court Judicial Data Base*. Available at ⟨www.as.uky.edu/polisci/ulmerproject/sctdata.html⟩.

Szmer, John, & Donald Songer (2005) "The Effects of Information on the Accuracy of Presidential Assessments of Supreme Court Nominee Preferences," 58 *Political Research Q.* 151.

Zivot, Eric, & Donald Andrews (1992) "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," 10(3) *J. of Business & Economic Statistics* 251.

# Appendix: Preconfirmation Measures of Nominee Ideology, 1937–2006

| *Nominee* | *Segal-Cover Score* | *Modified Segal-Cover (MSC) Score* | *Best NOMINATE Score* | *Factor* | *NOMINATE-Scaled Perceptions (NSP) Score* | *Sources of Best NOMINATE Score** |
|---|---|---|---|---|---|---|
| Alito | 0.100 | 0.359 | 0.536 | 1.784 | 0.509 | A(P) |
| Black | 0.875 | –0.355 | –0.482 | –1.589 | –0.482 | DW |
| Blackmun | 0.115 | 0.345 | 0.147 | 1.083 | 0.305 | A(P) |
| Bork | 0.095 | 0.363 | 0.558 | 1.832 | 0.523 | A(P) |
| Brennan | 1.000 | –0.470 | 0.034 | –0.962 | –0.302 | P |
| Breyer | 0.475 | 0.013 | –0.455 | –0.706 | –0.246 | A |
| Burger | 0.115 | 0.345 | 0.1470 | 1.083 | 0.305 | A |
| Burton | 0.280 | 0.193 | 0.148 | 0.740 | 0.203 | DW |
| BWhite | 0.500 | –0.010 | –0.550 | –0.922 | –0.268 | P |
| Byrnes | 0.330 | 0.147 | –0.320 | –0.275 | –0.062 | DW |
| Carswell | 0.040 | 0.414 | 0.356 | 1.599 | 0.456 | A |
| Clark | 0.500 | –0.010 | –0.412 | –0.684 | –0.214 | P |
| DGinsburg | 0.125 | 0.336 | 0.653 | 1.933 | 0.562 | A(P) |
| Douglas | 0.730 | –0.221 | –0.381 | –1.112 | –0.341 | P |
| Fortas1 | 1.000 | –0.470 | –0.460 | –1.812 | –0.549 | P |
| Fortas2 | 0.845 | –0.327 | –0.428 | –1.433 | –0.436 | C |
| Frankfurter | 0.665 | –0.162 | –0.381 | –0.976 | –0.301 | P |
| Goldberg | 0.750 | –0.240 | –0.550 | –1.444 | –0.438 | P |
| Harlan | 0.875 | –0.355 | 0.034 | –0.701 | –0.222 | A |
| Haynsworth | 0.160 | 0.303 | 0.147 | 0.989 | 0.277 | A |
| Kennedy | 0.365 | 0.115 | 0.343 | 0.897 | 0.248 | A(P) |
| Marshall | 1.000 | –0.470 | –0.460 | –1.812 | –0.603 | A(P) |
| Miers | 0.270 | 0.202 | 0.572 | 1.490 | 0.433 | P |
| Minton | 0.720 | –0.212 | –0.461 | –1.228 | –0.375 | DW |
| Murphy | 1.000 | –0.470 | –0.381 | –1.676 | –0.509 | P |
| OConnor | 0.415 | 0.069 | 0.653 | 1.326 | 0.372 | P |
| Powell | 0.165 | 0.299 | 0.184 | 1.042 | 0.292 | P |
| RBGinsburg | 0.680 | –0.175 | –0.455 | –1.134 | –0.391 | A |
| Reed | 0.725 | –0.217 | –0.381 | –1.101 | –0.338 | P |
| Rehnquist1 | 0.045 | 0.409 | 0.184 | 1.293 | 0.367 | P |
| Rehnquist2 | 0.045 | 0.409 | 0.651 | 2.097 | 0.601 | C |
| RJackson | 1.000 | –0.470 | –0.381 | –1.676 | –0.487 | P |
| Roberts | 0.120 | 0.340 | 0.572 | 1.804 | 0.514 | A(P) |
| Rutledge | 1.000 | –0.470 | –0.381 | –1.676 | –0.509 | P |
| Scalia | 0.000 | 0.451 | 0.558 | 2.031 | 0.582 | A(P) |
| Souter | 0.325 | 0.151 | 0.412 | 1.100 | 0.307 | A |
| Stevens | 0.250 | 0.221 | 0.039 | 0.615 | 0.167 | A |
| Stewart | 0.750 | –0.240 | 0.034 | –0.439 | 0.113 | A |
| Stone2 | 0.300 | 0.174 | –0.309 | –0.089 | –0.038 | C |
| Thomas | 0.150 | 0.313 | 0.581 | 1.757 | 0.478 | A |
| Thornberry | 1.000 | –0.470 | –0.384 | –1.681 | –0.489 | A |

## Appendix *Continued*

| *Nominee* | *Segal-Cover Score* | *Modified Segal-Cover (MSC) Score* | *Best NOMINATE Score* | *Factor* | *NOMINATE-Scaled Perceptions (NSP) Score* | *Sources of Best NOMINATE Score** |
|---|---|---|---|---|---|---|
| Vinson | 0.750 | –0.240 | –0.410 | –1.203 | –0.368 | DW |
| Warren | 0.750 | –0.240 | 0.034 | –0.439 | –0.146 | P |
| Whittaker | 0.500 | –0.010 | 0.147 | 0.278 | 0.066 | A(P) |

*DW = Nominee's DW-NOMINATE, A = Nominee served on U.S. Courts of Appeal, and score is the Giles/Hettinger/Pepper score based on the appointing president's DW-NOMINATE and the NOMINATE score or scores of the sponsoring senator or senators at the time of appointment to the circuit courts (courtesy appointment), A(P) = Nominee served on U.S. Courts of Appeal, and score is the Giles/Hettinger/Pepper score but based only on the appointing president's DW-NOMINATE at the time of appointment to the circuit court (noncourtesy appointment), C = Common space Martin-Quinn, P = President's DW-NOMINATE at the time of appointment as Supreme Court justice.

Note: All NOMINATE scores are converted into the Senate space.