

## Twenty-First Century Split: Partisan, Racial, and Gender Differences in Circuit Judges Following Earlier Opinions

Stuart Minor Benjamin, Kevin M. Quinn & ByungKoo Kim\*

*Judges shape the law with their votes and the reasoning in their opinions. An important element of the latter is which opinions they follow, and thus elevate, and which they cast doubt on, and thus diminish. Using a unique and comprehensive dataset containing the substantive Shepard's treatments of all circuit court published and unpublished majority opinions issued between 1974 and 2017, we examine the relationship between judges' substantive treatments of earlier appellate cases and their party, race, and gender. Are judges more likely to follow opinions written by colleagues of the same party, race, or gender? What we find is both surprising and nuanced. We have two major findings. First, over the forty-four-year span we studied, we find growing partisan differences in positive treatments of earlier cases. The partisan differences are largest for treatments in ideologically salient categories of cases. Interestingly, the partisan differences arise more for treatments of opinions written by Democratic*

---

\* Benjamin is the William Van Alstyne Professor of Law at Duke Law School; Quinn is a Professor at Emory Law School and in the Department of Quantitative Theory and Methods at the Emory College of Arts and Sciences; Kim is an Assistant Professor at the KDI School of Public Policy and Management. More than the usual thanks are in order. We thank LexisNexis, who shared their data with us in 2018 and patiently tracked down the answers to our many questions about the data over the next twenty-seven months. A dataset so massive required similarly massive work to identify, categorize, and organize its elements. Tom Balmat at Duke and Sean Chen at the Duke Law Library were particularly valuable in that lengthy process. Jane Bahnson and Wickliffe Shreve, also at the Duke Law Library, were indefatigable in helping us find relevant materials about *Shepard's*, Lexis, and court practices. Gloria Han, Andrew Tisinger, and Wenyi Zhou provided excellent research assistance, and Leanna Doty and Balfour Smith provided excellent editing assistance. Over the five-year gestation of the project, we have benefited from many colleagues' comments, including Matt Adler, Joseph Blocher, Curt Bradley, Michael Frakes, Jack Knight, Maggie Lemos, Marin Levy, Mat McCubbins, Darrell Miller, Jeff Powell, Neil Siegel, Georg Vanberg, Jonathan Wiener, and Ernie Young. We also thank colleagues who attended our 2021 Duke Law workshop on an early version of this Article as well as members of the Emory Law faculty who attended a 2022 workshop on a version of this Article.

*appointees than for opinions written by Republican appointees, which we think is best explained by an accelerating movement among Republican appointees in a conservative direction compared to a steady move among Democratic appointees in a liberal direction. The increase in partisan differences is not a function of presidential cohorts or age cohorts. Second, there are intraparty racial and gender differences in positive treatments of past cases, and these differences are similar to the partisan differences. Within each party, Black and White judges differ in their treatments of opinions authored by Black co-partisans, Hispanic and White judges differ in their treatments of opinions authored by Hispanic co-partisans, and female and male judges differ in their treatments of opinions authored by female co-partisans. Similar to the partisan divergence noted above, we also find that some of these differences increase in magnitude over time—with particularly notable increases in the Black-White Democratic differences, Hispanic-White Republican differences, and female-male Republican differences. Notably, the racial and gender differences we find in positive Shepard's treatments are not mirrored in most studies of racial and gender differences in judicial behavior, which focus on merits votes and include a much smaller number of cases.*

*These results defy easy explanation. They do not support the proposition that party, race, and gender have always played a pervasive role for judges. Instead, our results provide evidence of increasing partisan, racial, and gender polarization among judges in recent years. For reasons we explain in the body of this Article, the partisan, racial, and gender differences we find appear to be a function of political ideology. Further, because the racial and gender differences are within parties, our results indicate that not only partisan differences but also intraparty racial and gender ideological differences have risen in recent years (particularly for Republican judges).*

*Our data thus reveal polarization among circuit judges and, as a result, in their shaping of the law. Many groups in the United States have become more ideologically polarized in recent years. Our data indicate that judges are one of them.*

## CONTENTS

I. THE IMPORTANCE OF TREATMENTS OF EARLIER OPINIONS .....	376
II. THE EMPIRICAL LITERATURE ON JUDICIAL BEHAVIOR – PARTY, GENDER, AND RACE	387
A. Political Party .....	388
B. Gender and Race .....	390
1. Gender .....	390
2. Race .....	393
III. OUR RESEARCH QUESTIONS .....	395
A. Partisan Differences .....	395
B. Beyond Partisanship – Racial and Gender Differences .....	398
C. Solidarity Effects .....	401
D. Measures and Data Limitations with Respect to Race and Gender .....	406
IV. DATA AND RESEARCH DESIGN .....	408
A. Data .....	408
B. Research Design .....	411
C. Estimation and Inference .....	412
V. PRIMARY RESULTS .....	414
A. Average Controlled Differences by Author Partisanship .....	414
B. Average Controlled Differences by Author Race .....	417
C. Average Controlled Differences by Author Gender .....	423
VI. DISCUSSION .....	426
A. Partisan Differences .....	426
1. Partisan Differences Are Not a Function of Presidential or Age Cohorts .....	427
2. The Best Explanation: An Accelerating Move Right Among Republicans but a Steady Move Left Among Democrats .....	436
A. Racial Differences .....	443
C. Gender Differences .....	449
CONCLUSION .....	452
APPENDIX .....	453

## INTRODUCTION

Many researchers have studied the relationship between judges' behavior and their political party, race, or gender. Most of these studies focus on whether judges' votes in a particular set of cases are associated with the judges' party, race, or gender.<sup>1</sup> As we discuss in more detail in Part II, these studies have generally found

---

1. Throughout this Article, we generally use the term "race" to refer to both race and ethnicity. And, as we discuss below, in determining party we look to the political party of the President who most recently appointed the judge. See *infra* note 42 and accompanying text. To avoid wordiness, we sometimes refer to Democratic and Republican appointees by the shorthand "Democrats" and "Republicans."

that Democratic and Republican appointees vote differently in some categories of ideologically salient cases, that judges of different races vote differently in a smaller subset of cases, and that male and female judges vote differently in an even smaller subset of cases.<sup>2</sup>

Judges' votes are very important, but judges of course issue opinions, and the reasoning in those opinions helps to shape later cases. One significant aspect of a given opinion is which earlier cases it follows and which it casts doubt on. Using an original dataset containing all the substantive *Shepard's Citations* (*Shepard's*) treatments in all federal appellate majority opinions from 1974–2017 (670,784 opinions in total), we examine how a judge's party, race, and gender are associated with changes in an opinion's likelihood of following an earlier circuit opinion.<sup>3</sup>

The vast amount of data we have allows us to make adjustments for each of the 521 circuit-year combinations in our data.<sup>4</sup> This contrasts with most empirical studies of federal appellate decision making, which do not make such adjustments and include far fewer cases (usually only hundreds of cases).<sup>5</sup> We make these adjustments for all our analyses to reduce the possibility of producing biased results (e.g., by one circuit having a larger number of judges of a particular party, race, or gender in earlier years and another circuit having a larger number in later years, or by turnover on party, race, or gender lines within a circuit over the long period of our study).

---

2. See *infra* notes 42–68 and accompanying text; on “male” and “female,” see *infra* note 7.

3. We, and *Shepard's*, focus on majority opinions. See *infra* notes 28 and 99 and accompanying text. For conciseness, we use the term “opinions” to refer to majority opinions. And because we are focusing on majority opinions, we largely use the terms “opinion” and “case” synonymously.

4. There are 521 circuit-year combinations because our data cover forty-four years and the twelve regional circuits (we do not include the Federal Circuit because it has relatively few of the ideologically, race-, and gender-salient cases that we want to measure; see *infra* text accompanying note 100). The Eleventh Circuit did not exist in the first seven years of our data (because it was part of the Fifth Circuit), thus yielding 521 circuit-year combinations instead of 528.

5. See *infra* notes 67–68 and accompanying text.

What might one expect to find? At one extreme, we might expect no meaningful differences related to partisanship,<sup>6</sup> race,<sup>7</sup> or gender in how later judges substantively treat earlier opinions: judges will follow, say, the canonical case rejecting a claim of ineffective assistance of counsel, and that canonical case will not have elements more likely to appeal to a later judge of the same party, race, or gender as the opinion author. The idea is that when judges choose which opinions to follow, party, race, and gender are irrelevant. On this account, judges are not influenced by the party, race, or gender of the authors whose opinions they follow (and may not even notice the party, race, or gender of the earlier author), and nothing in the earlier opinions of a judge of a particular party, race, or gender will be correlated with anything that a later judge might value.<sup>8</sup> If this account is correct, we would not expect to see any differences in substantive treatments correlated with party, race, or gender.

At the other extreme, we might expect pervasive differences related to party, race, and gender in how later judges substantively treat earlier opinions. The idea is that judges can choose among different opinions on ineffective assistance of counsel (to stick with the example), and they will tend to follow opinions written by judges with whom they share a party, race, or gender, because

---

6. We use the term partisanship simply to refer to political parties, not in the more informal sense of particularly strong support for a party or cause.

7. Our analyses of racial differences focus on Hispanic, Black, and non-Hispanic White judges because of the small number of judges in our data who self-identify with other racial/ethnic groups. We refer to Hispanic rather than Latinx judges because we are using the Federal Judicial Center's definitions. Relatedly, we use the gender binary "female" and "male" in referring to judges because the Federal Judicial Center uses only those categories for gender and there are no known transgender federal circuit judges.

We follow the most common conventions in judicial behavior studies in referring to judges appointed by Republican (Democratic) Presidents as Republican (Democratic) "appointees," because some judges may not be members of the President's party. But when discussing race and gender we use the term "judges," because the Federal Judicial Center data rely on judges' self-identified race and gender and we have no reason to doubt that self-identification. See *infra* text accompanying note 96; CASS R. SUNSTEIN, DAVID SCHKADE, LISA M. ELLMAN & ANDRES SAWICKI, ARE JUDGES POLITICAL?: AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY 5-14 (2006).

8. Note that this second proposition is entailed in this hypothesis of no meaningful differences. If, for example, a Republican appointee preferred to follow a substantively conservative opinion, then we might expect that Republican appointee to be more likely to follow an opinion written by another Republican even if the later judge was unaware of the earlier judge's party, on the theory that opinions written by Republicans are, on average, more likely to be substantively conservative than are opinions written by Democrats.

the language and reasoning in the opinions they follow are closer to the later judges' preferences. As with the possibility that there are no differences based on party, race, or gender, it need not be that judges pay attention to these characteristics of the judges whose opinions they follow. The later judge may simply find the reasoning of a particular earlier opinion attractive for reasons correlated with these characteristics and thus be more likely to choose to follow an opinion written by a judge of the same party, race, or gender even if the judge ignores the identity of the author of the earlier opinion. So a later judge's awareness of the party, race, or gender of an opinion author is not necessary to motivate this account. But for this account to be plausible without such awareness, there must be some element of an opinion correlated with party, race, or gender (such as ideology) such that a later judge is likely to prefer to follow an opinion written by a judge of the same party, race, or gender. If this account is correct, then we would expect to see pervasive differences in substantive treatments correlated with party, race, and gender: judges will be more likely to (consciously or unconsciously) follow opinions written by those of the same party, race, or gender.

Our findings are surprising – and more nuanced than either of these accounts would suggest. Pooling data over the full time span of our data (1974–2017), we find statistically significant, but substantively small, partisan differences and intraparty racial and gender differences in authoring judges' positive treatments of earlier opinions.<sup>9</sup> But looking at the pooled data masks a dramatic development that is the real story: Far from being stable over time, the differences we find are quite small in the early years of our study and rise dramatically over time, becoming large and thus substantively meaningful.<sup>10</sup> Further, the differences are greatest for

---

9. These findings are for opinions that follow an earlier opinion and thus treat that earlier opinion as controlling. See *infra* notes 30–35 and accompanying text. We do not find substantively meaningful or statistically significant differences in negative treatments of opinions, which likely reflects the smaller number of such negative treatments (especially given our controls for circuit and year). See *infra* text accompanying note 101.

10. Statistical significance refers to the ability to reject a particular null hypothesis (typically of no difference or no effect) regardless of the size of the difference or effect. Substantive significance refers to an estimate that is large enough to be of scientific or policy interest. With a large enough sample, minuscule differences or effects can be statistically significant, even if they are of no scientific or policy interest. See STEPHEN T. ZILIAK & DEIRDRE N. MCCLOSKEY, *THE CULT OF STATISTICAL SIGNIFICANCE: HOW THE STANDARD ERROR COSTS US JOBS, JUSTICE, AND LIVES* 31–32 (2008) (drawing this distinction); Richard Lempert, *The*

the most ideologically charged categories of cases. Put differently, we do not see the sort of pervasive partisan, racial, and gender differences over the full time period that would exist if judges were consistently influenced by these factors. Instead, we see a sharp rise in partisan differences, and we see a rise in racial and gender differences within parties (particularly for Republicans). These differences are most dramatic in cases with the most ideological salience. The fact that these racial and gender differences occur within parties highlights that these differences are not the result of statistical associations between partisanship and race or between partisanship and gender. To pick the clearest example, Black and White co-partisan judges treat opinions by Black co-partisans differently in ways that are not only statistically significant but also large and therefore substantively meaningful.

A closer look at our results reveals that the partisan differences in substantive treatments of opinions written by Democratic appointees are larger than the differences in treatments of opinions by Republican appointees. What could explain greater partisan differences in treatments of Democratic opinions than in treatments of Republican opinions? The best explanation involves an accelerating shift among Republican appointees in a conservative direction compared to a steady shift among Democratic appointees in a liberal direction. Such a pattern produces results strikingly similar to what our data show. To be clear, we cannot prove this explanation, but we think it is the most likely one.

Finally, the race and gender findings (which, again, are within party) are particularly interesting, because our study of substantive treatments finds significant Black-White differences and significant Hispanic-White and female-male differences among Republicans, whereas studies of voting have found relatively few significant differences across judges of different races and genders.<sup>11</sup> As with partisanship, these differences have risen in recent years and apply to our broad category of ideologically salient cases (not just the

---

*Significance of Statistical Significance: Two Authors Restate an Incontrovertible Caution. Why a Book?*, 34 L. & SOC. INQUIRY 225, 227 n.4 (2009).

We do not use the terms “substantive significance” or “substantively significant” in this Article. The reason is to avoid confusion with the similar sounding but distinct “statistical significance” and “statistically significant.” In place of “substantively significant” we use “substantively meaningful.”

11. See *infra* notes 51–66 and accompanying text.

subset of ideologically salient cases that are race-salient and gender-salient).

We think the best explanation of these racial and gender differences is that they are capturing an element of ideology that partisanship does not capture. Our results are not consistent with a desire of judges of a particular race or gender to enhance the status of those of the same race or gender, because that would not explain the rise over time. The temporally increasing racial and gender differences in our data are similar to the increasing partisan differences we find. Both increases align with the widely documented rise in polarization among U.S. elected officials and within U.S. society more broadly.<sup>12</sup> There is evidence that just as a partisan affiliation reveals information about political attitudes and ideology,<sup>13</sup> race and gender are associated with political attitudes and ideology.<sup>14</sup> In light of these correlations, intraparty race and gender provide a finer-grained proxy for a judge's ideological leanings. Consequently, the fact that we see Hispanic-White and female-male differences only within Republican judges provides some support for the proposition that White male Republican appointees are the central contributors to the accelerating rightward move among Republican appointees that we find.<sup>15</sup>

---

12. See Alan I. Abramowitz & Kyle L. Saunders, *Is Polarization a Myth?*, 70 J. POL. 542, 546–47 (2008); ALAN I. ABRAMOWITZ, *Partisan-Ideological Polarization*, in THE DISAPPEARING CENTER: ENGAGED CITIZENS, POLARIZATION, AND AMERICAN DEMOCRACY 34, 37, 39–40 (2010); Drew DeSilver, *The Polarization in Today's Congress Has Roots that Go Back Decades*, PEW RSCH. CTR. (Mar. 10, 2022), <https://www.pewresearch.org/fact-tank/2022/03/10/the-polarization-in-todays-congress-has-roots-that-go-back-decades>; MICHAEL DIMOCK, JOCELYN KILEY, SCOTT KEETER & CARROLL DOHERTY, *POLITICAL POLARIZATION IN THE AMERICAN PUBLIC: HOW INCREASING IDEOLOGICAL UNIFORMITY AND PARTISAN ANTI-PATHY AFFECT POLITICS, COMPROMISE AND EVERYDAY LIFE* 19–20, 24 (2014), <https://www.pewresearch.org/politics/2014/06/12/political-polarization-in-the-american-public>; Cynthia R. Farina, *Congressional Polarization: Terminal Constitutional Dysfunction?*, 115 COLUM. L. REV. 1689, 1697 (2015); Jeff Lewis, *Polarization in Congress*, VOTEVIEW.COM (Jan. 20, 2022), [https://voteview.com/articles/party\\_polarization](https://voteview.com/articles/party_polarization); see also Jennifer McCoy & Benjamin Press, *What Happens When Democracies Become Perniciously Polarized?*, CARNEGIE ENDOWMENT INT'L PEACE (Jan. 18, 2022), <https://carnegieendowment.org/2022/01/18/what-happens-when-democracies-become-perniciously-polarized-pub-86190>.

13. See, e.g., DIMOCK ET AL., *supra* note 12 at 16.

14. See, e.g., Vincent L. Hutchings & Nicholas A. Valentino, *The Centrality of Race in American Politics*, 7 ANN. REV. POL. SCI. 383, 401 (2004); *Gender Gap Public Opinion*, CTR. FOR AM. WOMEN & POL., <https://cawp.rutgers.edu/gender-gap-public-opinion> (last visited Oct. 3, 2023).

15. This would comport with some suggestive evidence regarding the preconfirmation political ideology of district court judges (as opposed to their judicial behavior) indicating



Part of what is striking about the racial and gender differences in substantive treatments is that prior research on merits votes has found differences in voting behavior between female and male judges, or judges of different races, only in limited subsets of cases. By contrast, the intraparty race and gender differences we find suggest the existence of subtle ideological differences operating outside of race- or gender-salient cases—differences that previous studies have failed to identify.

Our results do not support the proposition that judges have always been pervasively influenced by party, race, or gender. Instead, our results provide evidence of increasing partisan, racial, and gender differences among judges in recent years that reflect some combination of greater ideological differences and a greater willingness to let ideological differences influence how opinions are written.<sup>16</sup> Judges have some insulation from the increasing ideological polarization in the country, but that insulation goes only so far.

The rest of the Article proceeds as follows. In Part I, we discuss the importance of majority opinions' substantive treatments of earlier opinions. Substantive treatments are not mere citations. Following an opinion means relying on it as controlling authority. *Shepard's* is the most studied and accepted source of substantive treatments, and we rely on it here. Part II discusses the empirical literature on differences in judicial behavior, which has focused on party, gender, and race. That literature has focused mainly on judges' votes and found partisan differences in some ideologically salient case categories, racial differences in a few case categories, and gender differences primarily in sex discrimination cases. In Part III we lay out our research questions. We focus on partisan differences as well as intraparty differences with respect to race and gender. In Part IV we present our data and research design, which are unique within the literature. Part V presents our primary results. In Part VI we discuss our findings and some of the

---

that Black, Hispanic, and female Republican district judges were more liberal (as measured by campaign contributions) than their White male Republican counterparts. See Maya Sen, *Diversity, Qualifications, and Ideology: How Female and Minority Judges Have Changed, or Not Changed, over Time*, 2017 WIS. L. REV. 367, 394 tbl.4 (2017) (finding preconfirmation ideological differences among district court judges within party based on race and gender).

16. For a similar finding in terms of how Supreme Court justices approach oral argument, see Tonja Jacobi & Matthew Sag, *The New Oral Argument: Justices as Advocates*, 94 NOTRE DAME L. REV. 1161, 1173–77 (2019) (discussing judicial polarization).

interesting questions they present, such as the greater partisan divergence for treatments of opinions written by Democratic appointees than for those written by Republican appointees—a difference that is consistent with an accelerating conservative shift among Republican appointees as opposed to a steady liberal shift among Democratic appointees. A brief conclusion follows.

### I. THE IMPORTANCE OF TREATMENTS OF EARLIER OPINIONS

The vast majority of the scholarship on judicial behavior focuses on merits votes.<sup>17</sup> Federal appellate judges' votes to affirm or reverse a lower court's decision are obviously important. Which party prevailed is the most concrete outcome of an appeal, and the one that likely matters most to the parties in the case. Judges' votes both help to shape the law and reveal valuable information about the judges' preferences. But votes alone are a fairly crude metric.

At the outset, it bears noting that studies of votes rely on contestable (and contested) ideological coding of how conservative or liberal a given decision is.<sup>18</sup> For example, should *Gonzales v. Raich*<sup>19</sup> (upholding Congress's authority to criminalize marijuana production notwithstanding a state law allowing it) be coded as

---

17. See *infra* notes 44–46, 51–68 and accompanying text.

18. For arguments against the reliability of coding, see Hon. Harry T. Edwards & Michael A. Livermore, *Pitfalls of Empirical Studies that Attempt to Understand the Factors Affecting Appellate Decisionmaking*, 58 DUKE L.J. 1895, 1925 (2009):

[I]t is very difficult to characterize many case outcomes. For example, the general rights embraced by freedom of religion and freedom of expression sometimes conflict with the exercise of other rights; it may not be clear how presumed liberal or conservative judges should be expected to vote in such cases. Cases may be disposed of on procedural grounds that are essentially nonideological, leading to coding errors when the outcome must be coded as liberal or conservative. A court's interpretation of a statute may defy ideological description (*e.g.*, rate allocations in a matter before the Federal Energy Regulatory Commission, where the parties before the court are competing companies) . . . [M]any appeals involve multiple, complex issues, thus making it impossible to describe the appellate court's disposition as liberal or conservative.

See also Carolyn Shapiro, *Coding Complexity: Bringing Law to the Empirical Analysis of the Supreme Court*, 60 HASTINGS L.J. 477, 480–81 (2009) (criticizing ideological coding of cases); Anna Harvey & Michael J. Woodruff, *Confirmation Bias in the United States Supreme Court Judicial Database*, 29 J.L. ECON. & ORG. 414, 415–21 (2013) (finding, as the title suggests, confirmation bias in the ideological coding of cases).

For arguments in favor of the reliability of coding, see, for example, Tracey E. George, *Developing a Positive Theory of Decisionmaking on U.S. Courts of Appeals*, 58 OHIO STATE L.J. 1635, 1673 n.129 (1998) (arguing in favor of ideological coding).

19. *Gonzales v. Raich*, 545 U.S. 1, 74 (2005).

conservative, because it upheld drug laws, or liberal, because it upheld congressional authority under the Commerce Clause?<sup>20</sup>

Even if a coding scheme could deal with cases like *Raich* in a principled way, a judge's understanding of what counts as a conservative or liberal outcome is almost certainly highly timebound. What was regarded as a conservative decision in 1975 is likely different from what was regarded as conservative in 2015.<sup>21</sup> This will create difficulties for studies that attempt to analyze the conservatism or liberalism of merits votes over time.

But there is a deeper problem with focusing on judges' votes: the law is shaped by reasoning and by precedent—courts' treatments of earlier cases. The most important part of a given majority opinion's reasoning is its articulation of the test or factors that lead the court to decide as it does. Lawyers and judges interpreting a given opinion will look first to the court's articulation of its holding. But other aspects of opinions are significant, and revealing, even though they are not as important as the holding. One of the other significant aspects of an opinion is its treatment of earlier cases. The treatment of earlier cases helps to shape the law and concomitantly helps to reveal judges' preferences.

An opinion's treatment of earlier cases is important for the law's development in two related ways. First, the treatment of earlier cases is an important element of an opinion's reasoning. Opinions follow the precedents they deem controlling, and overrule, question, criticize, limit, or distinguish the opinions they deem not controlling or poorly reasoned. In a common law system, treatments of earlier cases are the building blocks for the substance of new opinions. Second, and relatedly, the substantive treatments of earlier opinions help shape legal doctrines. If later opinions repeatedly criticize or question a given opinion, a lawyer would be foolish to blithely rely on that case. Conversely, the more a given

---

20. See Ernest A. Young, *Just Blowing Smoke? Politics, Doctrine, and the Federalist Revival After Gonzales v. Raich*, 2005 SUP. CT. REV. 1, 11-14 (discussing the difficulties in coding *Raich*, and arguing that ideological coding of case outcomes is fraught with difficulties and that "the different factors used to code a case as 'conservative' or 'liberal' may cut in different directions within the confines of a single case").

21. To foreshadow one of our key results, such a change in perception as to what counts as conservative or liberal is consistent with our data and the associated explanation we put forth *infra* in Section VI.A.

case is followed, the stronger its precedential authority becomes.<sup>22</sup> If no court follows the reasoning of a given opinion, then that opinion does not shape the law. Thus, when a court follows a given opinion or diminishes its significance, that treatment not only constitutes part of the court's reasoning but also sends a signal to future courts about which opinions merit following and which do not.

This leads to an expectation that judges who are hostile to an earlier opinion on ideological (or other) grounds will be less likely to explicitly follow that opinion, and more likely to diminish its significance, and that the opposite will be true for judges who are supportive of an earlier opinion. Other aspects of opinion writing might be more important to judges than the treatment of earlier cases, but judges of course understand the significance of their treatments of earlier cases in shaping the law. So when it comes to responding to earlier cases, we would expect judges to be more likely to cast doubt on the opinions to which they are hostile and to treat as guiding precedents the opinions with which they agree.

Votes are not only significant but also more readily identifiable than are treatments of earlier cases. Insofar as judges want to move the law in their preferred ideological direction without attracting much notice (e.g., to avoid other judges on the panel or in the circuit objecting), we might expect the lower profile decision of how to treat a case to vary more with judge ideology than the higher profile decision of how to vote in a case. But the opposite seems at least as likely—that judges' greater focus on votes than on treatment of earlier cases will lead them to vote ideologically more than they treat earlier cases ideologically. The larger point is that, insofar as judicial behavior is correlated with ideology, we would expect that

---

22. See THOMAS G. HANSFORD & JAMES F. SPRIGGS II, *THE POLITICS OF PRECEDENT ON THE U.S. SUPREME COURT* 42–49 (2006), which develops a theory of how Supreme Court justices develop doctrine that measures opinions' positive and negative treatments and then empirically tests that theory. A key part of the Hansford and Spriggs theory is that a precedent has an amount of "vitality," or legal authority, that varies over time. Positive treatments of a precedent increase the vitality of that case while negative citations diminish vitality. According to Hansford and Spriggs, the ideological predispositions of the justices are moderated by the vitality of relevant precedents. A justice who prefers an outcome that is inconsistent with a vital key precedent will find it difficult to reach that outcome until the legal authority of the case has been chipped away over time via negative treatments. *See also* text accompanying note 48.

ideology would be associated with both voting and the treatment of earlier cases.

This raises the question of how to measure judicial treatment of earlier opinions. Citations are an obvious source. The most extensive study of judges' responses to their colleagues' opinions focuses on citations, on the theory that opinions cite to cases the author thinks are important.<sup>23</sup> But citations alone are a crude measure because they do not capture the nature of a court's treatment of an earlier opinion. Some citations occur in the context of a court relying on, and thus following, an earlier opinion. Some citations are negative (e.g., criticizing or questioning a precedent). And some may not be significant. A bare citation in a string of citations with no accompanying discussion of the cases does not provide much information.

To classify and measure the substantive treatment of earlier opinions in each majority opinion, we rely on *Shepard's*, an approach that has become standard in the literature.<sup>24</sup> *Shepard's* is a

---

23. See Stephen J. Choi & G. Mitu Gulati, *Bias in Judicial Citations: A Window into the Behavior of Judges*, 37 J. LEGAL STUD. 87, 94 (2008).

24. See Ryan C. Black & James F. Spriggs II, *An Empirical Analysis of the Length of U.S. Supreme Court Opinions*, 45 HOUS. L. REV. 621, 669-76 (2008); Chad Westerland, Jeffrey A. Segal, Lee Epstein, Charles M. Cameron & Scott Comparato, *Strategic Defiance and Compliance in the U.S. Courts of Appeals*, 54 AM. J. POL. SCI. 891, 896-98 (2010); James F. Spriggs II & Thomas G. Hansford, *Explaining the Overruling of U.S. Supreme Court Precedent*, 63 J. POL. 1091, 1097, 1100 (2001) [hereinafter Spriggs & Hansford, *Explaining the Overruling*]; Pamela C. Corley & Justin Wedeking, *The (Dis)Advantage of Certainty: The Importance of Certainty in Language*, 48 L. & SOC. REV. 35, 41-42, 47-48 (2014); HANSFORD & SPRIGGS, *supra* note 22, at 43-62; James F. Spriggs II & Thomas G. Hansford, *The U.S. Supreme Court's Incorporation and Interpretation of Precedent*, 36 L. & SOC. REV. 139, 146-47 (2002) [hereinafter Spriggs & Hansford, *U.S. Supreme Court's Incorporations*]; Christina L. Boyd & James F. Spriggs II, *An Examination of Strategic Anticipation of Appellate Court Preferences by Federal District Court Judges*, 29 WASH. U. J. L. & POL'Y 37, 63-67 (2009); Ryan C. Black & James F. Spriggs II, *The Citation and Depreciation of U.S. Supreme Court Precedent*, 10 J. EMPIRICAL LEGAL STUD. 325, 335 (2013); Frank B. Cross, *The Ideology of Supreme Court Opinions and Citations*, 97 IOWA L. REV. 693, 723-24 (2012); Stuart Minor Benjamin & Georg Vanberg, *Judicial Retirements and the Staying Power of U.S. Supreme Court Decisions*, 13 J. EMPIRICAL LEGAL STUD. 5, 10-11 (2016); Jeffrey Budziak, *The Effect of Visiting Judges on the Treatment of Legal Policy in the U.S. Courts of Appeals*, 38 JUST. SYS. J. 348, 349 (2017); Thomas G. Hansford, James F. Spriggs II & Anthony A. Stenger, *The Information Dynamics of Vertical Stare Decisis*, 75 J. POL. 894, 898 (2013); Lee Epstein, William M. Landes & Adam Liptak, *The Decision to Depart (or Not) from Constitutional Precedent: An Empirical Study of the Roberts Court*, 90 N.Y.U. L. REV. 1115, 1133 (2015); Ali S. Masood & Benjamin J. Kassow, *The Sum of its Parts: How Supreme Court Justices Disparately Shape Attention to Their Opinions*, 101 SOC. SCI. Q. 842, 853 (2020); Sara C. Benesh & Malia Reddick, *Overruled: An Event History Analysis of Lower Court Reaction to Supreme Court Alteration of Precedent*, 64 J. POL. 534, 540-41 (2002); Matthew P. Hitt, *Measuring Precedent in a Judicial Hierarchy*, 50 L. & SOC'Y REV. 57, 66-67 (2016); Michael P. Fix, Justin T. Kingsland &

widely used commercial legal research service that employs attorneys to examine every state and federal court opinion and code the content of every citation within each opinion. Sometimes a judge will cite an opinion in a way that seems to give no useful information about the citing judge's substantive treatment of that opinion (for instance, a string of citations with no discussion). If a citation refers to a case but has no meaningful substantive reaction to it, *Shepard's* does not put the citation into a substantive category. Such bare citations constitute the majority of citations.<sup>25</sup> *Shepard's* classifies citations that are accompanied by a substantive treatment of an opinion (i.e., a discussion of and substantive response to an opinion rather than a mere mention of it) into the following main categories: overruled, questioned, limited, criticized, distinguished, explained, harmonized, paralleled, and followed.<sup>26</sup> *Shepard's* characterizes overruled, questioned, limited, criticized, and

---

Matthew D. Montgomery, *The Complexities of State Court Compliance with U.S. Supreme Court Precedent*, 38 JUST. SYS. J. 149, 155-56 (2017); Scott D. McClurg & Scott A. Comparato, *Rebellious or Just Misunderstood?: Assessing Measures of Lower Court Compliance with U.S. Supreme Court Precedent*, RESEARCHGATE, 12-17 (Jan. 2004) (unpublished manuscript), ([https://www.researchgate.net/profile/Scott-Mcclurg/publication/228504744\\_Rebellious\\_or\\_Just\\_Misunderstood\\_Assessing\\_Measures\\_of\\_Lower\\_Court\\_Compliance\\_with\\_US\\_Supreme\\_Court\\_Precedent/links/0912f50f41e34e92d5000000/Rebellious-or-Just-Misunderstood-Assessing-Measures-of-Lower-Court-Compliance-with-US-Supreme-Court-Precedent.pdf](https://www.researchgate.net/profile/Scott-Mcclurg/publication/228504744_Rebellious_or_Just_Misunderstood_Assessing_Measures_of_Lower_Court_Compliance_with_US_Supreme_Court_Precedent/links/0912f50f41e34e92d5000000/Rebellious-or-Just-Misunderstood-Assessing-Measures-of-Lower-Court-Compliance-with-US-Supreme-Court-Precedent.pdf)); Neal Devins & David Klein, *The Vanishing Common Law Judge?*, 165 U. PA. L. REV. 595, 605 (2017); Michael J. Nelson & Rachael K. Hinkle, *Crafting the Law: How Opinion Content Influences Legal Development*, 39 JUST. SYS. J. 97, 103-05 (2018); Joshua Boston, *Strategic Opinion Language on the U.S. Courts of Appeals*, 8 J.L. & CTS. 1, 8 (2020); Rachael K. Hinkle, *Strategic Anticipation of En Banc Review in the U.S. Courts of Appeals*, 50 L. & SOC'Y REV. 383, 393-94, n.3 (2016); Linda L. Berger & Eric C. Nystrom, *"Remarkable Influence": The Unexpected Importance of Justice Scalia's Deceptively Unanimous and Contested Majority Opinions*, 20 J. APP. PRAC. PROCESS 233, 252-54 (2019); Benjamin Kassow, *The Impact of Ideology and Attorneys on Precedent Usage: An Analysis of State High Courts 99-101* (Jan. 1, 2013) (Ph.D. dissertation, University of South Carolina), (<https://scholarcommons.sc.edu/etd/2571>); Robert C. Wigton, *What Does It Take to Overrule? An Analysis of Supreme Court Overrulings and the Doctrine of Stare Decisis*, 18 LEGAL STUD. F. 3, 4 (1994); Michael C. Gizzi & R. Craig Curtis, *The Impact of Arizona v. Gant on Search and Seizure Law as Applied to Vehicle Searches*, 1 U. DENV. CRIM. L. REV. 30, 40-41 (2011); Rachael K. Hinkle, *Panel Effects and Opinion Crafting in the U.S. Courts of Appeals*, 5 J.L. & CTS. 313, 323 (2017) [hereinafter Hinkle, *Panel Effects*]; Pamela C. Corley, *Uncertain Precedent: Circuit Court Responses to Supreme Court Plurality Opinions*, 37 AM. POL. RSCH. 30, 36-37 (2009); Stuart Minor Benjamin & Bruce A. Desmarais, *Standing the Test of Time: The Breadth of Majority Coalitions and the Fate of U.S. Supreme Court Precedents*, 4 J. LEGAL ANALYSIS 445, 451-53 (2012).

25. See James F. Spriggs II & Thomas G. Hansford, *Measuring Legal Change: The Reliability and Validity of Shepard's Citations*, 53 POL. RSCH. Q. 327, 329 n.2 (2000).

26. *Shepard's* identifies these treatments as "overruled by," "questioned by," etc. In this Article we generally drop the "by" simply to avoid wordiness.

distinguished as negative treatments. More than 99.99% of the positive treatments in our dataset are followed, so in this Article we treat the following of an opinion and a positive treatment as synonymous and use the terms interchangeably.<sup>27</sup> Explained and harmonized are neutral treatments. They are rare and, as neutral treatments, are not indicative of a positive or negative response. Because we want to focus on substantive responses to other appellate opinions—that is, circuit court treatments that indicate some level of support or non-support—neutral responses are not of interest.<sup>28</sup>

The reliability of *Shepard's* treatments has been rigorously studied. Most notably, James Spriggs and Thomas Hansford undertook a careful study to measure the reliability of *Shepard's*. They took a stratified random sample of Supreme Court opinions citing earlier Supreme Court cases, yielding 602 citing opinions, and they coded all the citing opinions according to the coding rules in the *Shepard's* training manual. They found high levels of agreement between their coding and *Shepard's* coding.<sup>29</sup>

---

27. There are no instances of “paralleled” in our dataset. There are two other positive treatments, but they are quite rare, constituting less than .01% of the positive treatments: “extended by” (10 of the 648,226 treatments in our dataset) and “valid by” (1 of the treatments in our dataset).

Because “follow” is a verb, we generally use “follow” for the verb form and “positive treatment” for the noun form. Nothing substantive turns on this difference in terminology—we use it simply for clarity and ease of exposition.

28. *Shepard's* does not code the content of dissents, for good reason: by definition, substantive discussions in dissents do not represent the views of the majority and thus are not precedential. *Shepard's* does code concurrences, but it codes their treatments as neutral. See, e.g., LEXISNEXIS, *SHEPARD'S EDITORIAL PHRASES—ALPHABETICAL LIST* 33, <http://www.lexisnexis.com/pdf/lexis-advance/Shepards-Editorial-Phrases-Alphabetical-List.pdf> (last visited Oct. 2, 2023) (listing “Criticized” in a majority opinion as a negative treatment, but “Criticized in Concurring Opinion” as a neutral treatment, noting for the concurrence that it “may not have the authority to materially affect its precedential value”). Again, by definition the substantive discussions in concurrences are not part of the majority opinion. That said, on some occasions a concurrence by a judge necessary to form the majority may contain an influential substantive discussion of an earlier case. *Shepard's* does not attempt to determine which concurrences may have some force and thus arguably might merit designation as positive or negative, for the apparent reason that such determinations are highly debatable. In this way, *Shepard's* may not code as positive or negative some treatments in concurrences that arguably are at least mildly positive or negative. We have no reason to believe that the absence of such information biases *Shepard's* results, and no other studies have so suggested or found. See Benjamin & Desmarais, *supra* note 24, at 7.

29. See Spriggs & Hansford, *supra* note 25, at 333–34.

There are strong reasons to believe that *Shepard's* treatments are valid measures. *Shepard's* definitions comport with judges' and lawyers' understanding of the substantive treatments. The specification of "followed" is illustrative. *Shepard's* defines "followed" as "[t]he citing opinion relies on the case you are Shepardizing as controlling or persuasive authority."<sup>30</sup> *Shepard's* created a training manual for the lawyers who code citations, with thirteen single-spaced pages devoted to laying out detailed coding rules for the treatment categories. According to the manual, "followed" (which the manual denotes with an "f") entails a case the citing opinion "relied on as controlling authority. The majority opinion in the [citing case] has expressly relied on the cited case as precedent on which to base its decision. The citing opinion must in some firm way refer to the cited case as compelling precedent."<sup>31</sup> The manual adds that "[a] mere 'going-along' with the cited case would not be sufficient for assigning a letter 'f.' Merely citing or quoting, with nothing more, is not a sufficient expression of reliance to permit an 'f' (or any other letter, for that matter)."<sup>32</sup> The manual identifies the following as language meriting a "followed" designation: "We affirm on the authority of . . . , or on the teaching of . . . , or for the reasons stated in . . . or under the rationale of . . . ; [or] such a conclusion is required by . . . or governed by . . . ."<sup>33</sup> This definition and discussion capture lawyers' and judges' understanding of what it means to follow a case.<sup>34</sup> This is not surprising, given the large amounts of money that lawyers have paid for access to *Shepard's*. It has long been widely used by practicing attorneys and judges, indicating that legal professionals view *Shepard's* as providing legally relevant information.<sup>35</sup>

It bears noting that positive treatment is much more common than negative treatment. Indeed, in our dataset there are 451,277 followed treatments and 141,768 negative treatments. The reason for this difference seems reasonably straightforward. For *Shepard's*

---

30. See LEXISNEXIS, HOW TO SHEPARDIZE: YOUR GUIDE TO LEGAL RESEARCH USING SHEPARD'S CITATIONS 10 (on file with authors).

31. See SHEPARD'S COMPANY, SHEPARD'S CITATIONS IN-HOUSE TRAINING MANUAL 13 (1993) (unpublished manual) (on file with authors).

32. *Id.*

33. *Id.*

34. See Benjamin & Vanberg, *supra* note 24, at 14.

35. See, e.g., J. MYRON JACOBSTEIN & ROY M. MERSKY, FUNDAMENTALS OF LEGAL RESEARCH, 261-89 (1977).



to assign a negative treatment to a particular discussion in an opinion, that opinion must be explicit in its criticism, questioning, etc. of the earlier opinion. Such explicit negativity about an earlier opinion is fairly aggressive and might be perceived by other judges as uncollegial. No circuit judge wants her own opinion to be treated negatively by her colleagues in the future, and she might concomitantly be hesitant to treat her colleagues' earlier opinions negatively. Negative treatments are thus costly and relatively unusual. Following an earlier opinion has all the opposite attributes. As we noted above, it bolsters the opinion.<sup>36</sup> *Shepard's* characterizes following as the positive treatment for a reason—in our precedential system, following an opinion is the central form of praise. So the costs of following an opinion are quite low. The only disadvantage of following an earlier opinion for a later panel would arise if the later panel did not in fact want to provide support for the earlier opinion. If the later panel found the earlier opinion objectionable, then, and only then, would it have an incentive to avoid following that earlier opinion. Indeed, this last point highlights why examining positive treatments can be so revealing: we would expect judges to be more inclined to follow opinions with which they agree and less inclined to follow opinions with which they do not.<sup>37</sup>

Are these *Shepard's* treatments reflective of judges' choices? There are two possible ways in which an opinion's discussion of a previous opinion might not reflect a judge's meaningful decision. One is that the judge may effectively have no choice in the matter. Most obviously, if there is only one precedent that directly controls the question at issue, then we would expect (or at least hope) that any judge would follow that precedent. Insofar as existing precedents constrain judges, one element of that constraint is that there are some cases that can have only one possible result, because that is what precedent demands. After *Roe v. Wade*<sup>38</sup> (and before

---

36. See *supra* note 22 and accompanying text.

37. In this Article we often use the word "citing" to refer to *Shepard's* treatments for the sake of streamlining some sentences. But, to be clear, what we are studying is *Shepard's* substantive treatments, and, as we have just discussed, those treatments are much more than a mere citation. So any references to citations in our data are referring to substantive treatments, and we use the terms "citing" and "treating" interchangeably.

38. *Roe v. Wade*, 410 U.S. 113 (1973).

*Dobbs v. Jackson Women's Health Organization*<sup>39</sup>), for example, any lower court would be compelled by *Roe* to invalidate a statute that criminalized all abortions.

But many cases are not so clearly constrained by binding legal precedent. To return to the abortion example, a flat ban on abortion was foreclosed by *Roe*, but restrictions on some abortions were subject to differing interpretations (as has been the case when courts of appeals have reviewed post-*Roe* abortion restrictions).

A second limit to the force of the argument may be more significant: there will rarely be only one precedent on which a court can rely. Consider a question like the standard applied for issuing a preliminary injunction or summary judgment. There are thousands of cases laying out a standard, and they often differ, even if only slightly, in the wording they use. Some wording is slightly more favorable to those seeking the injunction or summary judgment, and some is slightly less favorable. Judges can choose among them when deciding which case to follow in the articulation and application of standards for an injunction or summary judgment.

And even in situations where there is only one precedent directly on point for a given case, as soon as that case is decided there will be two cases that are directly on point. So the next panel confronting the same issue will have a choice among two relevant precedents, and the panel after that will have three, and so on. And given that each of those opinions will differ slightly from the others and will have different authors, judges will be able to make some choices in determining which opinions they follow.

The second possible way in which opinions' discussion of cases might not reflect meaningful decisions by judges is that judges may leave those decisions to their clerks. Insofar as judges defer to their clerks (or anyone else) in their opinions' discussions of earlier cases, the judges are not making the meaningful decisions and we should not expect to see the differences in treatment behavior that we hypothesize.

The intuition behind this second possibility is that judges often rely on their clerks for the first draft of a majority opinion and in particular may rely on their clerks for matters as mundane as which cases to cite. We think that this intuition likely has particular force with respect to string cites in which an opinion states a basic legal

---

39. *Dobbs v. Jackson Women's Health Org.*, 142 S. Ct. 2228 (2022).

rule and then lists several cases that support that legal rule. It seems quite unlikely that judges choose (or even focus on) every case that they list in a string cite for a straightforward proposition. But, as we noted above, *Shepard's* does not code such citations as positive treatments. For *Shepard's* to code something as a treatment, the opinion must have a meaningful and significant discussion of the case—a bare citation does not count. As we also noted, in our precedent-based system discussions of earlier opinions, and decisions of which opinions to follow (and which to cast doubt on), are important elements of a given opinion. Judges are much more likely to make choices about crafting those elements than they are about what cases to list in a string cite.

But we recognize that judges may well defer to their clerks' choice of which cases to rely on (and thus follow for *Shepard's* purposes) and which cases to diminish (and thus overrule, criticize, question, distinguish, or limit for *Shepard's* purposes). Indeed, some judges in some opinions may well defer to their clerks on all aspects of an opinion.

This raises an important possible dampening effect. The possibilities of a single directly relevant precedent and the effect of clerks will tend to diminish the differences we are studying in this Article: the less that opinions reflect judges' choices, the less likely that there will be significant differences in the measures designed to capture those choices. Insofar as we find the differences we hypothesize, we find those differences despite the dampening impact of these possibilities.

We have no reason to believe that either of these possibilities would skew our data. As to clerks, some judges may focus heavily on ideology in choosing their clerks and choose clerks who are ideologically aligned with them, but that of course would be an accurate reflection of the judge's ideology.<sup>40</sup> For other judges, ideology may play no role in their choice of clerks. For those judges, the impact of clerks would be random and thus would mute

---

40. See, e.g., Adam Bonica, Adam S. Chilton, Jacob Goldin, Kyle Rozema & Maya Sen, *Measuring Judicial Ideology Using Law Clerk Hiring*, 19 AM. L. & ECON. REV. 129, 146 (2017) (“[C]lerk ideologies provide a window into the ideology of the hiring judge . . .”); Adam Bonica, Adam S. Chilton, Jacob Goldin, Kyle Rozema & Maya Sen, *The Political Ideologies of Law Clerks*, 19 AM. L. & ECON. REV. 96, 123–24, (2017) (finding evidence that judges hire clerks with similar ideologies).

the impact of judicial partisanship.<sup>41</sup> But what if, say, Republican appointees are more likely than Democratic appointees to choose clerks who are ideologically aligned with them? Under those circumstances, clerks' influence would push Democratic appointees toward the middle of the ideological spectrum while their Republican counterparts would be unchanged. That might affect attempts at measuring the absolute conservatism or liberalism of appellate opinions. But our focus is on the relative distance between Republican and Democratic appointees. We are not measuring whether Democratic or Republican appointees are more liberal or conservative than expected. We are simply measuring the divergence between Democratic and Republican appointees. If Democrats were less likely than Republicans to hire ideologically compatible clerks (or vice versa), this would tend to dampen partisan differences (the fewer the clerks hired for ideological compatibility, the less a clerk effect will lead to ideological differences in how judges treat earlier opinions). Relatedly, if clerks were more likely to be politically moderate than the judges for whom they clerk, that might mute the differences in behavior between Democratic and Republican appointees and thus mute any differences based on the partisanship of the judges.

By contrast, if it were the case that Republican appointees systematically choose clerks who are more conservative than they, and that Democratic appointees systematically choose clerks more liberal than they, then this heterogeneity among clerks could increase the partisan differences we see between Democratic and Republican appointees. But in such circumstances that heterogeneity would be a function of judges' decisions to choose clerks who are more ideologically extreme than they are. That is, the heterogeneity would reflect the judges' ideological disposition to have clerks who are more ideological than the judges themselves are. Differences in clerks would be attributable to the judges who hired them. Similar points apply to the race and gender of clerks. Insofar as some judges are disproportionately likely to hire clerks whose race or gender matches their own, the role of the clerk's race or gender plays a similar role to that of the judge's race or gender. Conversely, insofar

---

41. See Jeremy D. Fogel, Mary S. Hoopes & Goodwin Liu, *Law Clerk Selection and Diversity: Insights from Fifty Sitting Judges of the Federal Courts of Appeals*, 137 HARV. L. REV. (forthcoming Nov. 2023) (finding that "[m]ost judges disclaim any interest in ideological alignment when hiring clerks[,]" though of course this is self-reported).

as other judges hire without regard to race or gender, any impact the clerks have might reduce the measured effect of judges' race and gender on how they treat earlier cases.

As to precedents directly on point, there is also no reason to believe there is any skew that would affect our analysis. We have no reason to believe that there is any difference in the likelihood of Democratic versus Republican, White versus Hispanic versus Black, or female versus male judges to write the sorts of opinions that are likely to be followed. But even if that were true, then presumably Democratic and Republican, White, Hispanic, and Black, and female and male judges would follow those opinions to a similar degree. If, say, Republican appointees are more likely to write opinions that merit being relied on, then presumably that reliance would be across the board and there would be no differences in the likelihood of particular categories of later judges to rely on them. And if the response to that last point is that later Republican appointees (to stick with the example) are more likely than Democratic appointees to find merit in earlier opinions by their Republican colleagues – well, that is exactly what we are trying to measure. That would not be a skewing of our data; it would be a confirmation of our hypothesis.

## II. THE EMPIRICAL LITERATURE ON JUDICIAL BEHAVIOR – PARTY, GENDER, AND RACE

One of the central questions at the intersection of law and political science is to what extent judges' personal characteristics influence their judicial behavior. Can we learn anything from examining the relationship between some personal attributes of judges and what they do on the bench?

Many studies examine how the individual attributes of judges correlate with judges' behavior. Many of these studies focus on differences between Republican and Democratic appointees' judicial behavior, but some address the relationship between race or gender and differences in judicial behavior. These studies generally look directly at individual voting differences and are thus easily interpretable. Our analyses are similar, but we look at differences in majority opinion authors' treatments of earlier opinions.

*A. Political Party*

The most commonly studied characteristic of judges is the political party of the President who most recently nominated a given judge.<sup>42</sup> The underlying theory is that judicial behavior may be influenced by ideology as revealed by the President's party. The key elements of that theory are fairly straightforward:

- Democratic and Republican Presidents diverge ideologically;
- Party and ideology are not perfectly correlated, but there is a strong relationship between the two;
- Presidents choose circuit court nominees with whom they are ideologically compatible;
- Legal doctrine may impose meaningful constraints but often leaves room for judicial decisions that are not determined by legal doctrine and can be influenced by ideology; and
- One of the things judges seek to achieve (indeed, one of the reasons to want to be a judge) is to help move the law in a positive direction, and a given judge's definition of "positive" will be correlated with the judge's ideology, with the result that judges will thus want to push the law in an ideological direction (even though they may conceptualize the direction as "positive" rather than ideological).<sup>43</sup>

In light of the importance of determining the impact of the President's party on judicial behavior, many empirical studies have

---

42. See, e.g., Pauline T. Kim, *Deliberation and Strategy on the United States Courts of Appeals: An Empirical Exploration of Panel Effects*, 157 U. PA. L. REV. 1319, 1327 (2009) (noting the party of the most recent President to nominate a judge is the standard practice for identifying the ideology of a judge); SUNSTEIN ET AL., *supra* note 7, at 5-7 (using the President's party as the measure of judicial ideology); Adam B. Cox & Thomas J. Miles, *Judging the Voting Rights Act*, 108 COLUM. L. REV. 1, 3 (2008) (same); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1718-19 (1997) (same); see also Daniel R. Pinello, *Linking Party to Judicial Ideology in American Courts: A Meta-analysis*, 20 JUST. SYS. J. 219, 221 (1999) (providing "a compendium of empirical undertakings connecting party ID with judicial ideology"). Note that because every President has been a Republican or Democrat, dividing judges into Republican and Democratic categories captures all judges.

43. See, e.g., SUNSTEIN ET AL., *supra* note 7, at 3, 13, 22-27. Some scholars associated with the attitudinalist model argue that legal doctrine poses little or no constraint, and that judges decide cases primarily (and sometimes exclusively) based on their ideology, but one need not subscribe to that view to posit that ideology likely plays *some* role. See, e.g., JEFFREY A. SEGAL & HAROLD J. SPAETH, *THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED* 37-42 (2002).

attempted to measure that impact. Most such studies—indeed, most studies of judicial behavior—focus on judges’ votes as the relevant behavior to be measured. The most extensive study of judicial voting in the U.S. Courts of Appeals was conducted by Cass Sunstein, David Schkade, Lisa Ellman, and Andres Sawicki. They focused on judges’ votes in published opinions in twenty-four issue areas that might be expected to have a partisan valence, such as environmental law, sex discrimination, sexual harassment, disability discrimination, and campaign finance, and found partisan differences in fifteen of them.<sup>44</sup> Other, less comprehensive studies have found differences in judges’ votes in ideologically salient areas like voting rights, affirmative action, and employment discrimination.<sup>45</sup> But other studies have found an absence of differences in some politically charged areas (such as abortion and capital punishment).<sup>46</sup>

Votes are not the only outcome that can be studied. A natural alternative is to examine citation practices. These studies have generally focused on the Supreme Court and have found differences in citation behavior (relying on *Shepard’s*) that comport with differences in judicial voting. For instance, Hansford and Spriggs find that Supreme Court justices are more likely to

---

44. See SUNSTEIN ET AL., *supra* note 7, at 26–27.

45. See Cox & Miles, *supra* note 42, at 48; Jonathan P. Kastellec, *Racial Diversity and Judicial Influence on Appellate Courts*, 57 AM. J. POL. SCI. 167, 173 (2013) (finding that Democratic appointees were more likely than Republican appointees to vote in favor of affirmative action); Sean Farhang & Gregory Wawro, *Institutional Dynamics on the U.S. Court of Appeals: Minority Representation Under Panel Decision Making*, 20 J.L. ECON. & ORG. 299, 314 n.10, 320 (2004) (finding that judges appointed by more conservative Presidents were less likely to find for the plaintiff in employment discrimination cases); see also Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1742–43 (1997) (finding that the party of the appointing President is associated with differences in judges’ votes in environmental cases on the D.C. Circuit).

46. See Thomas J. Miles & Cass R. Sunstein, *The New Legal Realism*, 75 U. CHI. L. REV. 831, 839 (2008). There have been varying findings of partisan differences in religious freedom cases. An early study found partisan differences on voting in Free Exercise claims, particularly where the subjects at issue were politically charged, such as religious accommodations for children in school, with judges appointed by Republican Presidents being more likely to vote in favor of such accommodations. See Gregory C. Sisk, Michael Heise & Andrew P. Morriss, *Searching for the Soul of Judicial Decisionmaking: An Empirical Study of Religious Freedom Decisions*, 65 OHIO STATE L. J. 491, 602 (2004). In contrast, a later study found no statistically significant partisan difference in Free Exercise claims. Both studies, however, found that Republicans were less likely to vote in favor of claimants in Establishment Clause cases. See Sepehr Shahshahani & Lawrence J. Liu, *Religion and Judging on the Federal Courts of Appeals*, 14 J. EMPIRICAL LEGAL STUD. 716, 731–32 (2017).

positively treat ideologically proximate precedents and negatively treat ideologically distant precedents.<sup>47</sup> They go on to show that these citation practices affect the “vitality” of the cited case, with positive treatments increasing vitality and negative treatments diminishing vitality.<sup>48</sup> Chad Westerland et al. find that a lower court’s propensity to positively treat a Supreme Court opinion depends on the ideological distance between the enacting Supreme Court and the current Supreme Court, but not on the ideological distance between the lower court and either the enacting or current Supreme Court.<sup>49</sup> And Frank Cross et al. find that the ideological heterogeneity of the Supreme Court majority coalition is significantly, albeit moderately, associated with citation practices. Specifically, they find that more ideologically heterogeneous Supreme Court majority coalitions cite more opinions than ideologically homogeneous coalitions and the opinions that are cited have greater network centrality.<sup>50</sup>

### *B. Gender and Race*

Partisanship is a significant attribute that might affect judicial behavior, but it is not the only one. After all, parties, and cohorts of judges within those parties, have variation within them. After partisanship, the two most prominent attributes assessed in studies of judicial behavior are judges’ gender and race. These studies raise the possibility that within parties gender and race may shed light on judicial behavior.

#### *1. Gender*

Previous studies of racial or gender differences in judicial behavior have generally focused on judges’ votes in areas of law thought to activate gender and racial identities.<sup>51</sup> In the gender

---

47. See HANSFORD & SPRIGGS, *supra* note 22, at 94.

48. *Id.* For a discussion of vitality, see *supra* note 22.

49. See Westerland et al., *supra* note 24, at 905.

50. Frank B. Cross, James F. Spriggs II, Timothy R. Johnson & Paul J. Wahlbeck, *Citations in the U.S. Supreme Court: An Empirical Study of Their Use and Significance*, 2010 U. ILL. L. REV. 489, 551.

51. The most notable exceptions are two articles by Stephen Burbank and Sean Farhang finding some race and gender effects for class certification decisions and some gender effects for motions to dismiss for failure to state a claim after *Twombly* and *Iqbal*. See Stephen B. Burbank & Sean Farhang, *Politics, Identity, and Class Certification on the U.S. Courts*



context, Christina Boyd, Lee Epstein, and Andrew Martin identified these areas of law as case types “on which female judges may possess valuable expertise, experience, or information” or that are issues of concern to women broadly speaking, such as sex harassment and sexual discrimination.<sup>52</sup> Such cases are often called gender-coded or gender-salient, and we use the latter term.<sup>53</sup> The idea is that in specific kinds of cases, judges may have a particular understanding of and sensitivity to particular issues, perhaps flowing from their expertise and lived experiences.<sup>54</sup>

Studies have found relatively few gender differences in voting. Boyd, Epstein, and Martin identified thirteen gender-salient categories of cases and found gender differences in only one: sex discrimination in employment.<sup>55</sup> Similarly, Susan Haire and Laura Moyer found that “[m]en and women on the bench are quite similar in their voting behavior, with one exception: cases involving sex discrimination.”<sup>56</sup> Sarah Westergren, meanwhile, found that any

*of Appeals*, 119 MICH. L. REV. 231, 231 (2020) [hereinafter Burbank & Farhang, *Class Certification*] (finding that “the presence of one African American on a panel, and the presence of two women (but not one), is associated with procertification outcomes.”); Stephen B. Burbank & Sean Farhang, *Politics, Identity, and Pleading Decisions on the U.S. Courts of Appeals* 169 U. PA. L. REV. 2225, 2226 (2021) [hereinafter Burbank & Farhang, *Pleading Decisions*] (finding that in precedential cases “panels with one woman were more likely to decide precedential other civil rights claims in favor of plaintiffs, and that panels with two women (but not one) were more likely to do so in non-civil rights claims.”).

52. See Christina L. Boyd, Lee Epstein & Andrew D. Martin, *Untangling the Causal Effects of Sex on Judging*, 54 AM. J. POL. SCI. 389, 391 (2010).

53. See *id.* Gender-salient (and race-salient) categories of cases are a subset of the ideologically salient categories we identify. See *infra* note 82 & Appendix A1.

54. There are various theories behind this proposition: that female judges bring a unique knowledge base in key areas like sex discrimination based on their experiences; that men and women think, communicate, and view the world differently from one another; and/or that judges serve as representatives of their group and work to advance their group’s interests. See, e.g., Christina Boyd, *Representation on the Courts? The Effects of Trial Judges’ Sex and Race*, 69 POL. RES. Q. 788, 789–90 (2016); *infra* notes 81–82 and accompanying text.

55. See Boyd, Epstein & Martin, *supra* note 52, at 389 (noting that among the 13 areas of law with gender salience, they observed gender differences only for sex discrimination in employment claims).

56. SUSAN B. HAIRE & LAURA P. MOYER, *DIVERSITY MATTERS: JUDICIAL POLICY MAKING IN THE U.S. COURTS OF APPEALS* 53–54 (2015); see also *id.* at 48 (female judges “tend to decide cases similarly to their male colleagues” with a single exception: “Women judges are more likely to support plaintiffs in sex discrimination cases when compared to the votes of their male colleagues.”). Interestingly, Haire and Moyer found that this difference in votes in sex discrimination cases is a function of age and experience: older cohorts of women and men voted differently in sex discrimination cases, but more recent cohorts did not. See Laura P. Moyer & Susan B. Haire, *Trailblazers and Those that Followed: Personal Experiences, Gender, and*

gender differences in sex discrimination disappeared once she controlled for judges' partisanship.<sup>57</sup> And, strikingly, Jennifer Segal, examining a range of gender-salient cases decided by Clinton-appointed district court judges, found that "there are gender differences in cases involving women's issues, yet it is male judges who are more supportive of these claims."<sup>58</sup>

A few studies have looked beyond gender-salient case types to examine gender differences in areas that are ideological but have no obvious gender salience (such as voting rights and religious liberty), but those studies have not found gender differences in judges' votes in such cases.<sup>59</sup> For example, Sunstein, Schkade, Ellman, and Sawicki found no gender differences in a broad range of ideologically salient cases, and Haire and Moyer aggregated all case types and found no statistically significant gender differences.<sup>60</sup>

---

*Judicial Empathy*, 49 L. & SOC'Y REV. 665, 668 (2015). For a discussion of the hypotheses that flow from a possible trailblazer effect, see *infra* notes 93-94 and accompanying text. It also bears noting that in an earlier, much smaller study, Jennifer Peresie also found panel and judge effects in votes in sexual harassment cases. Jennifer L. Peresie, Note, *Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts*, 114 YALE L.J. 1759, 1776 (2005) ("[I]n Title VII sexual harassment and sex discrimination cases, . . . a judge's gender and the gender composition of the panel mattered to a judge's decision.").

57. See Sarah Westergren, *Gender Effects in the Courts of Appeals Revisited: The Data Since 1994*, 92 GEO. L.J. 689, 703 (2004) (finding no statistically significant effect of gender on judges' votes in sex discrimination cases, and that "any gender effect appears to be intertwined with the effect of political party affiliation of the appointing president"); see also Carol T. Kulik, Elissa L. Perry & Molly B. Pepper, *Here Comes the Judge: The Influence of Judge Personal Characteristics on Federal Sexual Harassment Case Outcomes*, 27 L. & HUM. BEHAV. 69, 80-81 (2003) (finding no gender effect on sexual harassment cases).

58. Jennifer A. Segal, *Representative Decision Making on the Federal Bench: Clinton's District Court Appointees*, 53 POL. RES. Q. 137, 144 (2000).

59. See, e.g., Cox & Miles, *supra* note 42, at 43 (finding no statistically significant effect of gender on judges' voting patterns in voting rights cases); Sisk et al., *supra* note 46, at 593 (finding no statistically significant effect of gender on judges' voting patterns in religious liberty cases); Kastlelec, *supra* note 45, at 178 (finding no statistically significant effect of gender on judges voting in affirmative action cases regarding race); Kenneth L. Manning, Bruce A. Carroll & Robert A. Carp, *Does Age Matter? Judicial Decision Making in Age Discrimination Cases*, 85 SOC. SCI. Q. 1, 12 tbl.2 (2004) (finding no statistically significant gender differences in judges' voting in age discrimination cases); Orley Ashenfelter, Theodore Eisenberg & Stewart J. Schwab, *Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes*, 24 J. LEGAL STUD. 257, 262 (1995) (finding only "modest" effects of gender on judges' voting in civil rights cases).

60. SUNSTEIN ET AL., *supra* note 7, at 167, 171, 185, 197; HAIRE & MOYER, *supra* note 56, at 47-48 (finding no evidence of female judges voting differently from their colleagues when aggregating all case types).

Thus, the studies have found very limited voting differences based on gender. As Christina Boyd and Adam Rutkowski put it, “[a] relatively large number of empirical studies . . . have failed to find evidence that female and male judges decide cases differently from one another, particularly outside of issue areas that are not closely related to ‘women’s issues’ like sex discrimination.”<sup>61</sup>

## 2. Race

Studies have found racial voting differences in some race-salient case types but not in broader categories of cases. For instance, Jonathan Kstellec found racial voting differences in affirmative action and death penalty cases. Looking at all affirmative action decisions regarding race in the U.S. Courts of Appeals between 1971 and 2008, Kstellec found that Black judges were more likely to support affirmative action programs.<sup>62</sup> In death penalty cases, Kstellec found that adding a Black judge to a non-Black panel significantly increased the chances of granting relief to defendants on death row when the defendant was Black.<sup>63</sup> Other studies have found similar differences in other race-salient cases, such as racial harassment, voting rights, and police misconduct cases.<sup>64</sup> The underlying theory with respect to race-salient cases is that non-White judges will approach these issues differently than their White counterparts because of their life experiences and views given the long history of racial discrimination in the United States.<sup>65</sup>

---

61. Christina L. Boyd & Adam G. Rutkowski, *Judicial Behavior in Disability Cases: Do Judge Sex and Race Matter?*, 8 POL., GRPS., & IDENTITIES 834, 837–38 (2020). One theory behind this outcome suggests that all judges, regardless of background, are so influenced by their training prior to taking the bench, and constrained by judicial norms and practices, that any differences from their background are offset and have no systematic impacts on their behavior. See Boyd, *supra* note 54, at 790.

62. Kstellec, *supra* note 45, at 179; see also, Peresie, *supra* note 56, at 1774, 1776 (same).

63. Jonathan P. Kstellec, *Race, Context, and Judging on the Courts of Appeals: Race-Based Panel Effects in Death Penalty Cases*, JUST. SYS. J., Nov. 11, 2020, at 410.

64. See, e.g., Pat K. Chew & Robert E. Kelley, *Myth of the Color-Blind Judge: An Empirical Analysis of Racial Harassment Cases*, 86 WASH. U. L. REV. 1117, 1134 (2009) (finding that plaintiffs in racial harassment cases are more than twice as likely to succeed under a Black judge than a White judge); Cox & Miles, *supra* note 42, at 30, 43 (finding that a judge’s race is associated with higher likelihood of voting in favor of liability in voting rights cases on the U.S. Courts of Appeals); Nancy Scherer, *Blacks on the Bench*, 119 POL. SCI. Q. 655, 668 (2004) (finding a “statistically significant difference in the voting behavior between black and white judges” in police misconduct cases).

65. HAIRE & MOYER, *supra* note 56, at 25, 32–33.

As with gender, studies have generally not found racial differences when studying more broadly ideologically salient cases or when aggregating all case types together.<sup>66</sup> Thus for both gender and race, voting differences have been confined to a small subset of cases that are gender- or race-salient. Scholars have typically not found broader differences that comport with political ideology more generally.

One common, and significant, limitation of the studies looking for party, gender, and racial differences was that very few were able to control for year and circuit.<sup>67</sup> A major reason for this is that previous researchers have worked with limited datasets, usually containing hundreds of cases,<sup>68</sup> which makes statistical adjustment difficult.

---

66. See *id.*, at 32; SUNSTEIN ET AL., *supra* note 7, at 167, 171, 185, 197; Farhang & Wawro, *supra* note 45, at 321 (“[R]acial minority judges on the federal Court of Appeals in the period sampled do not hold views different from White judges on employment discrimination claims, as measured by case outcome.”); see also HAIRE & MOYER, *supra* note 56, at 32 (finding that Black judges do not vote more liberally than White judges on the Courts of Appeals when data is pooled over many policy areas). *But see* Sisk et al., *supra* note 46, at 595–96 (finding that a judge’s minority race is associated with a higher likelihood of voting in favor of plaintiffs alleging religious discrimination, and that minority judges are more willing to take non-mainstream approaches in religious freedom cases).

67. See, e.g., Kastlelec, *supra* note 45, at 172–73 (neither controlling by year nor circuit); Sisk et al., *supra* note 46, at 553–55 (same); HANSFORD & SPRIGGS, *supra* note 22, at 42–46 (same); Hinkle, *Panel Effects*, *supra* note 24, at 322, 324 (same); Boyd, *supra* note 54, at 792–93 (same); Cox & Miles, *supra* note 42, at 21–22, 25–26 (same); Manning et al., *supra* note 59, at 7–8 (same); Chew & Kelley, *supra* note 64, at 1138 (same). Some were able to control by either year or circuit. See, e.g., Farhang & Wawro, *supra* note 45, at 315 (controlling for circuit); Peresie, *supra* note 56, at 1775–76 (same); Scherer, *supra* note 64, at 666 (controlling by region and defining some regions by circuit); Kulik et al., *supra* note 57, at 76, 80 (controlling by year); HAIRE & MOYER, *supra* note 56, at 159 (same). Very few were able to control for both. One of the few exceptions is Shahshahani & Liu, *supra* note 46, at 726 (controlling for both year and circuit).

68. See, e.g., Kastlelec, *supra* note 45, at 173 (studying a total of 182 cases between 1971 and 2008); Sisk et al., *supra* note 46, at 553 (studying 729 or fewer decisions depending on the model used); Boyd, *supra* note 54, at 793 (studying between 186 and 450 observations of judge voting); Cox & Miles, *supra* note 42, at 8 (studying 342 decisions); Manning et al., *supra* note 59, at 5 (studying 544 cases); Chew & Kelley, *supra* note 64, at 1138 (studying 428 cases); Farhang & Wawro, *supra* note 45, at 310 (studying 400 cases); Kulik et al., *supra* note 57, at 75 (studying 143 cases); Peresie, *supra* note 56, at 1767 (studying 556 cases); Scherer, *supra* note 64, at 672 (studying 550 cases). *But see* HANSFORD & SPRIGGS, *supra* note 22, at 51 (studying 6,363 cases); Hinkle, *Panel Effects*, *supra* note 24, at 322 (including 6,693 cases in its study); Shahshahani & Liu, *supra* note 46, at 721, 735–36 (studying 1,058 religious freedom cases and 2,100 cases not involving religion).

### III. OUR RESEARCH QUESTIONS

The questions we want to examine in this Article are all elaborations of a simple inquiry: How do judge-specific characteristics relate to the substantive treatment of earlier opinions? To answer these questions, we focus on majority opinion authors and estimate differences in substantive treatment practices across different types of authors.<sup>69</sup> Importantly, in all our analyses we adjust for circuit, year, and circuit-year effects to get as close as possible to apples-to-apples comparisons that are descriptively informative about substantive treatments.<sup>70</sup> And, in light of the possibility of temporal changes in behavior, particularly in light of increases in measures of polarization among decisionmakers in the many years our data cover, we also evaluate changes over time. To do this, we split our data into four equally sized time periods (1974–84, 1985–95, 1996–2006, and 2007–17) and estimate partisan, racial, and gender differences within each time period.

#### *A. Partisan Differences*

The first opinion-author-specific attribute that we examine is partisanship as proxied by the appointing President's party.<sup>71</sup> More specifically, we ask: Are opinion authors of a given party more likely than authors from the opposite party to follow opinions written by fellow members of their party?

If judges are pervasively partisan, then we might expect to find statistically significant results if we look at all panels in all types of cases. After all, insofar as judges are deeply partisan, we might expect their partisanship to arise across the board. But such pervasive partisanship may seem somewhat unrealistic. Even those in the political branches find room to agree on some relatively less ideological matters.

---

69. We focus on opinion authors in light of the centrality of their role on the panel in crafting the discussion contained in the majority opinion. *See, e.g.*, Chad M. Oldfather, *Remedying Judicial Inactivism: Opinions as Informational Regulation*, 58 FLA. L. REV. 743, 794–95 (2006). In a separate paper we measure the extent to which there are partisan panel effects (i.e., whether the party of other panel members influences substantive treatments contained in the majority opinion).

70. *See* Fan Li, Alan M. Zaslavsky & Mary Beth Landrum, *Propensity Score Weighting with Multilevel Data*, 32 STAT. MED. 3373–87 (2013) (describing such average controlled differences).

71. *See supra* notes 42–43 and accompanying text.

It is not clear that there are any categories of cases that are totally nonideological. Even run-of-the-mill torts and contracts cases can have some political valence (e.g., perhaps the average Republican appointee is relatively more sympathetic to defendants in torts cases and to enforcing contracts than is the average Democratic appointee). But we might expect cases involving politically charged issues like campaign finance and affirmative action to induce more ideological behavior than ordinary torts and contracts cases. And in fact most of the empirical studies on partisan judicial behavior focus on case topics that are expected to be ideological and thus polarized on partisan grounds.<sup>72</sup>

Thus a narrower form of the hypothesis above would expect larger partisan differences within case topics that are most likely to be ideologically charged along partisan lines. We canvassed previous studies for the case topics they identified as more likely to divide judges along political ideology lines, a category of cases we refer to as ideologically salient.<sup>73</sup> We then identified all the Lexis topics that involved one of these case topics. That yielded thirty-eight Lexis case topics.<sup>74</sup> Note that some of these categories had relatively few cases (e.g., Establishment Clause and abrogation of state sovereign immunity).<sup>75</sup>

A focus on more ideological cases implicates the distinction between published and unpublished cases. For much of the period our data cover, circuit rules prohibited or at a minimum disfavored citation of unpublished cases.<sup>76</sup> Circuit rules allow the ruling panel

---

72. See *supra* notes 44–46 and accompanying text.

73. See *supra* notes 44–46 and accompanying text.

74. We list the thirty-eight topics in Appendix A1. We identified a thirty-ninth ideologically salient topic (federalism), but Lexis did not use federalism as a case topic header in any of the cases in our dataset.

75. Previous studies may have chosen these categories in part because the number of cases was small enough to allow them to address all the cases. Our dataset contains the entire universe of cases and we wanted to separate by circuit and year to isolate effects, so the small numbers in some categories made it extremely unlikely that we would find statistical significance. The error bars in the accompanying figures reflect this.

76. See, e.g., *In re Citation of Unpublished Opinions/Orders and Judgments*, 151 F.R.D. 470 (Nov. 29, 1993) (replacing its prohibition on citation of unpublished opinions with the following rule: “Unpublished opinions and orders and judgments of this court are not binding precedents, except under the doctrines of law of the case, *res judicata*, and collateral estoppel. Citation of these unpublished decisions is not favored.”); FED. R. APP. P. 32.1. (“A court may not prohibit or restrict the citation of federal judicial opinions, orders, judgments, or other written dispositions that have been: (i) designated as ‘unpublished,’ ‘not for

to choose whether to publish an opinion based on its importance and precedential value.<sup>77</sup> The point of the unpublished designation is to allow judges to issue relatively insignificant opinions. Indeed, each year thousands of very short opinions (often only a paragraph or two long, sometimes only a single sentence) are issued, virtually all of which are unpublished. Some have suggested that judges have on occasion refrained from publishing a given opinion to diminish its significance.<sup>78</sup> This highlights that there is some discretion involved in the decision to publish an opinion. But there is no evidence of partisan, racial, or gender differences in decisions not to publish an opinion or to follow an unpublished opinion, so we have no reason to believe that the decision to publish or not publish a given opinion would affect our findings.<sup>79</sup>

---

publication,' 'non-precedential,' 'not precedent,' or the like; and (ii) issued on or after January 1, 2007.")

77. For example, as the Ninth Circuit outlines in its "CRITERIA FOR PUBLICATION:"

A written, reasoned disposition shall be designated as an OPINION if it: (a) Establishes, alters, modifies or clarifies a rule of federal law, or (b) Calls attention to a rule of law that appears to have been generally overlooked, or (c) Criticizes existing law, or (d) Involves a legal or factual issue of unique interest or substantial public importance, or (e) Is a disposition of a case in which there is a published opinion by a lower court or administrative agency, unless the panel determines that publication is unnecessary for clarifying the panel's disposition of the case, or (f) Is a disposition of a case following a reversal or remand by the United States Supreme Court, or (g) Is accompanied by a separate concurring or dissenting expression, and the author of such separate expression requests publication of the disposition of the Court and the separate expression.

9th CIR. R. 36-2 (Rev. Jan. 1, 2012).

78. Indeed, there is some evidence that a judge on a given panel may threaten to write a dissent to a proposed opinion unless the panel agrees to issue the opinion as unpublished, thus using the threat of a dissent to push the opinion into the less salient unpublished category. See Mitu Gulati & Catherine McCauliff, *On Not Making Law*, 61 L. & CONTEMP. PROBS. 157, 204 (1998) ("Two judges inclined to reverse in a close case might agree to affirm without opinion when the third judge threatens to dissent from a published opinion ordering reversal; on the other hand, the third judge may agree to vote for an affirmance if only a nonprecedential JO [Judgment Order] is used."). If we were focusing on the prevalence of dissents, the possible suppression of dissents to avoid publication might be relevant. But there is no evidence of partisan, racial, or gender differences with respect to which judges might threaten (or might respond to a threat) to issue a dissent unless an opinion is unpublished. So there is no reason to believe that our results are systematically affected by the effects of these threats.

79. See Rachel Brown, Jade Ford, Sahrula Kubie, Katrin Marquez, Bennett Osdiek & Abbe R. Gluck, *Is Unpublished Unequal? An Empirical Examination of the 87% Nonpublication Rate in Federal Appeals*, 107 CORNELL L. REV. 1, 35-41 (2021) (comprehensively reviewing differences between published and unpublished opinions).

*Shepard's* provides further reason not to include unpublished opinions in our study. As compared to published opinions, unpublished opinions contain fewer average *Shepard's* treatments of earlier cases. Published opinions contain an average of 1.9 *Shepard's* treatments, and unpublished opinions contain an average of 0.5.<sup>80</sup> This is not surprising, given that the unpublished designation is for opinions that are not designed to make new law for the circuit, and *Shepard's* treatments are extensive discussions of earlier cases that shed important light on their precedential value. Put differently, the point of *Shepard's* is to describe the way in which opinions grapple with earlier cases as they help to shape the law, and the point of unpublished decisions is to have a category of opinions that are particularly straightforward and thus do not need to grapple with earlier cases. Focusing on published opinions removes 66.4% of the cases, but only 11.1% of the *Shepard's* treatments, in our dataset.

Finally, denoting an opinion as unpublished not only indicates less significance but also helps ensure that the opinion will in fact have less significance. Just as legislators are more likely to focus on more important bills, we would expect judges to emphasize published opinions. Those published opinions are, by circuit rule and court practice, the opinions on which future judges are likely to rely. We can thus refine our hypotheses above by limiting our focus to published opinions, on the theory that such opinions are the ones on which judges will actually focus.

### *B. Beyond Partisanship – Racial and Gender Differences*

We turn now to race and gender. The underlying theory is that partisanship may not capture elements that judges find relevant in deciding which opinions to follow or cast doubt on, and that there may be commonalities among judges along race or gender lines that partisanship does not capture. But different theories yield different hypotheses.

---

80. These numbers count *Shepard's* treatments of U.S. Courts of Appeals opinions decided between 1974 and 2017, to make them comparable to the dataset that we use for our analysis. Many of the treating cases also contain treatments of Supreme Court opinions, district court opinions, and pre-1974 federal appellate opinions. The average number of total treatments in published opinions in our dataset is 2.7, and for unpublished opinions it is 1.6. These larger averages reflect all treatments, including those of Supreme Court opinions etc.



The studies of racial and gender differences in voting noted in Part II suggest one possibility: judges may have a particular understanding of and sensitivity to particular kinds of issues, perhaps flowing from their expertise and lived experiences.<sup>81</sup> A Black judge, for example, might have a deeper understanding than White judges (including within the same party) of the ways that White people engage in racial discrimination because she is more likely to have been subject to racial discrimination and to have seen multifarious forms of it. The same reasoning might apply to sex discrimination cases with respect to female versus male judges (again, including within the same party). If so, then we might expect to find differences between co-partisan judges of different races for race-salient cases and between co-partisan male and female judges for gender-salient cases.<sup>82</sup> Democratic Presidents have been more likely to nominate Black, Hispanic, and female judges than have Republican Presidents, so looking at racial and gender differences within party is important because it rules out the possibility that any observed differences are actually driven by differences in partisanship.

Because the theory is that judges are favorably treating opinions whose substantive approach they agree with, differences in these treatments based on race or gender can be understood as policy-motivated differences.<sup>83</sup> But note that such policy-specific motivations are distinct from broader, more all-encompassing ideological

---

81. *See supra* notes 51–54, 65 and accompanying text.

82. The terms race- and gender-salient refer to particular categories of cases that previous studies have suggested might divide judges along race or gender lines. *See supra* notes 51–58, 62–65 and accompanying text. For the list of the race-salient and gender-salient categories of cases that we compiled from previous studies and use in this Article, see *infra* Appendix A1. There are eighteen race-salient categories and twelve gender-salient categories. The race- and gender-salient categories are also included in the thirty-eight ideologically salient categories (unsurprisingly, categories that may cut along race or gender lines may also cut along ideological lines).

83. Given the overlap between race/gender and ideology, it may be that race- and gender-salient cases are the most ideological of all cases. If so, then the categories of race- and gender-salient cases would be best understood as purer measures of ideology than the broader category of ideologically salient cases. The literature has not established such a relationship among these categories, however. Instead, studies have put forward race- and gender-salient categories as likely to have particular significance along race or gender lines without indicating that they are more purely ideological. And we draw our categories of ideologically salient, race-salient, and gender-salient categories from the existing literature. So we have no basis for concluding that the race- and gender-salient categories are the most ideological of all cases.

motivations, as they arise from issues that are particularly salient with respect to the lived experience of race and gender.

A different possibility would suggest racial and gender differences in a wider range of cases than particularly race- or gender-salient cases: maybe there are broad ideological differences based on race or gender, similar to those based on partisanship, that go beyond what partisanship alone reveals. Indeed, there is some evidence that within a given party, Black and Hispanic judges differ in their political ideology from White judges, and female judges differ ideologically from male judges—with White men being the most conservative group within each party.<sup>84</sup> If so, race and gender might reflect some important elements of political ideology that partisanship does not capture. The existing evidence for this proposition comes from the analysis of pre-confirmation campaign contributions of judges, not judges' behavior.<sup>85</sup> And, as we noted in Part II, the studies addressing judges' behavior have focused on judges' votes and have generally found racial and gender differences only in a subset of race-salient and gender-salient cases. No other study has had access to our comprehensive data, and none has been able to measure racial or gender differences in substantive treatments or to measure racial or gender differences across ideologically salient cases. Our dataset, by contrast, allows us to examine opinions' reasoning for both the more specific categories of cases that studies have posited as particularly salient for race or gender purposes as well as the broader category of cases that studies have found to be ideologically salient more generally.

Insofar as there are broad ideological differences within parties based on race or gender akin to differences based on partisanship, then just as we might expect Democratic authors to differ from Republican authors in their treatments of earlier opinions across a wide range of case categories (because of ideological differences), we might expect similar differences within parties based on race and gender across a wide range of ideologically salient cases (again, because of ideological differences). So with respect to the thirty-eight ideologically salient case types we identified, we might expect

---

84. See Sen, *supra* note 15, at 394 tbl.4 (relying on preconfirmation campaign contributions to identify ideological differences within party based on race and gender).

85. *Id.*

female judges to be more likely than male co-partisans to follow majority opinions written by other female co-partisans, and for the converse to be true with respect to opinions by male judges.<sup>86</sup> And we might expect a similar effect based on race. Race and gender would provide information about political ideology beyond what party membership reveals about ideology.

These two possible effects are independent of each other. Maybe, for example, there is a gap between male and female judges in ideologically salient cases but a bigger gap between male and female judges in gender-salient cases, which would provide support for both types of effects. Or it could be that there is a gender difference for all ideologically salient cases but no greater difference for gender-salient cases, or conversely that there is a gender difference within gender-salient cases but not for ideological cases more generally. On the other hand, if one observed gender differences for the subset of ideologically salient cases that excludes gender-salient cases and smaller gender differences for the gender-salient cases, then this would indicate that the observed “gender” differences are less about gender and more about general ideology/political preferences. And the same possible comparisons exist for race—comparisons among all ideologically salient versus among race-salient cases would yield information about the degree to which differences were broadly ideological or more narrowly focused on areas relevant to expertise and lived experiences.

### *C. Solidarity Effects*

The general ideological and more policy-specific hypotheses discussed in the sections above are the ones with the strongest grounding arising out of those studies and the theory underlying them. But our data cannot establish that the explanation for any party differences or intraparty racial or gender differences is general ideology or more specific differences, as opposed to something else. What else can explain party, racial, or gender differences in treatments of earlier cases? We have no other

---

86. As we discuss in Sections V.B and V.C, because of the lack of Black, Hispanic, and female judges in the early years of our study, our investigation of racial and gender differences looks at substantive treatments of Black-authored, Hispanic-authored, and female-authored opinions and not White-authored or male-authored opinions.

hypotheses that are grounded in studies of judicial behavior. But another possibility occurs to us as plausible: perhaps there is an in-group preference that affects behavior, which we might call a solidarity effect.<sup>87</sup> On this theory, a judge might choose to follow an opinion written by a judge of the same party, race, or gender not because she had a greater affinity for the substance of that judge's opinion but instead because of their shared party, race, or gender. Following the opinion would be a way of supporting the colleague and the group.

This solidarity effect differs from the broad ideological and more policy-specific explanations in an important way regarding knowledge of, and interest in, the identity of opinion authors. Insofar as any of the broad ideological or more policy-specific differences discussed above exist, it could be that the later judge is influenced by the identity of the opinion author. The later judge could use the party, race, or gender of the opinion author as a relevant factor (or even the sole factor) in identifying substantively attractive opinions. In this way, the later judge would be using party, race, or gender as a marker of ideology/policy. But note that neither the general ideological nor the more policy-specific hypotheses discussed above depend on the later judge knowing the party, race, or gender of the authoring judge. As we noted in the introduction, a later judge might follow an ideologically congenial opinion without noticing the identity of the author. Similarly, a Black/Hispanic or female judge might follow an

---

87. A preference for members of one's own group is often called in-group preference, in-group favoritism, or in-group bias. See, e.g., Jerry Kang & Kristin Lane, *Seeing Through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. REV. 465, 476 n.37 (2010) (discussing in-group preference); Robert J. Smith, Justin D. Levinson & Joë Robinson, *Implicit White Favoritism in the Criminal Justice System*, 66 ALA. L. REV. 871, 895 (2015) (discussing in-group favoritism); John T. Jost, Mahzarin R. Banaji & Brian A. Nosek, *A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo*, 25 POL. PSYCH. 881, 902 (2004) (same); Laurie A. Rudman & Stephanie A. Goodwin, *Gender Differences in Automatic In-Group Bias: Why Do Women Like Women More than Men Like Men?*, 87 J. PERSONALITY & SOC. PSYCH. 494, 494-95 (2004) (discussing in-group bias); Celina M. Chatman & William von Hippel, *Attributional Mediation of In-Group Bias*, 37 J. EXPERIMENTAL SOC. PSYCH. 267, 271 (2001) (same); Jennifer A. Richeson & Nalini Ambady, *Who's in Charge? Effects of Situational Roles on Automatic Gender Bias*, 44 SEX ROLES 493, 494 (2001). We use the term "solidarity effect" for two reasons. First, some of the literature on in-group preferences uses the term to focus on implicit preferences, and in this discussion we are not assuming that any preference is merely implicit. See, e.g., Smith et al., *supra*. Second, the terms in-group preference, in-group favoritism, and in-group bias may have a negative connotation, and we want to avoid any such connotation.

opinion that discusses racial or sex discrimination in a sophisticated and powerful way without noticing the identity of the author. Indeed, insofar as party is correlated with ideology, race is associated with a more sophisticated understanding of race, and gender is associated with a more sophisticated understanding of sex, then we might expect party, racial, and gender differences in the substantive treatments of opinions in ideologically salient, race-salient, and gender-salient cases even in a counterfactual world in which judges' names (or other explicit identifiers of party, race, or sex) were not included in opinions: later judges would find the substance of the opinions written by those of the same party, race, or gender more attractive and thus presumably be more likely to follow them.

By contrast, the solidarity effect depends on later judges not merely noticing but in fact acting on the identity of opinion authors. The whole point of the solidarity effect is that later judges prefer to follow the opinions written by judges of the same party, race, or gender, and that of course requires that the later judges be aware of the party, race, or gender of the author.

This does not mean that the preference for following opinions by judges of the same party, race, or gender need be conscious: the later judges might have such a preference but not acknowledge it even to themselves.<sup>88</sup> But the solidarity effect does entail that judges in fact act on a preference to follow the opinions of those of the same party, race, or gender when opinions written by judges of a different party, race, or gender are at least as ideologically congenial to them.<sup>89</sup>

---

88. That said, the solidarity effect may be more likely to be conscious than an ideological/policy effect would be. As to the latter, there is no cognitive dissonance entailed in a judge thinking that a particular form of reasoning is powerful without taking the added step of associating that reasoning with a particular worldview. It may be naïve or betray a lack of intellectual curiosity to fail to consider how the preferred reasoning accords with a particular worldview, but it need not entail the judge hiding anything from herself. By contrast, if a judge prefers to follow the opinions of those of the same party, race, or gender because of that shared characteristic but does not acknowledge that preference to herself, she would seem to fail to understand her motivations. Whereas relating a particular form of reasoning to ideology/policy entails making a (possibly nonobvious) connection, no such connection is required in relating the identity of the opinion author to the identity of the opinion author—they are one and the same. It is of course possible that any in-group preference is unconscious (there is an extensive literature on implicit group preferences, *see supra* note 87). But it seems at least as possible that a judge who preferred to follow the opinion of someone of the same party, race, or gender would be conscious of doing so.

89. By hypothesis, judges are not choosing to follow opinions based on ideological congeniality.

The proposition that judges' decisions are influenced by the identity of earlier opinion authors may seem implausible. Circuit judges write many opinions. Is it likely that they note the identity of the opinion author of many of the opinions they follow? We cannot know, of course, but we believe it is plausible that judges are aware of the author's identity for at least some of the opinions they follow. First, the average published opinion follows 1.1 opinions and negatively treats 0.5 opinions, and the average unpublished opinion follows 0.4 opinions and negatively treats 0.1 opinions. A busy judge likely is not aware of the identity of the authors of all the opinions she cites, but she may well note the identity of the author of the one or two opinions she chooses to rely on as controlling.<sup>90</sup> Second, 71.2% of *Shepard's* treatments are to cases within the same circuit, the median time difference between the treating and treated opinion is 5.1 years, and the mean time difference is 7.5 years. So if later judges see the name of the opinion author, they are likely to know that author's identity.

As we noted above, the idea that ideology matters reflects an assumption that judges want to move the law in what they regard as a positive direction, and that a judge's definition of "positive" will be correlated with ideology, with the result that judges will thus want to push the law in an ideological direction.<sup>91</sup> The connection between identity and judicial motivations is less obvious. What is the non-ideological reason why a judge might choose to follow opinions based on shared party, race, or gender? The answer must be that one (or more) of those characteristics is important to judges, and that importance translates into influence on judicial behavior. As to importance, the idea is that some characteristics are likely to be particularly significant to the self-definition of those who share them. Race and gender, for example, are likely more central to many people's self-definition than are many other characteristics they may have (e.g., height).<sup>92</sup> That importance may manifest itself as influence for two related reasons. First, the importance of the characteristic may lead those who share it to feel an allegiance with one another and a desire to enhance the status of others with that characteristic. Second, enhancing the

---

90. On what is entailed in *Shepard's* identifying an opinion as following another opinion, see *supra* notes 30–35 and accompanying text.

91. See *supra* note 43 and accompanying text.

92. See *supra* notes 52, 54, 56, 64 & 87 and accompanying text.

status of others with a shared characteristic necessarily means enhancing one's own status.

This effect is probably fairly attenuated with respect to party. There are many judges of each party, and that large size would seem to diminish the level of solidarity and the attractiveness of trying to enhance the status of the members of one's party. For race and gender, though, the story seems more plausible. As with party, there are so many White judges that any solidarity effect among White judges would likely be very small. But there have been relatively few female, Black, and Hispanic judges, and it is possible that female, Black, and Hispanic judges might feel an allegiance to the small number of other judges who share their race or gender and be aware (consciously or unconsciously) that following the opinion of a judge who shares that characteristic enhances the status of everyone who shares it (including the judge who issues the positive treatment).

The discussion above leads to two possible forms of a solidarity effect. One is that judges have a fairly consistent general preference for enhancing the status of those of the same race or gender for non-ideological reasons. When considering which of several similar opinions to follow, judges will be inclined to follow the opinions of those of the same race or gender. If, say, female judges prefer to follow the opinions of other female judges, then we should expect to see intraparty gender-based differences in positive treatments as soon as there were enough female judges to allow for meaningful comparison that continue throughout the remainder of our study period.

A different possibility is that female, Black, and Hispanic judges felt a level of solidarity and kinship arising from their shared status as a small group of relative trailblazers. Solidarity arising from such a trailblazer effect would suggest a pattern to treatment differences based on race and gender: when there were very few female, Black, or Hispanic judges, we might expect greater differences between female and male judges, and between Black or Hispanic and White judges, on the theory that the benefits of positive treatments would be particularly meaningful for the first few judges with a particular characteristic.

Once there was a critical mass of female, Black, or Hispanic judges, we might expect those differences to be reduced, on the theory that, with a critical mass of judges with a particular attribute,

the sense of being a tiny cohort of trailblazers would no longer be as powerful.<sup>93</sup> With larger numbers, there might still be some solidarity effect, but it might be smaller because the relevant group would perceive itself to be more established and less in need of proving itself. For some (or all) of these groups, we might not yet have reached that point of critical mass, and current female, Black, and/or Hispanic judges might see themselves as trailblazers. If so, then we might expect to see a consistent difference in intraparty substantive treatments based on race or gender, rather than a decline in the most recent time period.

A trailblazer effect would thus suggest one of two possibilities. We might see larger gender differences in *Shepard's* treatments in the middle two time periods of our study (1985–95 and 1996–2006), on the assumption that there were too few female judges until 1985 and a sufficient number by the last time period in our study (2007–17) that female judges would be less likely to see themselves as trailblazing members of a tiny cohort. For Black judges, the trailblazer effect might have extended longer, and for Hispanic judges longer still.<sup>94</sup> Or, as suggested above, perhaps for some categories the trailblazer effect would persist, such that we might see a consistent effect with respect to race or gender in the last three time periods of our data (1985–2017).

#### *D. Measures and Data Limitations with Respect to Race and Gender*

We have identified three possible mechanisms that would give rise to racial and gender differences: racial or gender differences specific to issues about which Black, Hispanic, or female judges have particular expertise; racial or gender differences reflecting broader ideological differences, akin to those separating Democrats and Republicans; and racial or gender differences reflecting solidarity effects. We test all three by examining different categories

---

93. Laura Moyer and Susan Haire found such a trailblazer effect that dissipated over time. Specifically, they found differences in female and male judges' likelihood of voting in favor of plaintiffs in sex discrimination cases, but only for the earliest cohorts of judges. See Moyer & Haire, *supra* note 56.

94. There were at least two female judges in most circuits by 1992 (and all but two circuits by 1998). By contrast, in our time period three circuits never had more than one Black judge sitting at any given time, and it was not until 2002 that a majority of circuits had at least two Black judges. And the numbers for Hispanic judges are bleaker: in our time period, four circuits never had a Hispanic judge, and only four had more than one. See *infra* Appendix A2.



of cases, breaking them up into different time periods, and looking at differences within parties.

Specifically, we examine whether there are gender differences in treatments within party. We do this by subsetting the treating cases to just those written by the members of a given party and then looking to see whether female co-partisan opinion authors are more likely than male co-partisan opinion authors to follow opinions written by female co-partisans. We then do the same analysis focusing on gender-salient cases, more broadly ideologically salient cases, and the group of ideologically salient cases that excludes gender-salient cases. And we divide our forty-four years of data into four time periods to examine whether there are changes over time.

Similarly, we investigate whether the race of judges is associated with positive treatments of opinions written by same-race co-partisan judges – that is, holding partisanship constant.<sup>95</sup> We use biographical data from the Federal Judicial Center to label judges who self-identify as Black, non-Hispanic White, or Hispanic. And we do the analogous analysis focusing on race-salient cases, ideologically salient cases, and ideologically salient cases minus race-salient cases, and we divide our data into four time periods.<sup>96</sup>

Analysis of racial and gender differences depends on there being a sizable pool of cases written by Black, Hispanic, and female judges as well as at least one Black, Hispanic, or female judge in the circuit who can follow one of those earlier cases. And given the strong tendency of judges to follow cases within their circuit, the most substantively relevant results will arise when there are at least two Black, Hispanic, or female judges in a given circuit, such that all judges can choose to follow an opinion they did not write that was written by someone of the same or a different race or gender.

As section III.C indicated, this is a modest data limitation with respect to gender but a significant one with respect to race.<sup>97</sup> We could avoid these data limitations if we combined all circuit judges together into a single group undifferentiated by circuit or year, as

---

95. We provide a full explanation of these intraparty racial and gender measures in Sections V.B and V.C.

96. As with all the analysis in this Article, we estimate whether judges in each group are more likely to follow opinions from judges of their own group than are judges from the other group(s) after adjusting for circuit, year, and circuit-year combinations. See *supra* text accompanying note 70.

97. See *supra* note 94.

then there would be more than enough Black and Hispanic judges to allow for comparisons. We do not perform such an analysis because such an agglomeration of judges runs the risk of inaccurate findings. Because 71.2% of *Shepard's* treatments are to cases within a given circuit, it would be problematic for us to treat all judges or cases—even within a given year—as an undifferentiated whole. So just as we condition on each circuit-year combination when looking at partisanship, we do the same here.

With the advantage of avoiding spurious results comes the disadvantage of inferences that are only relevant for a narrowly defined population of cases from less than all circuit-years. The problem is most acute with respect to race. Our focus on circuit-year combinations means that for most such combinations there will not be enough Black or Hispanic judges in prior years to have many opportunities for substantive treatments of opinions written by minority judges. This is a data limitation that we have no control over. Importantly, it implies that our results—particularly those regarding White-Hispanic comparisons within Republican appointees—are only representative of a narrow set of cases from a limited set of circuits and years.

#### IV. DATA AND RESEARCH DESIGN

##### A. Data

To construct our data, we gathered all published and unpublished federal appellate opinions in the Lexis database issued between 1974 and 2017.<sup>98</sup> We separately identified each substantive *Shepard's*

---

98. We received the data directly from Lexis, which sent us all the circuit court opinions (including opinions issued by a single judge or two judges) in its database for our time period. We spent nine months analyzing the data for any possible lacunae or discrepancies and found none. We also compared the published cases from Lexis with the cases available from the Caselaw Access Project, <https://case.law>, and found greater than 99.9% agreement. We removed the opinions issued by a single judge or only two judges, which constituted approximately 1.4% of the observations, as we found that most of them were not decisions on the merits.

All the available online sources of opinions (including Lexis) fail to include some unpublished orders and opinions. See Michael Kagan, Rebecca Gill & Fatma Marouf, *Invisible Adjudication in the U.S. Courts of Appeals*, 106 GEO. L.J. 683, 685 (2018) (finding that some unpublished decisions in deportation proceedings are not in online databases, although Lexis has the best coverage); Merritt E. McAlister, *Missing Decisions*, 169 U. PA. L. REV. 1101, 1103 (2021) (finding that some unpublished orders and opinions are not in online databases). Apparently, the various circuits consider some unpublished orders and opinions to be

treatment in every court of appeals opinion. We obtained the data for court opinions, judges, and *Shepard's* treatments from LexisNexis. We focus on substantive treatments within majority opinions generated by panels of the federal courts of appeals.<sup>99</sup> Thus our measure of *Shepard's* treatments does not include treatments of district court opinions or Supreme Court opinions. In addition, we exclude the Federal Circuit from our analysis, because it has relatively few of the ideologically, race-, and gender-salient cases that we want to measure.<sup>100</sup> The full dataset comprises 670,784 federal appellate majority opinions and 648,226 *Shepard's* treatments.

In our data, 69.6% (451,289) of the treatments are positive, 8.5% (55,169) are neutral (which, as we noted previously, we drop precisely because they are neutral), and the remaining 21.8% (141,768) are negative.<sup>101</sup> As we noted in section III.C, the average published opinion has 1.1 positive treatments and 0.5 negative treatments of earlier opinions. More than 70% of the treatments address opinions decided within 10 years of the original opinion. Only 7.4% of the treatments in our data show a gap of more than 20 years from the original opinion, so it is relatively uncommon for an opinion to rely on opinions from an earlier era.

Lexis provided names of the authoring judge of an opinion for 301,337 cases, which amounts to 44.9% of our data. Most of these

---

sufficiently trivial that they are not even included among the unpublished orders and opinions that are sent to the online databases. As Kagan, Gill & Marouf note, the fact that these opinions are not available in online databases means that these opinions are not merely nonprecedential but also invisible. Kagan, Gill & Marouf, *supra*, at 689 (noting two categories of “invisible” decisions—“Nonprecedent, invisible decisions” and “Nonmerits decisions (invisible)”). Indeed, these opinions would be invisible not only to lawyers but also to judges and clerks, except for those few who might have worked on one of the invisible cases. See McAlister, *supra*, at 1149. The invisible opinions are thus not cited (much less discussed), which of course continues their invisibility. This point is significant for our purposes because our focus in this Article is on courts’ treatments of earlier opinions. Invisible opinions, by being invisible, are not available for later treatment and thus fall out of the denominator. That said, if there were patterns of invisible opinions that related to the party, race, or gender of the judges deciding the cases, that could bias our results. But there is no reason to believe that any such patterns exist. See McAlister, *supra*, at 1146–47; *supra* notes 76–80 and accompanying text.

99. As we note in Part I, *Shepard's* does not code dissenting or concurring opinions as positive or negative (for good reasons), so our data encompass majority opinions. See *supra* note 28.

100. See *supra* note 4.

101. As we noted in Part I, “followed” constitutes more than 99.99% of the positive treatments, so we refer to following and positive treatments interchangeably. See *supra* note 27 and accompanying text.

cases are published cases. We then used the data from the Caselaw Access Project to cross-reference and supplement the author names.<sup>102</sup> We dropped from our analyses 39,906 cases with author names indicating that they were per curiam opinions.<sup>103</sup>

We obtained data on judge-level characteristics such as judges' birth year, commission year, race, and gender from the Federal Judicial Center. Similarly, we used data from the Federal Judicial Center to identify the President who most recently nominated each judge.

Lexis assigns multiple topic headers to each case to identify the legal issue areas that a given case addresses. An average opinion has ten topic headers. This reflects the specificity of topic headers – an ordinary case does not cover ten completely different areas of law, but it might cover ten closely related and highly specific topics. Each topic header is a hierarchy moving from broad to more specific categories of law. A typical topic header is “Labor & Employment Law>Discrimination>Gender & Sex Discrimination>Evidence>Burden of Proof>Burden Shifting.”<sup>104</sup>

We categorized each opinion based on whether at least one of its topic headers contains the topic-identifying keyword for one of the thirty-eight ideologically salient issue areas, e.g., Search and Seizure, Immigration Law, and Sex Discrimination.<sup>105</sup> In our data, 213,619 cases are assigned topic headers that pertain to the thirty-eight ideologically salient issue areas we identified. We refer to this group of cases as the ideologically salient subset.

In each empirical analysis that follows, we fit the same model specifications to three different subsets – unpublished and

---

102. *Caselaw Access Project*, LIBR. INNOVATION LAB, <https://case.law> (last visited Sept. 16, 2023). The Caselaw Access Project provides open access to raw texts of all published opinions in U.S. courts. This allows us to compare the names of authoring judges of the published opinions in Lexis and Caselaw. We matched cases in the two datasets based on the case title and decision date. After multiple steps to adjust for different formats and styles, we supplemented the authoring judge for 9,901 of the opinions we received from Lexis with data from the Caselaw Access Project.

103. For some cases, the data shows the author's name as “Per Curiam,” and for others the data indicates that the case is per curiam but shows the names of all participating judges. We drop these cases and keep only opinions that show one judge name as the author.

104. This is a topic header from *Kidd v. Mando Am. Corp.*, 731 F.3d 1196 (11th Cir. 2013), a case we randomly selected from our dataset for purposes of illustrating the topic headers.

105. The thirty-eight ideologically salient issue areas are in Appendix A1. The topic categories are not mutually exclusive under our coding rule. In other words, an opinion can, and likely does, have multiple topic categories.

published combined, published only, and published and ideologically salient. There are 225,465 published opinions and 106,804 published and ideologically salient opinions in our data. In addition, some analyses make use of even more fine-grained subsets of published and race-salient cases (of which there are 70,066) and published and gender-salient cases (of which there are 46,458).

### B. Research Design

Because we have collected essentially all opinions issued by federal courts of appeals, we effectively have the entire population of data from the time period in question. We thus do not need to rely on random sampling of cases or other methods to ensure that our descriptive claims are accurate.

It is also worth reiterating that our decision to use *Shepard's* treatments as our primary outcome variable has several benefits. Our results do not depend on our own substantive judgments about how to code outcomes, and past research has shown the *Shepard's* measures to be reliable and valid.<sup>106</sup>

Further, we expect that our approach of looking at how sitting judges make use of the opinions of past judges is much more likely to produce measures that are comparable over time than approaches that focus on which litigant prevailed in a dispute and/or that attempt to discern the ideological valence of a decision. Measures that rely on the identity of the prevailing side run afoul of all manner of serious and not-so-serious selection issues, as strategic litigants will condition their litigation strategy on their expectation of prevailing on the merits.<sup>107</sup> The win rates of certain types of litigants are not of interest to us in this Article. We are also not interested in what would happen as a result of counterfactual changes in litigation strategies. In short, the behavior of litigants—along with the concomitant selection issues—is not of concern here.

Measures that require researchers to make substantive judgments about what the “liberal” or “conservative” outcome is within certain types of cases can miss key aspects of the legal reasoning and can also be difficult to compare over time. The meaning of “liberal” and “conservative” is timebound, and the types of disputes clearly change

---

106. See *supra* notes 30–36 and accompanying text.

107. See George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 11 (1984).

over time.<sup>108</sup> Because judges always have and will seek to justify their decisions by relying on earlier decisions, there is little question of the comparability of our outcome measures over time. Further, because we are interested in the body of law as it exists and not the body of law that would counterfactually exist had some disputes not settled (and we are not interested in the win rates of certain types of litigants), we do not need to be concerned about the litigation strategies of the parties.

As we noted in Part III, when comparing the behavior of Democratic appointees to Republican appointees, White judges to Black judges, White judges to Hispanic judges, and female judges to male judges, we attempt to get as close as we can to apples-to-apples comparisons. We do this by adjusting for circuit, year, and circuit-year fixed effects. The exact form of this adjustment is described in more detail in section C below. Further, we also report time-period specific differences to better assess the extent to which behavior is changing over time.

### *C. Estimation and Inference*

To adjust for circuit, year, and circuit-year effects, we estimate and report what are known as average controlled differences using overlap weights.<sup>109</sup> Put simply, for a particular comparison, say Democrat versus Republican, this approach works by first estimating the probability, within each circuit and year, that each case is authored by a Democrat and the probability that each case is authored by a Republican.<sup>110</sup> The behavioral difference of interest, say the difference in positive treatments of Democratic-authored opinions, is then defined as a weighted average difference where, in our running example, the weights would be proportional to the probability the case was authored by a Democratic appointee times the probability that the case was authored by a Republican appointee. In this running example, this average controlled difference is estimated by weighting the Democratic-authored opinions by the probability that they could have been authored by

---

108. See Carolyn Shapiro, *Coding Complexity: Bringing Law to the Empirical Analysis of the Supreme Court*, 60 HASTINGS L.J. 477, 480 (2009); *supra* notes 18–21 and accompanying text.

109. See Li, Zaslavsky & Landrum, *supra* note 70; Fan Li, Kari Lock Morgan & Alan M. Zaslavsky, *Balancing Covariates via Propensity Score Weighting*, 113 J. AM. STAT. ASS'N 390, 391 (2018).

110. These probabilities are referred to as propensity scores in the literature. Li, Morgan & Zaslavsky, *supra* note 109.

a Republican and weighting the Republican-authored opinions by the probability they could have been authored by a Democrat.<sup>111</sup>

As discussed by Fan Li, Kari Morgan, and Alan Zaslavsky, this approach to weighting the data serves to perfectly balance the measured covariate distributions across the comparison groups.<sup>112</sup> In other words, if we are looking at a comparison of Democratic judges to Republican judges, the weighted fraction of Democratic-authored opinions from a particular year and circuit will be the same as the weighted fraction of Republican-authored opinions from that same year and circuit. This is highly desirable, as it eliminates circuit-, year-, and circuit-year-specific factors as confounding variables.

Further, since the overlap weights are proportional to a probability (say the probability of a case being authored by a Democrat) times one minus that probability, the cases that will get the most weight are those with a 50-50 chance of being decided by either type of author in the comparison. Not only is this part and parcel with producing covariate balance, but it also importantly downweights cases from circuit-years where the comparisons of interest are simply difficult if not impossible to make in a credible fashion. For instance, if there are no Hispanic judges in a particular circuit and year, then it would not be meaningful to make White-Hispanic comparisons within that circuit and year. Our estimation approach automatically gives the substantive treatment decisions from that circuit and year zero weight.

We construct standard errors and confidence intervals using the nonparametric bootstrap.<sup>113</sup>

Finally, it is important to note that we estimate a large number of average controlled differences in this Article. A concern when conducting many hypothesis tests (or equivalently looking to see whether many *p*-values fall below a threshold) is the high likelihood of making many false discoveries (i.e., incorrect rejections of true null hypotheses). We guard against this by employing the methods

---

111. *See id.*

112. *See id.*

113. *See* BRADLEY EFRON & ROBERT J. TIBSHIRANI, AN INTRODUCTION TO THE BOOTSTRAP 42–45 (1993) (explaining the calculation of nonparametric bootstrap standard errors and confidence intervals).

of Benjamini and Yekutieli to control the false discovery rate.<sup>114</sup> We also use the associated method of Wright to adjust the  $p$ -values.<sup>115</sup>

## V. PRIMARY RESULTS

Having laid out our research questions, research design, and strategy for estimation and inference, we turn to the results of our study. In all these analyses, we attempt to minimize imbalances in the data by focusing on average controlled differences using overlap weights and adjusting for circuit, year, and circuit-year indicators.<sup>116</sup> As we noted in Part IV, this estimation strategy generates meaningful, apples-to-apples comparisons between the contrasting groups of interest (Democrat-Republican, Black-White, Hispanic-White, and female-male). It does so by downweighting data from circuit-years that are heavily skewed to one group – say, towards White judges and away from Hispanic judges. We begin with partisanship and then proceed to race and gender.

### A. Average Controlled Differences by Author Partisanship

We begin by examining the role played by partisanship, which we hypothesize to play a major role in structuring judicial behavior. We start by pooling our data over the full time span of the study (1974–2017) and looking at partisan differences in the propensity to follow Democratic-authored opinions and Republican-authored opinions. The results are broken down for all cases, published cases, and published and ideologically salient cases.

---

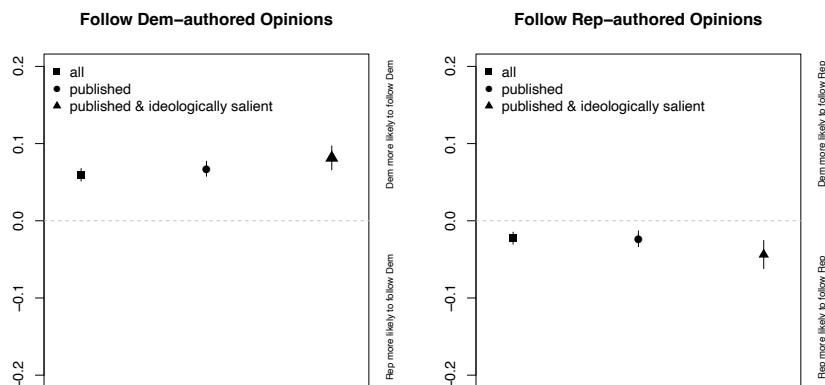
114. See generally Yoav Benjamini & Daniel Yekutieli, *The Control of the False Discovery Rate in Multiple Testing Under Dependency*, 29 ANNALS STAT. 1165, 1166 (2001) (providing a procedure to control the false discovery rate); S. Paul Wright, *Adjusted P-Values for Simultaneous Inference*, 48 BIOMETRICS 1005, 1007 (1992) (providing a correspondence between methods to control the false discovery rate and adjustment of  $p$ -values).

115. See Wright, *supra* note 114.

116. See *supra* note 109 and accompanying text.



Figure 1

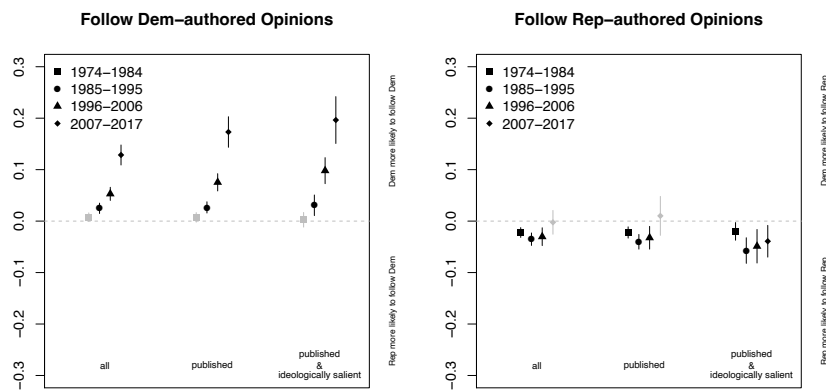


Partisan differences in positive treatments of opinions authored by a Democratic or Republican appointee for all, published, and published and ideologically salient cases. Points on the plot represent the difference in the propensity of Democratic authors versus Republican authors to follow opinions by other Democratic (left panel) or Republican (right panel) authors. Points above the horizontal line at 0 indicate a greater propensity of Democratic appointees than Republican appointees to follow the partisan-authored opinions in question. Points below the horizontal line at 0 indicate a greater propensity of Republican appointees than Democratic appointees to follow the partisan-authored opinions in question. Each panel plots three different average controlled difference estimates: the left point is for all cases (published and unpublished combined), the center point is for published cases, and the right point is for published and ideologically salient cases. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

Figure 1 displays these results. We find statistically significant partisan differences across all cases, published cases, and published ideologically salient cases. These differences are all in the expected direction. The left panel of Figure 1 shows that Democratic opinion authors are more likely to follow opinions authored by other Democratic appointees than are Republican opinion authors. Similarly, we see in the right panel of Figure 1 that Republican opinion authors are more likely than Democratic authors to follow opinions written by Republican appointees. For both plots in Figure 1 the partisan differences grow slightly stronger as we narrow our analyses down to published opinions and published and ideologically salient opinions. The partisan differences in positive treatments to Democratic-authored opinions are larger in absolute

value than the partisan differences in positive treatments to Republican-authored opinions. Further, the estimated differences are large enough to be substantively meaningful. For instance, the Democrat-Republican difference in the average positive treatments to Democratic-authored opinions is approximately 0.1, which corresponds to a Democrat giving one more positive treatment to a previous Democratic opinion than was given by a Republican in every 10 opinions.

Figure 2



*Partisan differences in positive treatments of opinions authored by a Democratic or Republican appointee over time for all, published, and published and ideologically salient cases. Points on the plot represent the difference in the propensity of Democratic authors versus Republican authors to follow opinions by other Democratic (left panel) or Republican (right panel) authors. Points above the horizontal line at 0 indicate a greater propensity of Democratic appointees than Republican appointees to follow the partisan-authored opinions in question. Points below the horizontal line at 0 indicate a greater propensity of Republican appointees than Democratic appointees to follow the partisan-authored opinions in question. The shape of the points denotes the average controlled difference for different time periods, with each period encompassing 11 years. For both panels, the left four points display the average controlled differences for all cases (published and unpublished combined), the center four points show the average controlled differences for published cases, and the right four points show the average controlled differences for published and ideologically salient cases. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.*

We next break these partisan author analyses down by four time periods (1974–1984, 1985–1995, 1996–2006, 2007–2017). Figure 2 plots partisan differences in the propensity to follow opinions

written by Democratic and Republican authors. Looking at *Figure 2*, we see that Democratic opinion authors are much more likely than Republican opinion authors to follow earlier Democratic-authored opinions. Further, these partisan differences grow dramatically larger over time with the size of the differences accelerating rapidly. The differences are largest, with the greatest acceleration, within the subset of published and ideologically salient opinions. In the most recent time period, Democrats make more than 1 additional positive treatment to a Democratic-authored opinion than do Republicans in every 5 opinions, as is seen in the estimated difference being greater than 0.2. There is some evidence that Republican opinion authors are more likely than Democratic opinion authors to follow opinions written by Republicans. While these differences are significantly different from zero in the first three time periods for all cases and published cases, and significantly different from zero in all four periods for published and ideologically salient cases, they are much smaller than the corresponding differences with respect to positive treatments of Democratic-authored opinions. Further, the differences do not become larger over time, in contrast to the positive treatments to Democratic-authored opinions.

*B. Average Controlled Differences by Author Race*

In this section we present results on the extent to which there are racial differences in substantive treatments. The history of Black and Hispanic representation on the federal courts of appeals gives rise to two related empirical patterns that need to be dealt with when conducting this analysis. First, Black and Hispanic judges appeared on the bench in substantial numbers only in the mid-1990s and after – over twenty years after the first year of our data.<sup>117</sup> Second, during the 1974–2017 time period Black and Hispanic judges were more likely to be appointed by a Democratic President than a Republican President. Both facts have implications for which outcome variables are most meaningful to study.

The fact that few Black and Hispanic judges were on the bench prior to the 1980s means that the vast majority of opinions written prior to the 1980s were written by White judges. It is thus not particularly useful to look at the propensity of White and minority

---

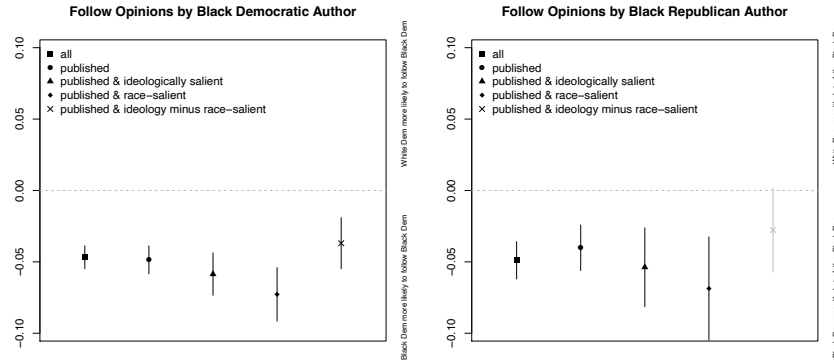
117. *See infra* Appendix A2.

judges to follow White-authored opinions without some form of adjustment, since both White and minority judges effectively had only White-authored opinions to treat until a sufficient body of minority-authored case law was developed well after the start of our data.

Relatedly, the fact that most Black and Hispanic judges—particularly in the early years of our study—were Democratic appointees means that there is also a substantial partisan skew to the minority-authored opinions that do exist. Failing to adjust for this partisan difference in the stock of minority-authored opinions that can be followed will also result in unreliable inferences about the role of race in structuring how judges treat earlier opinions.

The approach we take to deal with both issues is to condition our analysis on partisanship and to examine the extent to which White and minority judges from a given party positively treat opinions written by minority co-partisan judges. More specifically, we subset the data down to cases with substantive treatments authored by a given party and then estimate the average controlled difference between how White and either Black or Hispanic co-partisans positively treat past opinions written by Black or Hispanic co-partisans. Using positive treatments of minority-authored opinions as the outcome variable automatically adjusts for the later arrival of substantial numbers of minority judges in our data since both White and minority judges will have equal opportunity to positively treat minority-authored opinions as long as there are some minority judges (and thus some minority-authored opinions). Doing this within party adjusts for the partisan skew of the minority-authored opinions.

Figure 3

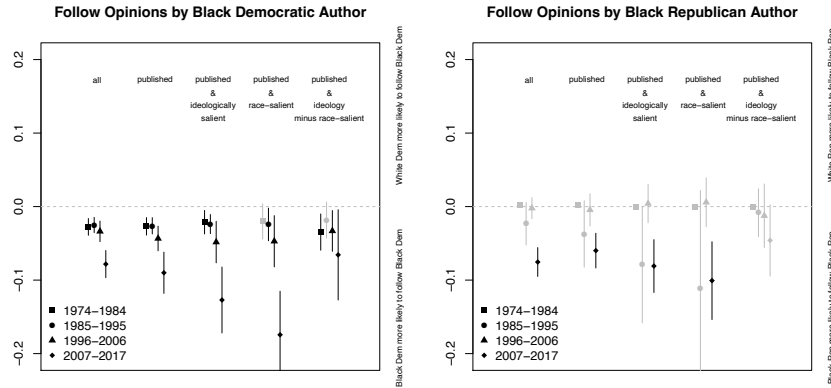


Intraparty racial differences in the positive treatments of opinions authored by a Black judge of the same party for all, published, published and ideologically salient, published and race-salient, and published and ideologically salient minus race-salient cases. Points on the left panel represent the difference in the propensity of White Democratic authors versus Black Democratic authors to follow opinions by Black Democratic authors. Points on the right panel represent the difference in the propensity of White Republican authors versus Black Republican authors to follow opinions by Black Republican authors. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

Figure 3 displays the average controlled White-Black intraparty differences (again controlling for year, circuit, and circuit-year interactions) aggregated over the full time span of our study but broken down by case type (all, published, published and ideologically salient, published and race-salient, and published and ideologically salient excluding race-salient).<sup>118</sup> As expected, within party, Black judges are more likely to follow the opinions of fellow Black judges than are White judges. These differences get somewhat larger as one moves from all cases to ideologically salient and race-salient cases and diminishes somewhat for the category of ideologically salient but not race-salient. That said, the differences are generally similar in magnitude across the various subsets of cases. We discuss the implications of this in section VI.B. Further, the size of these differences is similar across parties.

118. Recall that the race-salient cases (and the gender-salient categories) are subsets of the ideologically salient categories. See *supra* note 82.

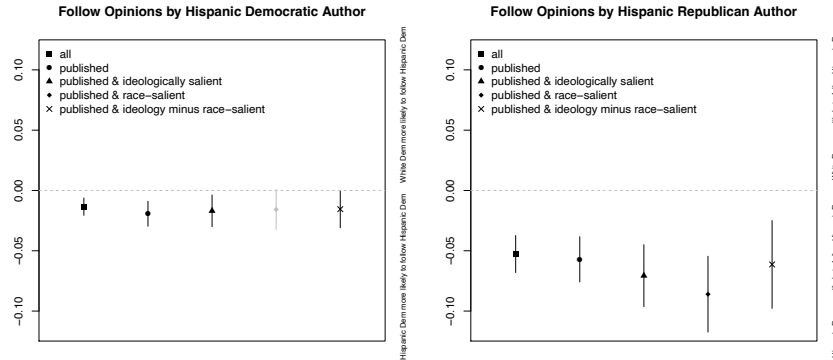
Figure 4



*Intraparty racial differences in the positive treatments of opinions authored by a Black judge of the same party over time for all, published, published and ideologically salient, published and race-salient, and published and ideologically salient minus race-salient cases. Points on the left panel represent the difference in the propensity of White Democratic authors versus Black Democratic authors to follow opinions by Black Democratic authors. Points on the right panel represent the difference in the propensity of White Republican authors versus Black Republican authors to follow opinions by Black Republican authors. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.*

Figure 4 disaggregates these White-Black intraparty differences by time period. Interestingly, the White-Black differences increase substantially over time. The within-Democratic White-Black differences are significantly different from zero in all but two case-subset-time-periods, but more notably the size of the difference generally accelerates rapidly over time, with the difference in the 2007–2017 time period several times larger than for the earlier periods. The size of these 2007–2017 differences is large and substantively meaningful. For instance, within the subset of published and race-salient cases, Black Democrats gave one more positive treatment to opinions written by Black Democrats in every five opinions than did their White Democratic colleagues. The 2007–2017 differences are also the largest for Republicans, although here the pre-2007 differences are not significantly different from zero.

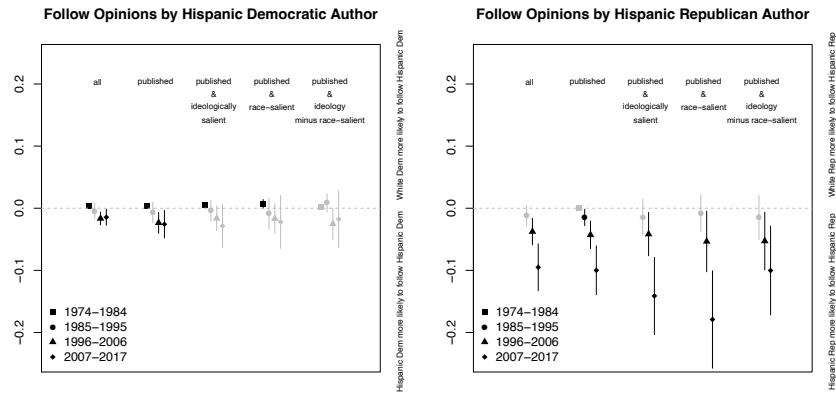
Figure 5



Intraparty racial differences in the positive treatments of opinions authored by a Hispanic judge of the same party, for all, published, published and ideologically salient, published and race-salient, and published and ideologically salient minus race-salient cases. Points on the left panel represent the difference in the propensity of White Democratic authors versus Hispanic Democratic authors to follow opinions by Hispanic Democratic authors. Points on the right panel represent the difference in the propensity of White Republican authors versus Hispanic Republican authors to follow opinions by Hispanic Republican authors. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

In Figure 5, we display the White-Hispanic intraparty average controlled differences (again controlling for year, circuit, and circuit-year interactions) aggregated over the full time span of our study and broken down by case type. All but one White-Hispanic difference is significantly different from 0 and in the expected direction—Hispanic authors being more likely than White authors to positively treat past opinions written by Hispanic co-partisans. The sole statistically insignificant difference is the partisan White-Hispanic difference within published and race-salient cases among Democratic appointees. Within party, the differences across different subsets of cases are not statistically distinguishable from each other. We return to this point and discuss its substantive interpretation in section VI.B. Interestingly, the White-Hispanic differences are larger within Republican judges than within Democratic judges.

Figure 6



*Intraparty racial differences in the positive treatments of opinions authored by a Hispanic judge of the same party over time for all, published, published and ideologically salient, published and race-salient, and published and ideologically salient minus race-salient cases. Points on the left panel represent the difference in the propensity of White Democratic authors versus Hispanic Democratic authors to follow opinions by Hispanic Democratic authors. Points on the right panel represent the difference in the propensity of White Republican authors versus Hispanic Republican authors to follow opinions by Hispanic Republican authors. Note that the data points for the first period in the right panel are not plotted due to a lack of data (insufficient Hispanic Republican appointees in 1974–1984). Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.*

Figure 6 breaks down the intraparty White-Hispanic differences by time period. Again, among Democrats, we see White-Hispanic differences that are close to zero, with most not statistically distinguishable from zero. In the first time period, the differences for published cases, published and ideologically salient cases, and published and race-salient cases are significantly different from zero but extremely close to zero. This is driven by the very small number of opinions by Hispanic judges in this time period. The available pool of opinions in the 1974–1984 period that a Hispanic judge could follow were overwhelmingly authored by White judges. There is perhaps some slight evidence that the differences are increasing over time for this group, but again, in the subset of cases where we would expect to see the largest differences—race-salient and ideologically salient cases—the White-Hispanic

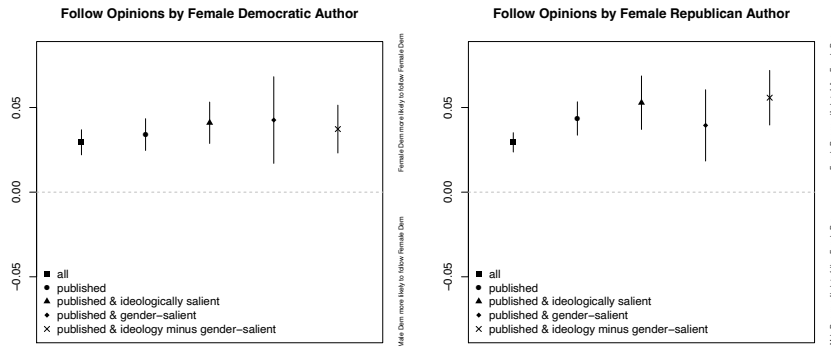


differences are not statistically significant (except for the 1974–1984 period, which is an artifact of the tiny number of cases with Hispanic judges in that time period). The White-Hispanic average controlled differences are much larger and grow much more rapidly over time for Republican appointees. Within this group, the White-Hispanic differences are either undefined or indistinguishable from zero in the first two time periods, but after 2007, we see that Hispanic Republicans are much more likely than White Republicans to positively treat Hispanic Republicans and, indeed, these differences accelerate over time.

C. Average Controlled Differences by Author Gender

The relatively recent and small female representation on the federal courts of appeals gives rise to the same sorts of issues discussed in section V.B, and we deal with this issue in the same way: we subset the data down to cases with substantive treatments authored by members of a given party and then estimate the average controlled difference between how female and male judges positively treat past opinions written by female co-partisan judges.

Figure 7



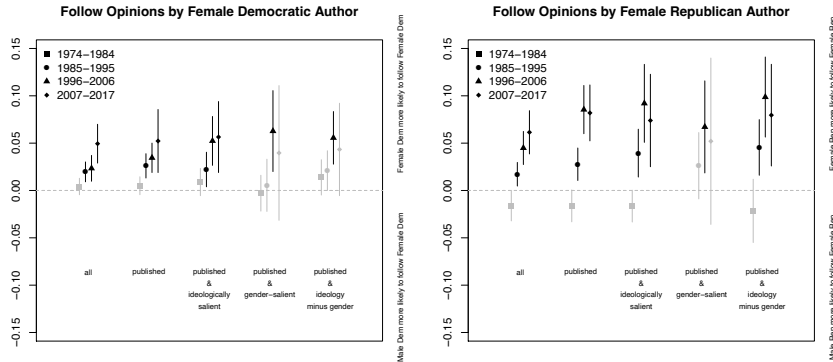
*Intraparty gender differences in the positive treatments of opinions authored by a female judge of the same party for all, published, published and ideologically salient, published and gender-salient, and published and ideologically salient minus gender-salient cases. Points on the left panel represent the difference in the propensity of female Democratic authors versus male Democratic authors to follow opinions by female Democratic authors. Points on the right panel represent the difference in the propensity of female Republican authors versus male Republican authors to follow opinions by female Republican authors. Differences that are statistically significant at the 0.05 level after adjusting for multiple*

*testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.*

*Figure 7* displays these estimated average controlled differences (again controlling for year, circuit, and circuit-year interactions) aggregated over the full time span of our study. As with the race results in section V.B, we break them down by case type. We focus on gender-salient rather than race-salient cases for the obvious reason that the literature on gender differences based on expertise and lived experiences focuses on gender-salient cases. And, analogous to race, we consider the ideologically salient cases that are not gender salient.

Looking at *Figure 7*, we see that after adjusting for partisanship, gender plays a role in guiding treatment practices—with female judges being more likely than male judges to positively treat the work of female co-partisans. The size of these differences is similar for both Democratic and Republican judges—within the subset of published and gender-salient cases, slightly less than one more positive treatment from a female judge to another female judge's opinion than from a male judge in every 20 opinions. Interestingly, the average controlled differences within the published and gender-salient subset of cases are not larger than the average controlled differences within the published and ideologically salient subset. Indeed, as with the Black and Hispanic results in section V.B, none of the intraparty differences corresponding to different subsets of cases are statistically distinguishable from each other.

Figure 8



Intraparty gender differences in the positive treatments of opinions authored by a female judge of the same party over time, for all, published, published and ideologically salient, published and gender-salient, and published and ideologically salient minus gender-salient cases. Points on the left panel represent the difference in the propensity of female Democratic authors versus male Democratic authors to follow opinions by female Democratic authors. Points on the right panel represent the difference in the propensity of female Republican authors versus male Republican authors to follow opinions by female Republican authors. The shape of the points denotes the averaged controlled difference for different time periods with each period encompassing 11 years. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

Figure 8 breaks down the average controlled differences in Figure 7 by time period. In each case, we see evidence that the intraparty gender differences are growing larger over time. This is most apparent within the Republican subset—particularly within the subset of published and ideologically salient cases. The relatively small number of published and gender-salient cases within each of the party-time-period combinations produces a great deal of estimation uncertainty which manifests as wide confidence intervals. This level of statistical uncertainty makes it difficult to draw strong conclusions about change over time within this subset of cases. Nonetheless, we do see strong evidence of increasing gender differences over time within all cases and published cases. Further, the aggregate results presented in Figure 7 provide strong evidence of intraparty gender differences over the entirety of our study (1974–2017).

## VI. DISCUSSION

We have examined judges' substantive treatments of opinions to determine if there are substantively meaningful and statistically significant differences in those treatments based on the partisanship, race, and gender of the judges. Our findings are more nuanced and interesting than we anticipated. Pooling the data over the full span of our study (1974–2017), we find statistically significant differences with respect to positive treatments (we do not find substantively meaningful differences for the less common negative treatments). But the real story arises from looking at changes in the partisan, race, and gender differences over time. Most of the differences we find are fairly small in the early periods of our study and rise dramatically over time, becoming large and therefore substantively meaningful. Further, most of the differences are greatest for the ideologically charged categories of cases.

*A. Partisan Differences*

As we noted above, treatment of earlier cases is a central element of an opinion's reasoning. And, relatedly, following a case helps to increase its importance. If judges were consistently inclined to act in a partisan manner, we would expect them to treat earlier opinions by the members of their party better than earlier opinions by members of the opposite party.

Looking at all time periods combined, we find statistically significant differences in how Democratic appointees versus Republican appointees treat earlier opinions, with somewhat greater partisan differences for earlier Democratic-authored opinions than for earlier Republican-authored opinions.

But once we break the cases down into four time periods, we see two different progressions in the point estimates over time (though some are within the reported confidence intervals): first, the magnitude of the partisan point estimates increases; second, the increase over time is greatest for ideologically salient published opinions. In other words, the point estimates suggest that partisanship in general increases over time, and partisanship with respect to the most ideological cases increases the most. And the final period (2007–2017) presents a particularly sharp increase, resulting in partisan differences that are not merely only statistically significant but also large and thus substantively meaningful.

How can we explain that shift over time? We begin by considering whether factors such as a replacement via presidential cohorts or the aging of judges are likely explanations. The data reveal that these factors are unlikely to explain our primary results. We then examine a possibility that is consistent with our findings – different rates of acceleration of ideological change across parties.

### *1. Partisan Differences Are Not a Function of Presidential or Age Cohorts*

One might imagine that part of the story has to do with the increasing politicization of the nomination and confirmation process that started in the Carter and Reagan administrations and carries through to this day.<sup>119</sup> More specifically, one might suspect that the increasing size of partisan differences in recent time periods has something to do with a presidential cohort effect, in which some presidential administrations might have outsized abilities to shift the ideological makeup of the courts. It is widely believed that there are nontrivial ideological differences between presidential cohorts of the same party.<sup>120</sup> If so, those differences might shed light on the partisan differences we find. Perhaps it is not that Democratic and Republican appointees in general are becoming more polarized but that earlier Presidents in our sample appointed moderate Democrats and Republicans and later ones appointed more extreme Democrats and Republicans, with the result that the replacement of the earlier presidential cohorts by the later cohorts produces the polarization we find.

We evaluate that possibility in this section. To summarize, our data do not (somewhat to our surprise) support the claim that the observed increases in partisan differences are due to presidential cohorts, or more generally to the replacement of moderate judges by more extreme judges.

A launching point is the fact that there have been changes in the presidential selection process. For most of the twentieth century, the party of the appointing President was not a particularly strong indicator of ideology. Presidential administrations deferred to a considerable degree to Senators' preferences in choosing circuit nominees, and those Senators often did not prioritize ideological

---

119. See *infra* notes 123–126 and accompanying text on the roles of Carter and Reagan in changing the process by which judges were chosen.

120. See *infra* notes 123–127 and accompanying text.

commitment to the national party in their choices.<sup>121</sup> Indeed, they were often patronage positions.<sup>122</sup>

The first big move away from senatorial influence came under President Carter, who appointed nominating commissions for each circuit.<sup>123</sup> Those commissions took recommendations from Senators, but the commissions made their own recommendations to the President (to the great annoyance of many Senators).<sup>124</sup> Carter proved to be a way station toward the more complete control that began in the Reagan Administration.<sup>125</sup> Reagan moved to a model of judicial selection that prioritized presidential discretion over senatorial influence, with a small group within the Reagan Administration choosing circuit nominees after engaging in extensive screening that emphasized ideology—“the most systematic judicial philosophical screening of judicial candidates ever seen in the nation’s history.”<sup>126</sup> Centralized control remained

---

121. See AMY STEIGERWALT, *BATTLE OVER THE BENCH: SENATORS, INTEREST GROUPS, AND LOWER COURT CONFIRMATIONS* 3–5 (2010). The Eisenhower Republicans on the Fifth Circuit were by many measures more liberal than the Kennedy Democrats, because the former came from the desegregationist party in the South and the latter from the segregationist party in the South. See also JACK BASS, *UNLIKELY HEROES* 84–96 (1981); VICTOR S. NAVASKY, *KENNEDY JUSTICE* 269 (1971); Kenneth N. Vines, *Federal District Judges and Race Relations Cases in the South*, 26 J. POL. 337, 348 (1964).

122. See NANCY SCHERER, *SCORING POINTS: POLITICIANS, ACTIVISTS, AND THE LOWER FEDERAL COURT APPOINTMENT PROCESS* 13 (2005) (noting that judgeships were “distributed to friends and campaign contributors”).

123. A Nixon aide named Tom Charles Huston recommended that Nixon focus on judicial nominations. The memo was forwarded to Nixon’s Deputy Attorney General with Nixon’s endorsement, but it did not lead to the centralization of the process within the Nixon Administration. See, e.g., Elliot E. Slotnick, *Federal Judicial Selection in the New Millennium*, 36 U.C. DAVIS L. REV. 587, 590 (2003) (“Despite [the Huston] memo, it is a bit too easy to point to the Nixon administration as the historical point in time where the most significant changes took place in the nature of federal judicial selection. The policy implications of judicial selection, which Huston spoke of, were not fully realized until the centralization of the judicial selection process during the Reagan years. More accurately, the modern era of contentious, politicized judicial selection politics can best be traced to the Carter administration.”).

124. See SHELDON GOLDMAN, *PICKING FEDERAL JUDGES: LOWER COURT SELECTION FROM ROOSEVELT THROUGH REAGAN* 126–31 (1997).

125. See Slotnick, *supra* note 123, at 593 (“Once the genie of openly avowed policy considerations in judicial selection had been let out of the bottle [under Carter], and once the White House’s political role in judicial selection increased, it would be difficult to return to the old ways. In the wake of the Carter years, during the two-term presidency of Ronald Reagan, the policy agenda of the president and centralized White House control of judicial selection was a major facet of selection processes.”).

126. Sheldon Goldman, *Reagan’s Judicial Legacy: Completing the Puzzle and Summing Up*, 72 JUDICATURE 318, 319–20 (1989); SCHERER, *supra* note 122, at 161.

for later administrations, but many researchers suggest that there are substantively meaningful differences in presidential cohorts of the same party – for instance, that the Reagan and George W. Bush judges were more conservative than the George H.W. Bush judges.<sup>127</sup>

The large size of our dataset gives us leverage to examine the relationship among presidential cohorts. More specifically, our data allow us to examine whether a judge's presidential cohort is associated with that judge's treatments of earlier opinions. We examine this question directly. To shed light on the proposition that increasing political polarization is linked to selection of judges, we look at substantive treatments within a cohort on a year-by-year basis to see whether any cohort-specific differences become larger over time.<sup>128</sup>

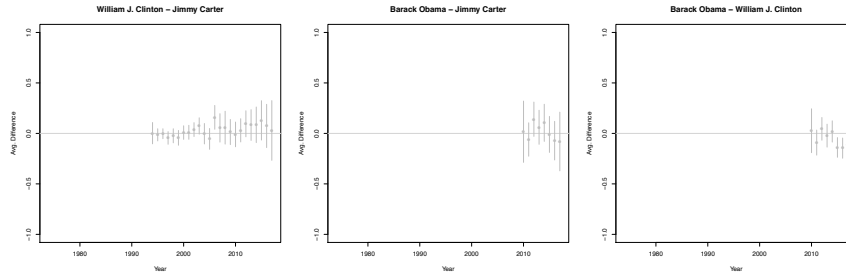
We begin by examining the intraparty differences between presidential cohorts with respect to the propensity of judges from each cohort to follow opinions written by Democratic authors and Republican authors. We examine this by looking at differences in the average number of positive treatments coming from judges within the various intraparty cohorts after adjusting for circuits. *Figure 9* displays these differences over time for the three pairings of Democratic presidential cohorts.

---

127. See, e.g., SUNSTEIN ET AL., *supra* note 7, at 113 (“By common lore . . . President Reagan was determined to ‘stock’ the federal bench with conservative judges, whereas President George H. W. Bush was significantly more moderate and President George W. Bush behaved more like President Reagan.”).

128. It is worth noting that the analysis of cohorts faces a fundamental identification problem. Namely, if there are age effects (effects specific to judges of a certain age regardless of calendar year or presidential cohort), period effects (effects that are present for all judges in a particular calendar year regardless of cohort and age), and cohort effects (effects specific to a cohort of judges regardless of the judge's age or the calendar year), it is impossible to separate these effects. See Willard L. Rodgers, *Estimable Functions of Age, Period, and Cohort Effects*, 47 AM. SOCIO. REV. 774 (1982). For instance, if cohorts are defined by a judge's birth year, then calendar year - age = cohort, which results in perfect collinearity among the right-hand-side variables of a regression. Thankfully, we do not face this severe problem. Because we are defining cohorts in terms of each judge's appointing President, and because Presidents nominate judges of multiple ages within and across multiple years, our age, calendar year, and presidential cohort variables are not perfectly collinear. Nonetheless, these variables are correlated, which creates challenges for inference.

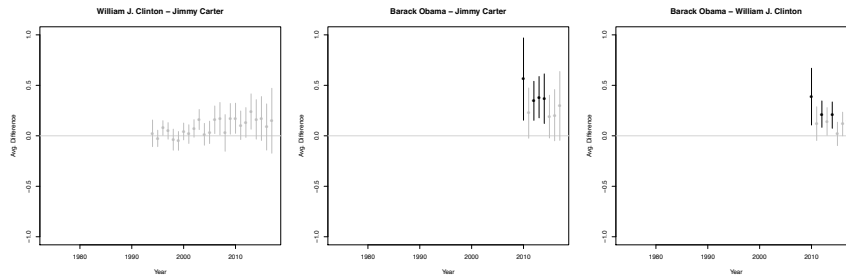
Figure 9



*Differences in the average number of positive treatments to opinions authored by Democratic appointees by Democratic presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.*

We see that there are effectively no differences between Democratic presidential cohorts in terms of their propensity to follow previous opinions authored by Democratic appointees.

Figure 10



*Differences in the average number of positive treatments to opinions authored by Republican appointees by Democratic presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.*

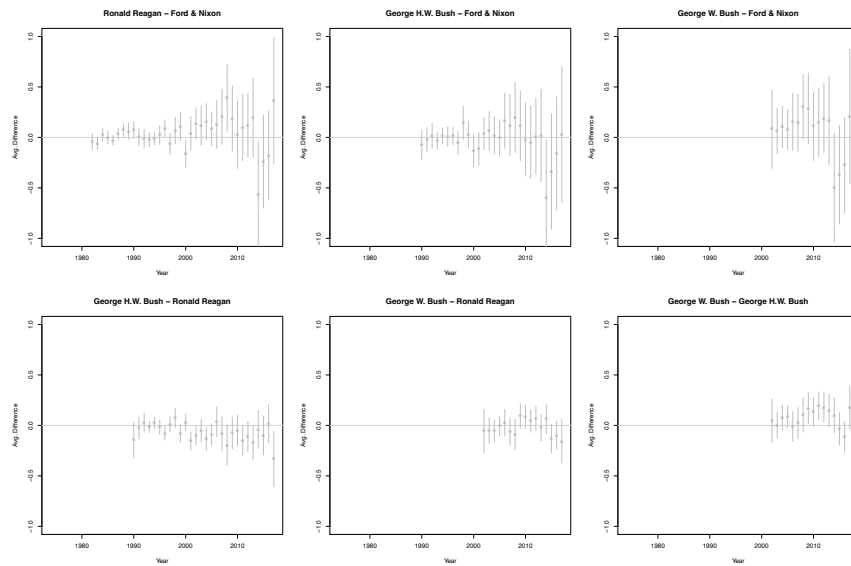
Figure 10 plots similar results comparing the extent to which different Democratic presidential cohorts follow opinions written by Republican appointees. Again, most of the year-by-year differences are not significantly different from zero. The few differences that are significant occur in the 2010s where there is some evidence that the Obama cohort is more likely to follow Republican-authored opinions than the Carter cohort and, to a lesser extent, the Clinton cohort. Note that this pattern of more



recent appointees being more likely than earlier appointees to positively treat opinions authored by judges of the opposing party is exactly the opposite of what we would expect if the increased politicization of the nomination and confirmation process is driving increases in partisan differences.

We repeat the same analysis for the Republican presidential cohorts. Again, we look at differences in the average number of positive treatments coming from judges within the various intraparty cohorts after adjusting for circuits.

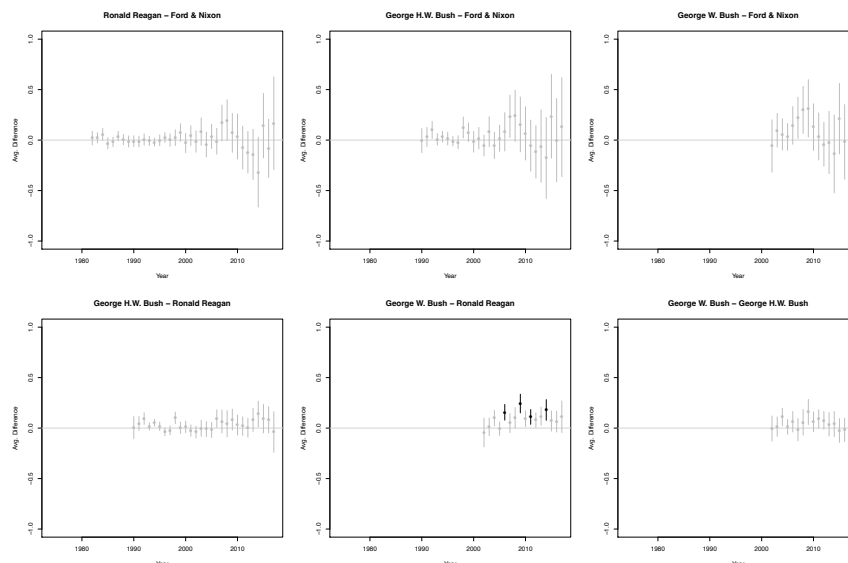
*Figure 11*



*Differences in the average number of positive treatments to opinions authored by Republican appointees by Republican presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.*

*Figure 11* displays the resulting differences over time in how six pairings of Republican presidential cohorts positively treat earlier Republican-authored opinions. As we saw with the Democratic presidential cohorts above, there is no statistically significant difference between any of the Republican presidential cohorts.

Figure 12



*Differences in the average number of positive treatments to opinions authored by Democratic appointees by Republican presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.*

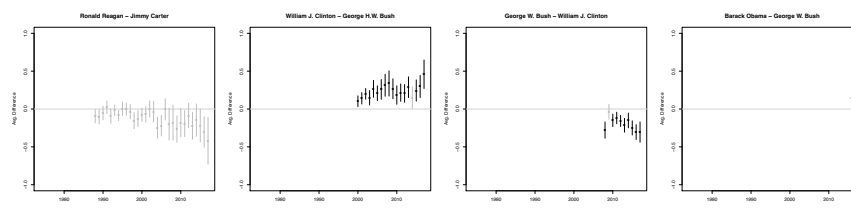
Figure 12 reports similar results for the differences in how these six pairings of Republican presidential cohorts positively treat opinions authored by Democrats. The vast majority of these year-specific differences are not statistically different from zero.

Taken as a whole, the evidence in Figures 9–12 paints a picture in which there are few, if any, differences among intraparty cohorts in how earlier opinions are substantively treated. Figures 9–12, together with Figures 1 and 2, indicate that there are large, substantively meaningful differences between Democratic and Republican appointees but very small differences among Republican appointees and among Democratic appointees. These results are consistent with a provocative claim: scholars may attribute too much significance to intraparty presidential cohorts. Partisanship matters, but within a given party the identity (and ideology) of the appointing President does not. This goes against most of the conventional wisdom on the importance of presidential cohorts within the same party.

An issue with the analyses in *Figures 9–12* is that judges are entering and leaving the court at different times. Accordingly, the over-time comparisons involve different judges. To address that potential issue, we next look at cross-party comparisons of temporally adjacent presidential cohorts that are also restricted to only those judges who served from the last year of the second presidency in each pairing to 2017. By holding the set of judges fixed while also adjusting for circuit we hope to eliminate any changes in outcomes due to changes in the pool of judges being compared.

Since we are now looking at cross-party differences in substantive treatments, we expect to see statistically significant differences. However, the real question involves the trajectory of those cross-party differences in behavior. If the differences are constant over time, with the differences between some pairs much larger than others, this would be consistent with presidential cohort effects. On the other hand, if the differences we see tend to be trending over time and largest in the most recent periods, this would be more consistent with increasing partisan polarization—driven not by presidential-cohort fueled replacement on the bench (since we are looking at the same judges throughout the period covered in the figures) but rather by factors related to increasing political polarization in the country.

*Figure 13*



Differences in the average number of positive treatments to opinions authored by Democratic appointees for temporally adjacent cross-party presidential cohorts. The judges included are only those from the cohort pair in question who served for the entirety of the period from the last year of the second presidency in each pairing to 2017. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

*Figure 13* plots the cross-party differences in the average number of positive treatments to Democratic-authored opinions for these four temporally adjacent cross-party cohorts. Note that we

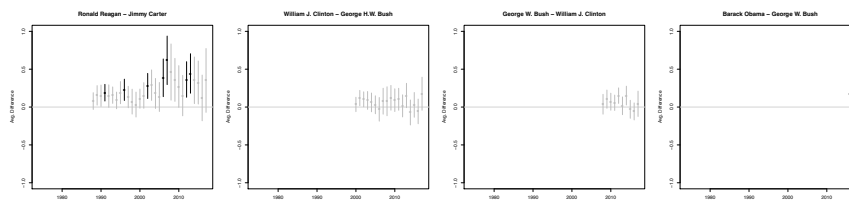
see significant differences for three pairs of adjacent cohorts. In each case, the differences are in the expected partisan direction. Note that these differences appear to be trending away from zero (at least for the pairs with more than two years of data). Further, the Clinton-H.W. Bush plot is essentially the mirror image of the W. Bush-Clinton plot for the years 2008 to 2017 when there are data for both pairs. All this is not surprising given our finding of similarities among presidential cohorts within parties: given that presidential cohorts within parties are similar, we would expect to see roughly similar trends when we compare the appointees from any two presidential cohorts from different parties. Again, this is more consistent with a partisan polarization story than a presidential cohort story.

The one exception is surprising: there are no statistically significant differences between the Reagan cohort and the Carter cohort for any year from 1988 to 2017. Given the primacy that the Reagan administration's approach to judicial appointments has in the conventional wisdom regarding the politicization of the judiciary, we might have expected large differences between the Carter and Reagan cohorts. We find none. This is all the more interesting given that the Carter administration is viewed by some scholars as the actual starting point for the politicization of judicial appointments.<sup>129</sup> That said, while there are no statistically significant differences, the point estimates do trend toward a partisan difference in the expected direction—and, importantly, this trend begins only in recent years. This is suggestive of partisan polarization rather than a presidential cohort effect; however, it is important to emphasize that this trend is not statistically significant at conventional levels.

---

129. See *supra* notes 123–125 and accompanying text.

Figure 14



Differences in the average number of positive treatments to opinions authored by Republican appointees for temporally adjacent cross-party presidential cohorts. The judges included are only those from the cohort pair in question who served for the entirety of the period from the last year of the second presidency in each pairing to 2017. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

We can also look at how these adjacent cohort pairs compare in their propensity to positively treat opinions written by Republicans. This information is displayed in *Figure 14*. In some sense, these results are even easier to interpret, as nearly all the within-pair differences are not statistically different from zero. Once again, we see very little difference between the Carter judges and the Reagan judges (only one difference is significantly different from zero prior to 2007). At that point the differences do trend up in a partisan direction and two become significant. Again, this is more consistent with partisan polarization in the 2000s than with presidential cohort effects.

To save space, we present related aggregated analyses in Appendix A4. These results compare particular Democratic presidential cohorts to all Republicans and particular Republican presidential cohorts to all Democrats. These results tell a similar story to those above in *Figures 13 and 14*.

An alternative explanation is that what we are assuming is a partisan polarization effect could be an age effect. Perhaps judges become more ideologically extreme (or at least less likely to positively treat opinions by judges of the opposing party and more likely to positively treat opinions by co-partisans) as they become older. To investigate this, we examine the cross-party difference in the average number of positive treatments to Democratic-authored opinions within judges born in the same year over time as well as the cross-party difference in the average number of positive treatments to Republican-authored opinions within judges born in

the same year over time. These analyses also adjust for circuit. The resulting plots are displayed in Appendix A3 for reasons of space. The key takeaway from this analysis is that within judges with the same birth year there are very few year-specific partisan differences in the propensity to positively treat either Democratic or Republican opinions. The within-birth-year partisan differences that do emerge are very largely in the years after 2000. These differences are not found predominantly among the earlier birth-year cohorts. Indeed, the one birth-year group that exhibits a consistently significant partisan difference in the average number of positive treatments to Democratic opinions after about 2004 is the group of judges born in 1948, which is about halfway between the earliest birth-year group (1926) and the most recent (1967). As with presidential cohorts, partisan polarization in the 2000s is a more compelling explanation than an account based on aging.

*2. The Best Explanation: An Accelerating Move Right Among Republicans but a Steady Move Left Among Democrats*

Given that explanations focusing on presidential cohorts or the aging of the judges do not hold water, how might we explain the growing magnitude of partisan differences in our data? Some form of fairly recent ideological divergence between Democratic and Republican appointees seems like the best answer. This implicates another finding that surprised us: the partisan differences in treatment of Democratic-authored opinions and of Republican-authored opinions are both statistically significant, but they are larger for Democratic-authored opinions. At first blush, this asymmetric pattern of behavior does not seem consistent with simple explanations rooted in ideology. What we might call the simple ideological account is the view that judicial behavior depends only on the relative ideological locations of current judges. According to this simple ideological account, sitting Democratic and Republican appointees should behave as mirror images of each other. We do not see that in the data, so this simple ideological account cannot be correct.

What explains this pattern of larger partisan differences for the treatment of Democratic-authored opinions than Republican-authored opinions? In general, focusing on attributes and characteristics of the later judges does little to explain this pattern. If some attribute of later judges distinguished later Republican

appointees from later Democratic appointees, presumably such a distinction would apply equally to their treatment of earlier Democratic-authored opinions and Republican-authored opinions. For example, the possible increasing partisanship of the later judges would not explain the finding: if, say, the Republicans and/or Democrats on the later panels were more partisan, then we would expect that divergence to show up in the treatment of earlier Democratic-authored opinions and earlier Republican-authored opinions.

We believe that the key variable must be some difference in the earlier Democratic-authored opinions compared to the earlier Republican-authored opinions—or, more to the point, in the later judges' perception of those earlier opinions. What aspect of these earlier opinions might be driving the result we observe? While we cannot be sure what the relevant aspect is, we do have strong reasons to exclude some things from consideration.

One might be tempted to think that the relevant difference between these earlier Democratic- and Republican-authored opinions has something to do with the types of cases heard by panels with different partisan compositions. But the judges and cases are randomly assigned to panels within a circuit and year, so there should not be any such differences within a particular circuit and year.

Nonetheless, one might wonder about the stock of earlier opinions that can be followed and whether partisan imbalances in the circuit in question might create more opportunities for selective treatment of the Democratic-authored opinions than the Republican-authored opinions. This is unlikely. Note that the median time difference between the treating and treated opinion is 5.1 years and the mean time difference is 7.5 years. In other words, most positive treatments are to opinions that were issued relatively recently. Further, recall that our estimation approach weights the data so that treating opinions that are equally likely to be authored by each side of the comparison (e.g., as likely to be authored by a Democrat as a Republican) get the highest weight.<sup>130</sup> The weights go to zero as the likelihood of each type of author becomes increasingly unequal. Finally, note that the partisan composition of most circuits changes fairly slowly over time. Taken together, these three facts

---

130. See Section IV.C.

mean that the treating cases that receive the most weight in our analysis have a stock of cases to follow that is composed of a roughly similar mix of opinions written by Democratic and Republican appointees.

Other possible differences in earlier opinions do not have much explanatory force. For instance, if we imagine that earlier Republican-authored opinions were more attractive for some nonpartisan reason (e.g., they had better reasoning), that would not explain why there is a divergence in how later judges treat Democratic-authored versus Republican-authored opinions. If later Democratic and Republican appointees recognized that better reasoning and deemed it important, then presumably they would similarly treat (well) the earlier Republican-authored opinions and similarly treat (poorly) the earlier Democratic-authored opinions.

How, then, can we explain the differences in the substantive treatment of Democratic-authored opinions versus Republican-authored opinions? We surmise that Republican-authored opinions were perceived as more equally acceptable to later Republican and Democratic appointees than Democratic-authored opinions were. The most obvious possible basis for a difference in acceptability is a difference in *perceived ideological distance of the earlier opinions from the preferences of later judges*. If later Republicans became dramatically more conservative but later Democrats did not become dramatically more liberal, then we might expect to see the pattern in our data.<sup>131</sup> Later Democrats would see significant differences between earlier Democratic-authored and Republican-authored opinions and would find earlier Democratic-authored opinions much more acceptable. Later Republicans would find earlier Republican-authored opinions slightly more acceptable than earlier Democratic-authored opinions but would find neither set particularly attractive because (by hypothesis) the later Republicans would have become so much more conservative.

The result would be that later Democrats and Republicans would respond differently to earlier Democratic-authored opinions (because later Democrats would find earlier Democratic-authored opinions quite attractive, and later Republicans would find earlier Democratic-authored opinions unattractive); but later Democrats

---

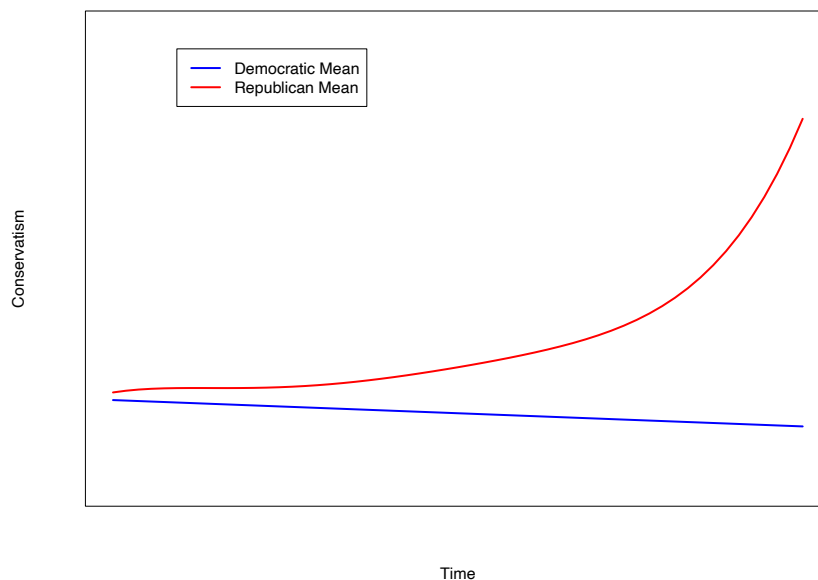
131. Note that such party-specific differences in ideological movement are consistent with accounts of ideological polarization in Congress. See, e.g., DeSilver, *supra* note 12; Lewis, *supra* note 12.



and Republicans would respond more similarly to earlier Republican-authored opinions (because later Democrats would find them unattractive, and later Republicans would find them only marginally more attractive than later Democrats do, because Republicans would be so much more conservative than their Republican predecessors that these earlier Republican-authored opinions would not be significantly more attractive). What separates this nuanced ideological account from the simple ideological account discussed above is that the ideology of earlier judges (and by extension their opinions) enters the explanation, whereas the simple account looks at only the ideology of current judges deciding whether to follow earlier opinions.

To investigate our surmise that partisan differences in ideological distance from earlier Democratic-authored and Republican-authored opinions might explain our findings with respect to those opinions, we created a simulation model. The model assumes that when deciding which earlier opinions should be followed, judges look back to earlier opinions within a temporal window with the probability of a positive treatment being a decreasing function of the distance between the treating judge's ideology and the earlier opinion author's ideology. Simply stated, the model assumes that more ideologically proximate authors are more likely to be followed. The model assumes that Democratic appointees, on average, have different ideologies from Republican appointees. It also assumes that these ideological differences change over time. Specifically, it assumes the average ideology of a Republican appointee is distinct from the average ideology of a Democratic appointee and that Republican appointees are becoming exponentially more conservative over time while Democratic appointees are trending in the opposite direction in a linear fashion. *Figure 15* displays this hypothetical ideological divergence.

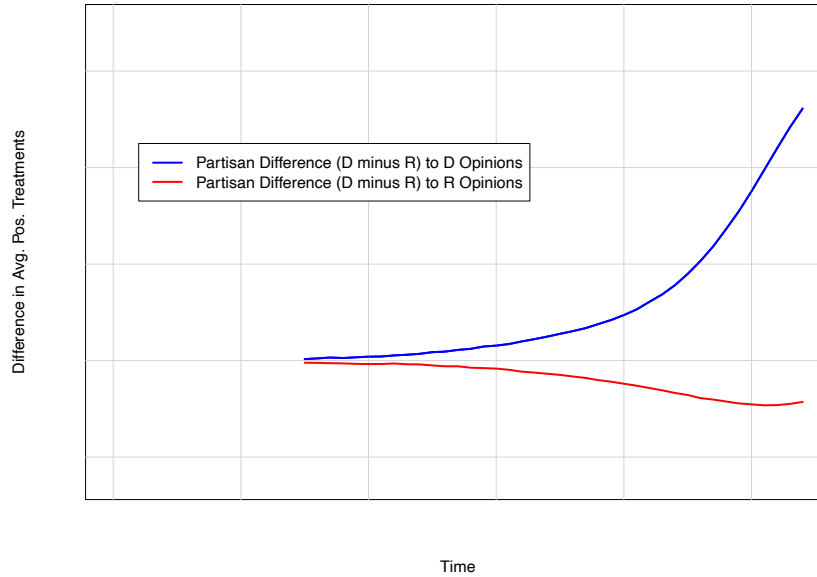
Figure 15



*Hypothetical change in the average ideological location of Democratic and Republican appointees over time. This is the divergence that is posited in the simple simulation study discussed in the text. The key point that drives the results is the accelerating change among Republican appointees compared to the steady change of Democratic appointees.*

Putting these assumptions together and letting our simulated judges make simulated decisions to follow opinions produces the partisan differences in positive treatments to Democratic-authored opinions and Republican-authored opinions depicted in *Figure 16*. Note that the partisan difference in the propensity to follow Democratic-authored opinions increases rapidly over time. The partisan difference in the propensity to follow Republican-authored opinions grows more slowly than the partisan difference to follow Democratic-authored opinions and eventually stabilizes. This pattern largely mirrors the average controlled differences by author partisanship over time in *Figure 2*.

Figure 16



*Partisan differences in positive treatments of Democratic-authored opinions and Republican-authored opinions produced by the simulation model discussed in the text.*

What is the intuition behind these results? The crucial step is the assumed accelerating movement in a conservative direction among Republican appointees compared to a steady movement in a liberal direction among Democratic appointees. The result is that Republican-authored opinions decided considerably earlier are roughly halfway between the positions of current Democratic appointees and current Republican appointees. That is, such temporally distant Republican opinions are roughly equally attractive to current Democratic appointees and Republican appointees – although more recent Republican opinions will still be substantially more attractive to Republicans than to Democrats. This will attenuate the partisan differences in the propensity to follow Republican opinions. By contrast, because Democratic appointees are not trending to the left at an increasing rate, there will be much less attenuation of partisan differences in the propensity to follow Democratic-authored opinions. If Republican appointees are becoming exponentially more conservative over time and Democrats' move to the left is only linear, then Republican appointees will be only mildly more positive about older

Republican as opposed to older Democratic opinions, whereas Democrats will have a stronger preference for Democratic over Republican opinions. If so, then we would expect exactly the results that the model suggests in *Figure 16* and that we find in this Article.

None of this is to say that Republican appointees have become, in some objective sense, more ideologically extreme than Democratic appointees. The starting location prior to the ideological divergence is arbitrary. It is consistent with a world where both Democratic and Republican appointees start off on the extreme left in some objective sense and the Republican appointees are tracking back to moderate positions at an accelerating pace while Democratic appointees move to more extreme left positions at a steady rate.

So our data suggest that there has been a rapidly accelerating divergence in the relative conservatism of appeals court judges across parties and that this accelerating divergence is driven mainly by an accelerating move to the right by Republican appointees. While we did not set out to provide direct measures of judicial ideology for the federal courts of appeals, our analysis does provide indirect evidence of such an accelerating divergence in judicial ideology, driven primarily by an accelerating rightward move by Republican appointees. Absent a more compelling alternative explanation, which we have not been able to articulate, we believe this accelerating divergence is the best explanation for the differences we find in the treatments of Democratic-authored opinions versus Republican-authored opinions.<sup>132</sup>

---

132. The recent rise in partisan differences raises a possibility suggested by our colleague Maggie Lemos: Could it be that there was a significant broadening in access to Lexis and Westlaw immediately preceding the rise, such that judges would have newly gained access to these vast online databases and thus were able, for the first time, to comprehensively look for earlier opinions written by ideologically compatible judges? The answer is no, for two reasons. First, Lexis and Westlaw sent us chronologies of their availability, and both were widely available to judges by the early 1980s, with full coverage of circuit court opinions. Second, this could not explain the recent divergence in Republican and Democratic appointees' treatments of earlier Democratic-authored opinions but not earlier Republican-authored opinions. Insofar as the availability of Lexis and Westlaw allowed Democratic appointees to do a better job of finding Democratic-authored opinions that they could follow, or allowed Republican appointees to do a better job of finding ways not to follow Democratic-authored opinions, we would expect to see a similar pattern with respect to Republican-authored opinions – and we do not.

### A. Racial Differences

Our investigation of the relationship between a judge's race and positive treatments of opinions by Black and Hispanic authors reveals patterns that are fairly similar to our findings on partisanship, particularly for opinions by Black authors. And our results are not consistent with the solidarity effect that we posited in section III.C.

Starting with opinions by Black authors, *Figure 3* indicates that across the entirety of the time period Black judges are more likely than White judges to follow opinions written by Black co-partisans. The differences are statistically significant and comparable in size to the partisan differences discussed above. Breaking the data into four time periods in *Figure 4*, we see progressions over time that resemble those for partisan differences. First, the magnitude of racial differences increases. Second, the increase over time is greatest for ideologically salient and race-salient published opinions. And the final period presents a particularly sharp increase. In that final period, the differences are statistically significant and substantively meaningful, and they are particularly large for opinions authored by Black Democratic judges.

The White-Hispanic differences we see are more muted than the White-Black differences. When we pool the data over the full time span of the study, the differences are statistically significant (except for race-salient opinions by Democratic authors) but substantively small. The differences are smaller, and thus less meaningful, for opinions by Hispanic Democratic authors. It is particularly interesting to note that the difference between how White and Hispanic Democratic opinion authors treat opinions written by Hispanic Democrats in race-salient cases is not significantly different from 0—suggesting that racial/ethnic identity is not driving the White-Hispanic differences (at least within Democrats).

Turning to the four time periods, for Hispanic Republicans the clearest pattern is that the differences increase over time. As for different types of cases, there are relatively modest differences in the point estimates (and all are well within the confidence intervals). For opinions by Hispanic Democrats, the differences broken into different time periods are quite small and generally are not statistically significant. There is a modest upward trend in the point estimates over time for ideologically salient and race-salient

cases, but given the lack of statistical significance we do not draw any conclusions from it.

Relatedly, as we noted above, the relatively small number of Black (and the even smaller number of Hispanic) judges means that our results are based on fewer cases from an idiosyncratic selection of circuits with much more weight given to recent cases. Because 71.2% of *Shepard's* treatments occur within circuits, comparisons depend on sufficient numbers of judges with the relevant attribute in a given circuit, and there are fewer circuit-year combinations that have Black or Hispanic judges than have female judges. As we also noted above, our decision to focus on average controlled differences with overlap weights (and conditioning on circuit, year, and circuit-year combinations to avoid spurious results) further serves to downweight data from many circuits and years. To be clear, this downweighting of data from circuit-years where good apples-to-apples comparisons are not available is a strength of our research design, but it does mean that our results need to be interpreted with care.

The fact that we see intraparty racial differences is notable because it indicates behavioral differences among judges of the same party along racial lines. The Black-White racial differences (particularly for Black Democratic authors, who are more numerous than Black Republican authors) are particularly striking: Black Democratic appointees significantly diverge from their White Democratic counterparts in their treatment of cases written by Black Democratic authors, and to a lesser extent Black Republican appointees differ from White Republican appointees in their treatment of opinions by Black Republican authors, even though in both cases we are comparing co-partisans' treatment of their fellow partisans' opinions. A possible explanation for the smaller differences for Hispanic Democratic judges relative to White Democratic judges is that Hispanic Democratic judges may not be appreciably more liberal than White Democratic judges.<sup>133</sup> Given the small number of Hispanic Republican judges, the presence of statistically significant results for them is surprising. But the small number of such judges also means that our results reflect the decisions of a relatively small number of judges, giving us less

---

133. See Sen, *supra* note 15, at 394 tbl.4 (finding, based on preconfirmation campaign contributions, that Hispanic Democratic judges are ideologically closer to White male Democratic judges than are female or Black Democratic judges).

confidence with respect to Hispanic Republican judges than for Black judges that we have uncovered an important pattern.

So what to make of these intraparty racial differences, particularly for opinions by Black authors? The data are not consistent with the solidarity effect that we posited in section III.C. As we noted there, a solidarity effect would presumably either be constant or decrease over time. Either Black judges would prefer to follow opinions issued by another Black judge, or trailblazing Black judges might prefer to support other Black judges when they were very few in number and struggling to be accepted. The first possibility would lead to a constant effect and the second would likely lead to a diminishing effect in the most recent time period, or perhaps also a constant effect (insofar as current Black judges see themselves as trailblazers).<sup>134</sup> These theories are not consistent with the results we find, where the observed differences rise dramatically in the most recent time period. Rather than evidence for solidarity effects, we instead see the trend that we also saw with respect to partisan differences.

By contrast, the general ideological and more policy-specific hypotheses we discussed in section III.B seem to fit well with our results. We think that the data better fit an explanation that places more of the weight on broad ideology than on expertise and lived experiences, for two main reasons. First, recall that race-salient cases are a subset of ideologically salient cases, because the categories that researchers have identified as race-salient have also been identified as ideologically salient.<sup>135</sup> The key difference is that the subset of race-salient cases, in addition to reflecting ideology generally, may reflect differences in lived experiences. The subset of ideologically salient cases that does not include race-salient cases is a purer measure of ideology (i.e., without the additional element of expertise and lived experiences). Because this subset of ideologically salient cases excludes the cases for which lived experiences have been posited as relevant, the large difference for this subset indicates that ideology is playing a large role. The fact that the differences in race-salient cases are statistically indistinguishable from the differences in the subset of ideologically

---

134. See *supra* notes 93–94 and accompanying text.

135. See *supra* note 82. Gender-salient cases are also a subset of ideologically salient cases. *Id.*

salient cases that exclude race-salient cases suggests that lived experience is, at most, playing a modest incremental role beyond that played by general ideology.

Second, the rise in the differences between Black and White judges follows the same basic timing pattern as the differences based on partisanship: for both race and partisanship, the differences rise over time, particularly in the most recent time period. As we noted above, what the category of ideologically salient cases is designed to measure is similar to what partisanship measures.<sup>136</sup> In both cases, we are measuring ideology broadly understood. The striking similarity of the rise in partisan differences and differences in ideologically salient cases suggests that there is a commonality between them in the form of broad ideology.

It is of course possible that the rise in partisan differences happened to coincide with a rise in intraparty racial differences in relying on expertise and lived experiences, but we put less weight on such a possibility. We can think of three reasons why the two might coincide. First, it might be that race-salient cases are connected to partisanship such that the rise in both is not just a coincidence. Insofar as the race-salient category measures differences in expertise and lived experiences, it is not clear what that connection would be. The theory behind the emphasis on expertise and experiences, after all, is that it is different from ideology. This relates to a second possibility: it might be that the category of race-salient cases in fact measures ideology akin to partisanship, rather than measuring expertise and lived experiences, such that both reflect a rise in ideological differences. But, if so, then this is just an ideological story after all: the category of race-salient cases would in fact be measuring ideology broadly understood. Third, it might be that the rise in partisan differences and in intraparty differences for race-salient cases is just a coincidence. We cannot rule out that possibility, of course. But given the connection between what partisanship measures and broad ideology (as measured by ideologically salient cases), we think a correlation between partisanship and broad ideology is more likely than the happenstance of two unrelated phenomena showing the same pattern over the same period of time.

---

136. *See supra* text accompanying notes 84–86.



The fact that the recent rise in intraparty racial differences accords with the rise in partisan behavior does not mean that we have proved that ideology is the explanation for racial differences—or partisan differences, for that matter. For both race and partisanship, there may be a non-ideological explanation for our results. But the most obvious non-ideological explanation is a solidarity effect, and our data are not consistent with such an effect.

We think the best explanation is that this increase in racial differences is indeed akin to the increase in partisan differences, reflecting the same trend—greater ideological polarization over time. Insofar as race may reflect components of ideology that partisanship does not reflect, increasing racial differences likely reflect aspects of ideological polarization that partisanship does not capture. Simply stated, it appears that judges' substantive treatments of their colleagues' opinions are reflecting greater polarization that is manifested in both partisan and intraparty racial differences.

The implications of these results relate to broader issues of race and ideology. As we noted in section II.B.ii, previous researchers have attempted to identify differences in judicial behavior between Black and White judges by examining their votes in a variety of case types, and they have found differences in relatively few categories of race-salient cases.<sup>137</sup> The null results in most categories of cases have been interpreted as indicating that there are no significant differences in the behavior of Black and White judges outside of this small number of race-salient categories.<sup>138</sup> But, as we have noted in this Article, votes are only one form of judicial behavior. Our dataset allows us to carefully examine racial differences in substantive treatments, and our findings indicate that there are substantively meaningful racial differences in behavior across a wide range of cases. By looking at substantive treatments, we have found differences that others have missed.

---

137. See *supra* notes 62–66 and accompanying text.

138. See, e.g., Burbank & Farhang, *Pleading Decisions*, *supra* note 51, at 2246 (canvassing the null results outside of sex discrimination in employment and concluding: “In sum, the employment discrimination studies revealing gender differences in Court of Appeals decision-making are islands in a sea of null results.”).

Our results suggest that Black judges ideologically differ from their White co-partisans.<sup>139</sup> And those differences are consistent with the idea that Black judges are, on average, to the left of White judges. Party is a measure of ideology. It appears that race can refine that measure.

The results for differences between Hispanic and White judges are less meaningful, for two reasons. First, the results for Democratic appointees are small and not substantively meaningful (and often statistically insignificant). Over the full 1974–2017 span of our study, the White-Hispanic average controlled differences are always substantively small. Further, the White-Hispanic average controlled difference within the subset of race-salient cases is not different from 0. And when we break down the results into our four time periods, most of the results for Hispanic Democrats lack statistical significance. All of this points to limited influence of Hispanic identity on Democratic appointees' following of earlier opinions.

Second, the White-Hispanic differences we report are based on limited data. Our decision to focus on average controlled differences with overlap weights that adjust for circuit, year, and circuit-year means that circuit-year combinations that have no Hispanic judges do not enter into our analyses. Even circuit-years with some Hispanic judges may get very little weight if there is, say, only one Hispanic judge in that circuit-year. In the time span of our study, nine circuits never had a Hispanic Republican judge, and only two had more than one. Practically, this means that the White-Hispanic comparisons we report are best thought of as very localized comparisons of White and Hispanic judges in those (relatively few) circuit-years with multiple Hispanic judges and White judges. Indeed, our results with respect to White and Hispanic Republican appointees derive, to a very large extent, from three circuits—the First, Fifth and Ninth. Our results reflect the universe of meaningfully comparable White and Hispanic Republican appointees, but that universe is a very limited one.

Although we place little weight on these results, the pattern over time for Hispanic Republicans is fairly similar to those for Black judges (and, as we shall see, for female Republican judges):

---

139. And that Hispanic Republicans ideologically differ from White Republicans, and are, on average, to the left of White Republicans.

differences become larger over time, with the differences in the 2007–2017 time period being quite large. The confidence intervals for all the types of cases in the last time period have substantial overlap, suggesting that the differences we find do not depend appreciably on the subset of cases examined. But these findings can be taken as suggestive of substantively meaningful differences between White and Hispanic Republican judges. And, as with differences between Black and White judges, these differences are consistent with the proposition that Hispanic Republicans are, on average, less conservative than White Republicans.

### *C. Gender Differences*

Turning to gender differences, we see in *Figure 7* that, pooling the data across the entirety of the 1974–2017 time period, female judges are more likely than male judges to follow opinions authored by co-partisan women. The differences are statistically significant, but they are not large and not substantively meaningful. Breaking the data into four time periods in *Figure 10*, the results remain substantively small, but we see different patterns for Republican and Democratic judges. For Republican judges, there are increasing gender differences over time, particularly between the second and third time periods, and particularly for ideologically salient cases (and less so for gender-salient cases). For Democratic judges, the results are less clear: many of the results are not statistically significant, and most of the increases over time are well within the confidence intervals and thus not substantively meaningful. Interesting in this regard is the absence of statistical significance for the ideologically salient and gender-salient cases for three of the time periods. The analogous measures for A) gender differences for Republicans, B) Black-White differences, and C) Hispanic-White differences for Republicans were generally statistically significant (and changed the most over time). But for Democratic gender differences, the main statistical significance occurs in all cases and all published cases.

So, what should we make of these gender results? If we focus on Republican judges, the pattern is fairly similar to the patterns for partisan differences, Black-White differences, and Hispanic-White differences for Republicans (though the gender differences are considerably smaller): the differences increase over time and are greatest for the most ideologically salient cases.

This increase over time for Republican judges is not consistent with the solidarity effects we outlined in section III.C. As we noted in the previous section, a solidarity effect would presumably either be constant or decrease over time. The observed increase in gender differences over time, like that for Black-White differences, and Hispanic-White differences for Republicans, suggests that something other than solidarity effects is the explanation.

Also notable is the fact that the gender difference for Republican judges is not higher for gender-salient cases than for ideologically salient cases or for ideologically salient cases excluding gender-salient cases. Indeed, the latter two categories are statistically significant for all three time periods with sufficient data, and the gender-salient cases are significant only for the 1996–2006 time period. So the case for placing most of the weight on broad ideology rather than expertise and lived experiences is even stronger for the gender differences for Republicans than for Black-White differences. All the same points we made in section VI.B apply, and here the lack of statistical significance for gender-salience in two of the time periods further weakens the case for the gender-salient cases doing much work. All this suggests that the gender differences are better explained by ideological differences, as opposed to the lived experience differences that the gender-salient cases are designed to measure.

The discussion in this section so far has focused on gender differences among Republicans. But when we turn to gender differences among Democrats, the picture becomes much murkier. As we noted above, if we pool the data among all four time periods, there are statistically significant, but small, gender differences. When we break them up into the four time periods, there is a fairly clear pattern of increases in gender differences for all cases. And there are statistically significant differences in all three time periods with a meaningful number of female judges for all and all published cases, and in two of the time periods for ideologically salient cases minus gender-salient cases, but in only the 1996–2006 period for ideologically salient and gender-salient cases (and we do not see a consistent rise in gender differences for ideological or gender-salient cases). The point estimate for gender-salient cases in the 1996–2006 period is a bit higher than for ideologically salient and ideologically salient minus gender-salient cases, but all are well within the confidence intervals. Given the weakness of these

results, it would be a stretch to attempt to draw any meaningful conclusions about the role of broad ideology versus expertise and lived experiences, or the role of any solidarity effect. There simply are not enough data to draw strong conclusions. So if we consider the gender differences for Democrats in isolation, the results are not very substantively meaningful: there are statistically significant differences over the entire time period, but when we break the results into time periods the results are more often statistically insignificant or small and not substantively meaningful.

The murkiness of the gender results for Democrats has implications for the gender results overall. At a minimum, combining the Democratic and Republican gender results means that we cannot identify meaningful consistent patterns about the role of gender (whereas we can identify such patterns for partisanship and for Black-White differences). There is no requirement that we consider the Democratic and Republican gender results together, but we did not hypothesize different gender differences for Democrats versus Republicans.

The most we can say is that the gender results for Democratic judges are not particularly meaningful, but the results for Republicans are suggestive of ideological differences along gender lines akin to those that we saw for racial differences: gender (for Republicans) seems to reveal a growing ideological separation over time. That is, for Republicans gender may illuminate ideological differences that party membership does not reveal. And those gender differences are consistent with the proposition that female Republican judges are, on average, less conservative than male Republican judges.

An ideological explanation for the gender differences among Republicans suggests a possible refinement to the partisan explanation at the end of section VI.A. Recall that our explanation for greater partisan differences for opinions by Democratic judges than for Republican judges involved great Republican rightward acceleration compared to steadily leftward movement among Democrats. The gender differences among Republicans raise the possibility that all Republicans did not move equally quickly in a conservative direction, and that in fact male Republicans accelerated rightward more rapidly than did female Republicans. The Black-White Republican differences suggest that White Republicans accelerated in a conservative direction relative to Black

Republicans. And although the number of observations is smaller and thus the data are less informative, the change over time for Hispanic-White Republican differences is consistent with White Republicans moving rightward more rapidly than Hispanic Republicans. The data are only at best suggestive. But they do provide some mild support for the proposition that an accelerating rightward movement among White male Republican judges had the greatest contribution to the recently rising differences we find in this Article.

#### CONCLUSION

If we had written this Article twenty years ago, we would have been able to tell a heartening story about ideology – that it was not a significant factor in judges’ decisions to follow earlier opinions. But today the story is different: In recent years, polarization has risen along party lines, within both parties for Black versus White judges, and within Republicans for Hispanic versus White and female versus male judges. And these differences are greatest in the most ideologically salient cases. The best explanation is ideological polarization between and within political parties.

And there is an interesting wrinkle: The rising partisan polarization we find in judges’ treatment of earlier opinions is greater when a Democrat wrote the earlier opinion. That pattern is consistent with Republican appointees rapidly becoming more conservative and Democrats appointees more slowly becoming more liberal.

More broadly, our data indicate that judges were not particularly ideological in their substantive treatments of earlier opinions in earlier decades, but in the twenty-first century have become increasingly so, with attendant impacts on the shape of legal doctrines. Judges, it seems, are subject to the polarization that affects the rest of us.

## APPENDIX

## A1. CASE COUNTS ON IDEOLOGICALLY SALIENT TOPIC SUBSETS

No.	Topic	Number of Cases
1	Search & Seizure	25,699
2	Search Warrants	9,702
3	Right to Counsel	9,528
4	Effective Assistance of Counsel	16,221
5	Prisoner Rights	14,049
6	Habeas Corpus	32,707
7	Capital Punishment	3,244
8	Immigration Law	33,657
9	Controlled Substances	17,029
10	Miranda	4,893
11	Guilty Pleas	18,491
12	Civil Rights	44,359
13	Title VII	8,430
14	Affirmative Action	1,495
15	Labor & Employment Law	47,528
16	Discrimination	26,114
17	Employment Law Discrimination	23,124
18	Disability Discrimination	5,142
19	Racial Discrimination	4,492
20	Sex Discrimination	3,332
21	Sexual Harassment	2,182
22	Sexual Orientation	138
23	Workers' Compensation Workers Compensation	11,858
24	Wage & Hour	3,027
25	Collective Bargaining & Labor Relations	14,671
26	Abortion	298
27	Voting Rights	452
28	Campaign	3,424
29	Establishment Clause	147
30	Environmental	10,483
31	Clean Water Act	907

32	Property Rights	1,701
33	Piercing the Corporate Veil	866
34	State Sovereign Immunity	3,110
35	Abrogation of Immunity	487
36	NEPA	4,844
37	NLRB	3,311
38	Campaign Finance	1,176

*List of ideologically salient topics and the total number of cases in each topic category. A given opinion typically has more than one topic (including one or more topics that are not ideologically salient).*

Table A2

No.	Topic	Number of Cases
1	Civil Rights	44,359
2	Title VII	8,430
3	Affirmative Action	1,495
4	Labor & Employment Law	47,528
5	Discrimination	26,114
6	Employment Law Discrimination	23,124
7	Disability Discrimination	5,142
8	Racial Discrimination	4,492
9	Sex Discrimination	3,332
10	Sexual Harassment	2,182
11	Sexual Orientation	138
12	Abortion	298

*List of gender-salient topics and the total number of cases in each topic category. Gender-salient topics are a subset of ideologically salient topics.*



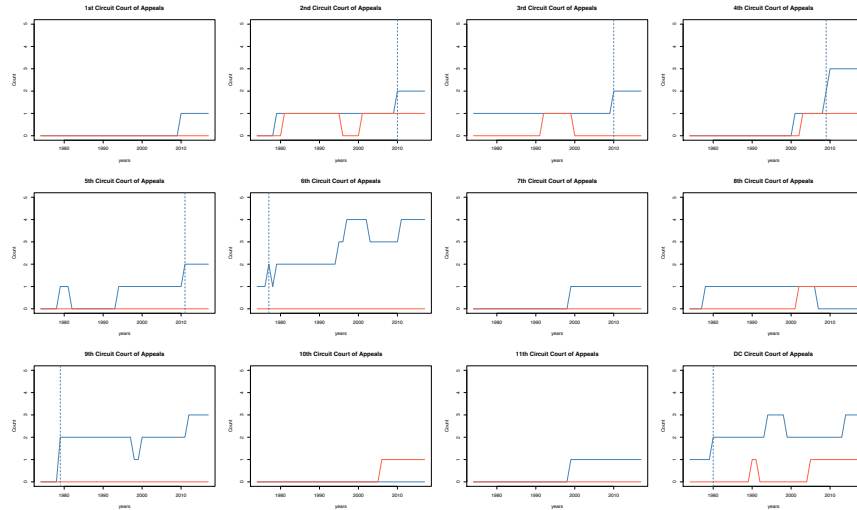
*Table A3*

No.	Topic	Number of Cases
1	Search & Seizure	25,699
2	Search Warrants	9,702
3	Right to Counsel	9,528
4	Effective Assistance of Counsel	16,221
5	Prisoner's Rights	14,049
6	Habeas Corpus	32,707
7	Capital Punishment	3,244
8	Immigration Law	33,657
9	Controlled Substances	17,029
10	Miranda	4,893
11	Guilty Pleas	18,491
12	Civil Rights	44,359
13	Title VII	8,430
14	Affirmative Action	1,495
15	Discrimination	26,114
16	Employment Law Discrimination	23,124
17	Racial Discrimination	4,492
18	Voting Rights	452

*List of race-salient topics and the total number of cases in each topic category. Race-salient topics are a subset of ideologically salient topics.*

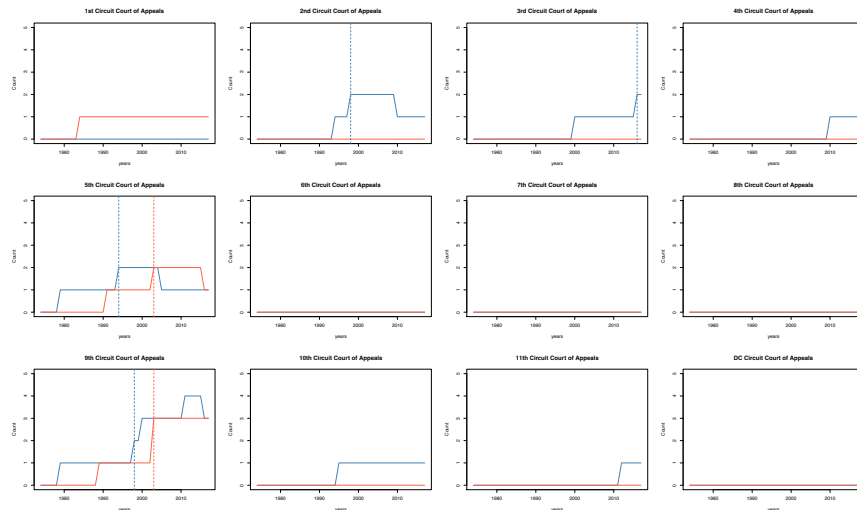
A2. NUMBER OF BLACK, HISPANIC, AND FEMALE JUDGES BY CIRCUIT AND YEAR

Figure A1



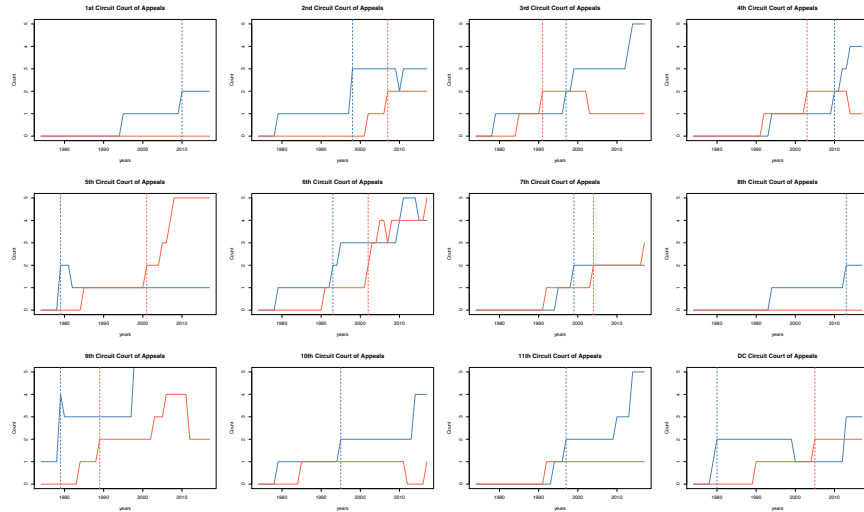
Total number of Black judges for each court and year by partisanship. Red and Blue dotted vertical lines are added to the years that the given circuit has two Black Republican judges and Black Democratic judges respectively. Data from FJC.

Figure A2



Total number of Hispanic judges for each court and year by partisanship. Red and Blue dotted vertical lines are added to the years that the given circuit has two Hispanic Republican judges and Hispanic Democratic judges respectively. Data from FJC.

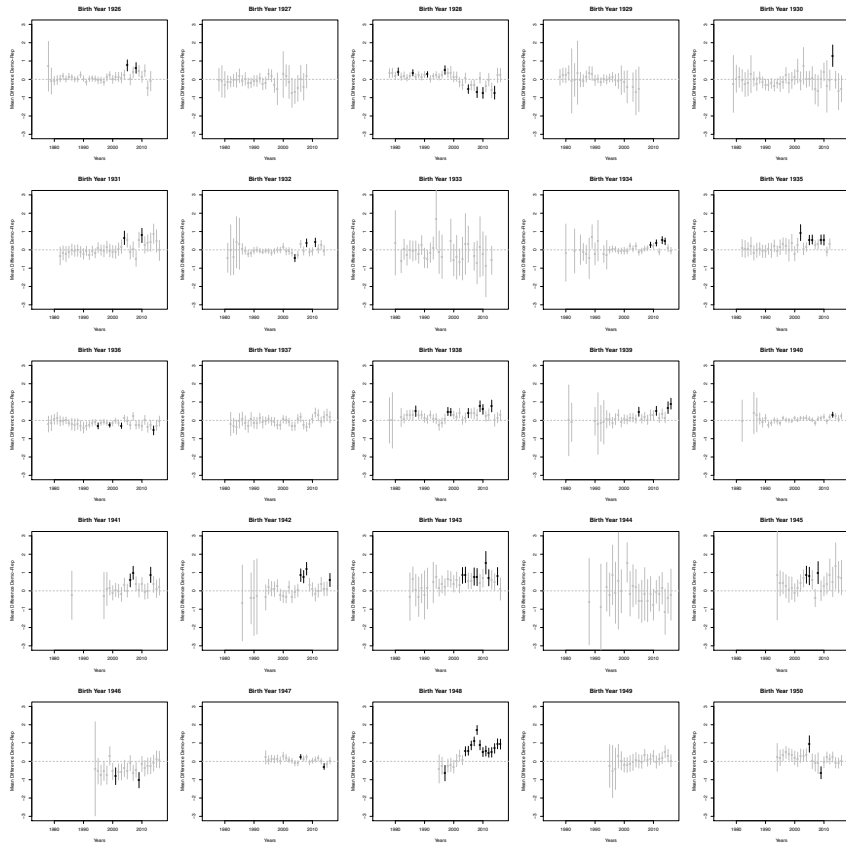
Figure A3



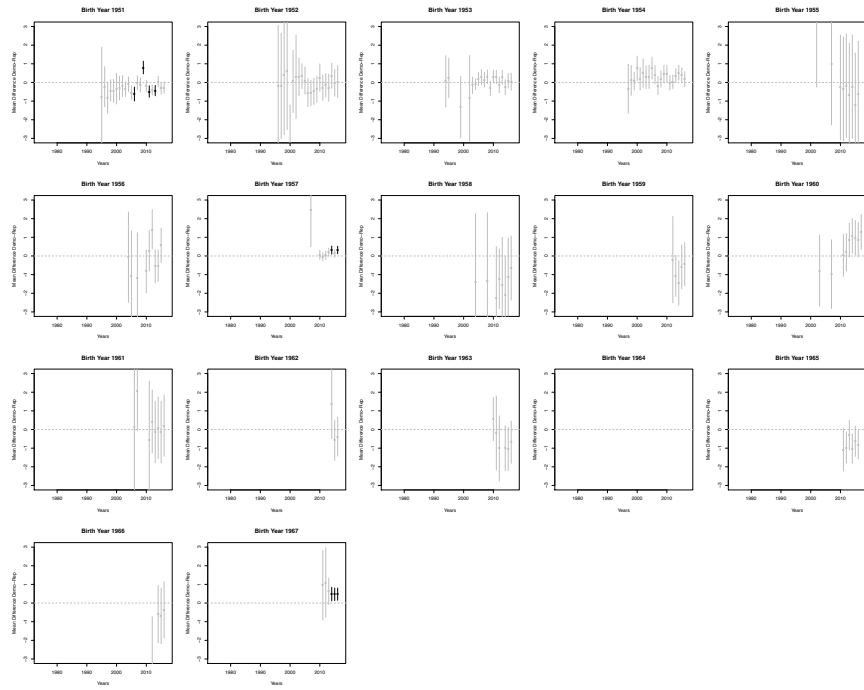
Total number of female judges for each court and year by partisanship. Red and Blue dotted vertical lines are added to the years that the given circuit has two female Republican judges and female Democratic judges respectively. Data from FJC.

A3. AGE OF JUDGES AND THE PARTISAN DIFFERENCES

Figure A4

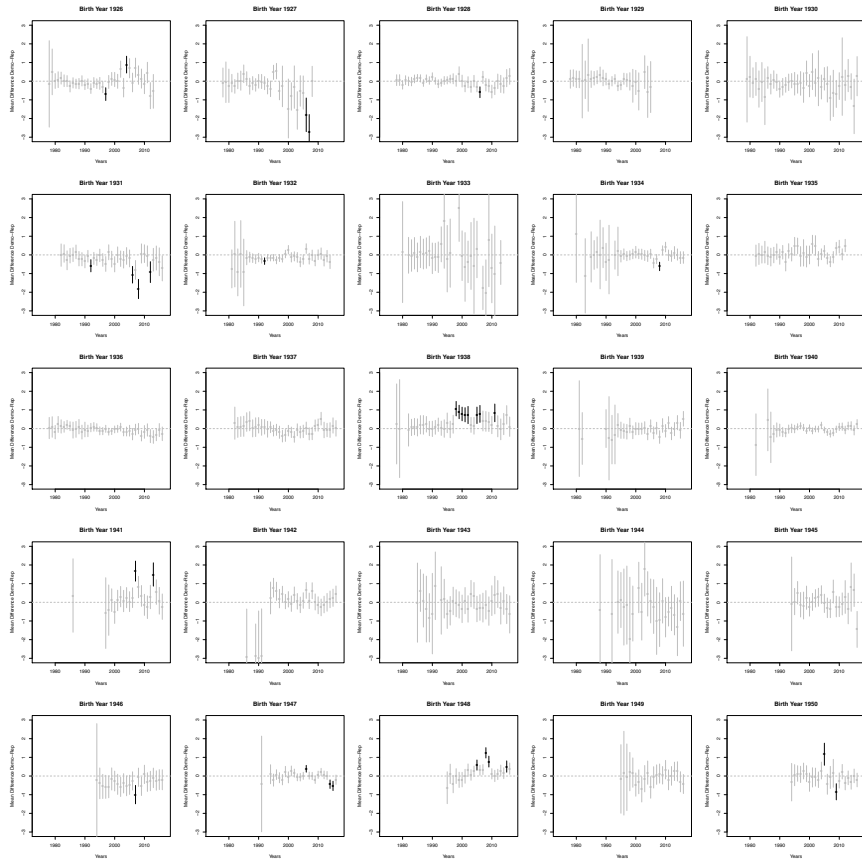


(continues on next page)

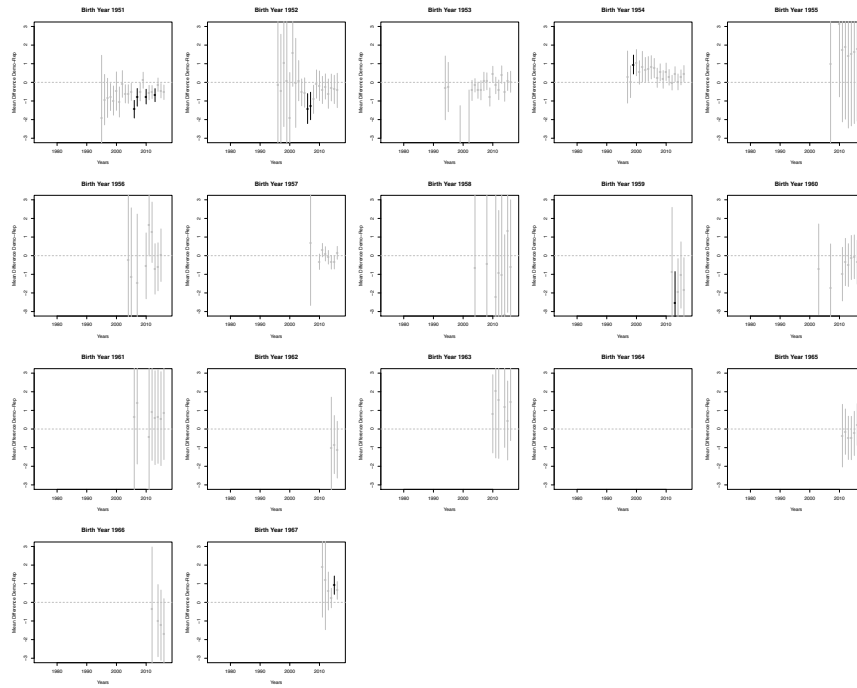


Effect of birth year of appointees on the positive treatment of opinions written by Democratic authors. Each point is a predicted difference between Democratic authors and Republican authors in the propensity to follow opinions by Democratic authors. To obtain predicted values, we fit an OLS regression of the count of positive treatment to Dem-authored opinions on the interaction of author's party and year. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

Figure A5



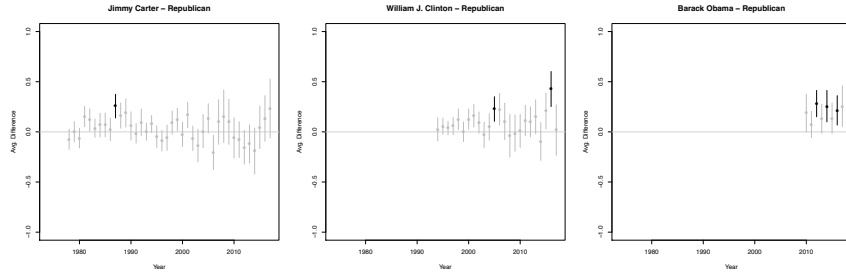
(continues on next page)



Effect of birth year of appointees on the positive treatment of opinions written by Republican authors. Each point is a predicted difference between Democratic authors and Republican authors in the propensity to follow opinions by Republican authors. To obtain predicted values, we fit an OLS regression of the count of positive treatment to Rep-authored opinions on the interaction of author's party and year. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

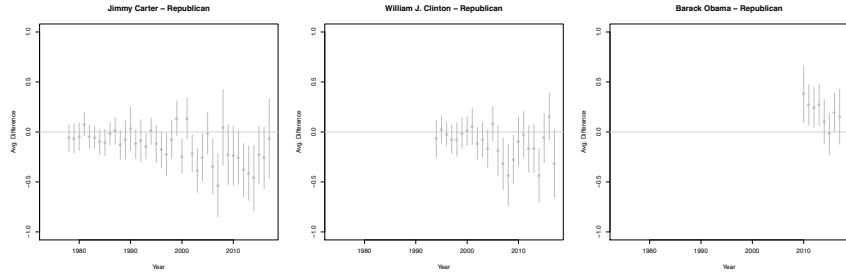
A4. PRESIDENTIAL COHORTS AND PARTISAN DIFFERENCES

Figure A6



Differences in the average number of positive treatments to opinions authored by Democratic appointees for Democratic presidential cohorts compared to all Republican presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

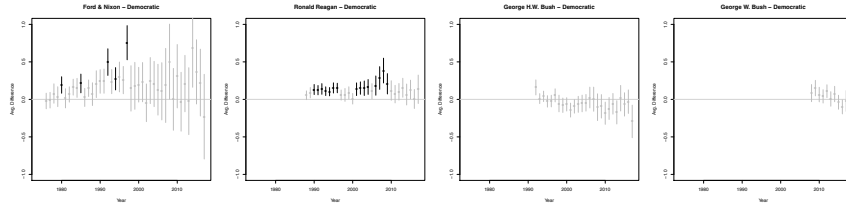
Figure A7



Differences in the average number of positive treatments to opinions authored by Republican appointees for Democratic presidential cohorts compared to all Republican presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

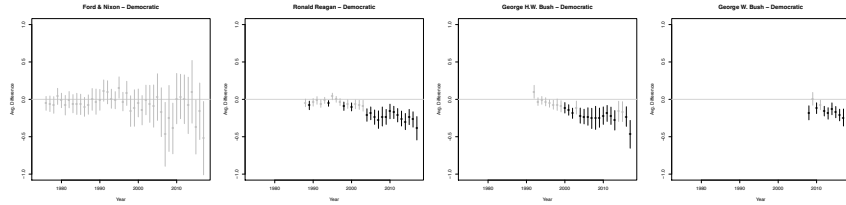


Figure A8



Differences in the average number of positive treatments to opinions authored by Republican appointees for Republican presidential cohorts compared to all Democratic presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

Figure A9



Differences in the average number of positive treatments to opinions authored by Democratic appointees for Republican presidential cohorts compared to all Democratic presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.