Is Justice Really Blind? Race and Reversal in US Courts

Maya Sen

ABSTRACT

I use two newly collected data sets to demonstrate that black federal district judges are consistently overturned on appeal more often than white district judges, with a gap in reversal rates of up to 10 percentage points. This gap is robust and persists after taking into account previous professional and judicial experience, educational background, qualification ratings assigned by the American Bar Association, and differences in appellate panel composition. In total, I find that approximately 2,800 additional cases authored by black judges have been reversed over the last 12 years. This study is among the first to explore how higher-court judges evaluate opinions written by judges of color, and it has clear implications: despite attempts to make the judiciary more reflective of the general population, racial disparities in the legal system appear to persist.

1. INTRODUCTION

In 1961, Illinois state judge James Parsons was at his summer home when he got a call that changed his life. The call was from President John F. Kennedy, and over the course of it, Kennedy asked Parsons if he would accept a federal judgeship at the US District Court for the Northern District of Illinois. As Parsons later recalled, "I said, 'As a former naval officer, aye, aye sir,' and he said, 'Carry on'" (*Jet* 1993, p. 4). The significance of this conversation—an otherwise routine exchange between a

MAYA SEN is Assistant Professor at the John F. Kennedy School of Government, Harvard University. A previous version of this paper was awarded the 2012 Best Graduate Student Paper Prize from the American Political Science Association Law and Courts Section. I am grateful to Paul Brace, Matthew Blackwell, Adam Glynn, Lee Epstein, Anna Harvey, Jennifer Hochschild, Jonathan Kastellec, Gary King, Bethany Lacina, Clayton Nall, Richard Nielsen, Alexandra Pagano, Kevin Quinn, Joseph Ura, and participants at the University of Chicago conference A Rational Choice Approach to Judging (October 4 and 5, 2013) for helpful feedback. Thanks also to the Federal Judicial Center and to the Harvard Law Library and Westlaw staff.

[*Journal of Legal Studies*, vol. 44 (January 2015)]
 0047-2530/2015/4401-0016$10.00

president and a potential judicial nominee—was that Parsons was black, and his investiture made him the first black appointed to the US District Courts.

Thanks to jurists like Parsons, numerous men and women of color now occupy roles in the upper echelons of the judiciary, not just in state and federal courts but also in other countries and at the international level. And while social scientists have an increasing understanding of how characteristics such as race influence decision making, less well understood is how the legal system has incorporated these actors—that is, how the decisions rendered by minority and women judges have been evaluated by higher courts, whether they have been treated on equal footing, and how influential they have been. On the one hand, the increased appointment of women and minorities serves to make the judiciary more reflective of the population it serves. On the other, if these judges are more likely to be overturned, then we must consider whether more needs to be done to achieve the goals of descriptive representation in the courts.

In this paper, I examine how higher-court judges evaluate opinions written by minority judges. I leverage several new data sets that include the personal characteristics of approximately 1,500 federal district judges and their corresponding appeal and reversal rates on cases decided between 2000 and 2012. By then controlling for measures of partisanship, qualifications (including ratings awarded by the American Bar Association [ABA]), experience, and jurisdiction, I find that cases decided by black lower-court judges are up to 10 percentage points more likely to be overturned than are opinions written by similar white judges. This gap is significant and robust and appears to be particularly strong among judges appointed by Democratic presidents. I confirm these results using an existing data set of randomly selected published appeals cases compiled by Songer (2007) and Kuersten and Haire (2011). In additional results presented in the Appendix, I also present matching and sensitivity analyses showing that these results are probably not due to fundamental imbalances in the data or to omitted-variable bias. In terms of meaningful impact, this gap is not insubstantial: if blacks were reversed at whites' comparably lower reversal rates, some 2,800 cases authored by black judges would have been upheld on appeal over the last 12 years.

Although I explore several possible explanations behind this finding, the underlying mechanism is not straightforward. One possibility is that the racial gap is explained by differences in ideological views, perhaps because black judges are more liberal and are therefore overturned more,

even when appointed by the same president as comparable white judges. To test this, I examine the composition of reviewing appeals panels using data from Kuersten and Haire (2011). I find that the difference between black and white judges in terms of reversal does not vary across more or less conservative higher courts. Nor are the results driven by distinctive voting by blacks on civil rights or affirmative action issues, on which previous scholarship has suggested differentiated voting patterns (Kastellec 2013; Cox and Miles 2008). These results suggest that something more than simple ideological differences are at play; a more likely explanation is that the racial gap is driven by an amalgam of factors, possibly including differences in lower-court-judge ideology, but also possibly implicit biases of higher courts. However, the results are clear: a factor in predicting whether a judge will be reversed is, surprisingly, his or her race.

This paper proceeds as follows. I discuss theories linking race, decision making, and reversal in Section 2. In Section 3, I discuss the data, which comprise two newly collected data sets and one existing data set on US district judges. Section 4 presents the core results: black judges are indeed more likely to be reversed than white judges, and the difference is robust. In Section 5, I discuss possible explanations behind this finding, which are (1) the possibility that black lower-court judges have more, or different, kinds of cases appealed to higher courts, (2) differences in professional experience, (3) differences in ideology, (4) differences in voting patterns in certain issue areas, and (5) implicit racial bias on the part of higher courts. I conclude in Section 6 with a discussion of the limitations and implications of this research. Additional results from matching analyses are presented in the Appendix.

2. THEORETICAL FRAMEWORK

2.1. Race and Judicial Decision Making

Ever since President Jimmy Carter began nominating women and minority judges in large numbers, scholars have stressed their potential importance and focused closely on their impact. On the one hand, there is a view that simply having a diverse bench might be normatively desirable (Pitkin 1967) and that it has the potential to increase the institutional legitimacy of the courts (Scherer and Curry 2010). Another view is that descriptive representation can also be instrumentally important by bringing in viewpoints that might otherwise be unshared. Descriptive representa-

tion can therefore often (although it does not always) result in substantive representation (Krislov 1974).

Empirical studies on this topic have mostly focused on this second question—whether women and minority judges decide cases differently than their white male counterparts. For the most part, the answer to this question has been yes, but they do so in the context of substantively salient issues. To take some examples relevant to the minority judges who are the focus of this study,[1] first, Kastellec (2013) finds that black judges are more likely to vote in favor of affirmative action policies and that having blacks on an appellate panel changes how that panel votes. Cox and Miles (2008) similarly find that the addition of a black judge to a panel increases the likelihood that it will find a violation of the Voting Rights Act. Other studies find a difference in voting in related civil rights areas. For example, Pinello (2003) finds that black judges are more likely than white judges to side with lesbian, gay, bisexual, or transgender claimants, and Martin and Pyle (2000) find that black judges are more likely to rule in a liberal direction in discrimination and gender-related cases. (On this last point, however, Segal [2000] finds evidence to the contrary.) Importantly, a number of these studies have found this difference between white and black judges when the race of the parties is a salient issue. In the criminal context, Scherer (2004) finds that black judges are more likely to accept black defendants' claims of police misconduct, while Welch, Combs, and Gruhl (1988) and Gottschall (1983) find that black judges are more lenient with black defendants than are white judges (but see Spohn [1990], who finds no differences). A number of other studies have found no differences across other legal areas (Walker and Barrow 1985; Gottschall 1983).

A thread running through this literature is that differences in voting may be due to different personal and professional experiences, which in turn inform legal views. Blacks on the bench tend to be clustered in certain (oftentimes urban) districts, with a greater share having experience as public defenders, government lawyers, and law professors (Sen 2014b). In addition, a number of black judges have historically come

1. A similar literature addresses differences in voting by male and female judges. This scholarship suggests that there are differences in the way that male and female judges vote, but only in the context of gender-salient cases—for example, those dealing with sex discrimination (Boyd, Epstein, and Martin 2010; Baldez, Epstein, and Martin 2006; Peresie 2005; Segal and Spaeth 2002). Others have found little or no effect associated with a judge's gender (Manning, Carroll, and Carp 2004; Kulik, Perry, and Pepper 2003; Ashenfelter, Eisenberg, and Schwab 1995). With regard to reversal rates, I find no differences between male and female judges.

from the trenches of the civil rights movement (for example, Thurgood Marshall, Constance Baker Motley, and Matthew Perry), or black judges may have more experience as lawyers working within the criminal justice system (Scherer 2004). These different experiences could introduce or reinforce distinct attitudes about affirmative action, civil rights, and voting rights. A number of these studies examining voting differences between black and white judges do attempt to control for different political views (by way of judicial common-space [JCS] scores or other proxies for judicial ideology); however, that differences between black and white judges persist despite controlling for such measures suggests that blacks' voting transcends measurements of political ideology, particularly on civil rights issues. That is, the views of black judges differ from the views of otherwise similarly liberal whites.

However, differences in voting are only half of the story, at least regarding substantive representation. Although we know that minority judges vote differently once on the bench, we have little sense of how they are perceived or evaluated—that is, what kind of impact these judges make. The question is key for understanding the effect of descriptive representation; after all, if these judges have diminished impact on account of consistent reversal or lessened influence, then their substantive impact will be lessened as well. Underlying this concern is the possibility of implicit biases against minority actors, especially in the form of more appeals and increased reversal. In this regard, a number of studies have demonstrated implicit biases against blacks in a host of settings, including high-level business organizations (Castilla 2008; Bielby and Baron 1986; Fernandez, Castilla, and Moore 2000), law (Greenwald and Krieger 2006; Banks, Eberhardt, and Ross 2006; Bagenstost 2006; Kang 2004), public health (Krieger et al. 2010), academia (Ginther et al. 2011), employment (Bertrand and Mullainathan 2004; Fryer and Levitt 2004a), housing (Yinger 1986), and even the halls of Congress (Butler and Broockman 2011). That the same could apply to the judiciary may be problematic, but perhaps unsurprising.

Despite substantial literature in other fields, the literature on how the legal system incorporates (non-criminal-defendant) minorities is limited. Some insight comes from state-level analyses—specifically, attempts to quantify judicial performance in anticipation of judicial elections. Such judicial evaluations have been implemented in 19 states and usually involve surveys of local attorneys about judicial performance (Pelander 1998; Gill, Lazos, and Waters 2011). Gill, Lazos, and Waters (2011) find

that attorney surveys routinely award lower scores to women and minorities, even after controlling for experience and reversal rates. At the federal level, no study has looked at the comparative performance of minority or women judges or at how often these judges are overturned by higher courts. Perhaps the only measure of judicial quality comes in the form of ratings awarded by the ABA; black judges have been shown to receive lower ratings in some studies (Lott 2001; Sen 2014a), but not others (Smelcer, Steigerwalt, and Vining 2011). In addition, in public opinion, Scherer and Curry (2010) find that many perceive black judges to be more liberal.

2.2. How Race and Reversal May Be Related

In this study, I explore the judicial analogy to those outcomes explored in other implicit-bias studies: reversal by higher courts. Of the few studies examining individual-level judicial reversal rates, most agree that reversal is costly (for example, Epstein, Landes, and Posner 2013; Choi, Gulati, and Posner 2012). Reversal may result in increased workload as judges have to revisit cases, forcing them to allocate scarce resources to the task and to deal with the higher court's instructions; all of this comes with no reduction in the number of incoming cases. In addition, a higher reversal rate could bring with it reputation costs, especially as lower-court judges consider actions that could make them palatable candidates for promotion to higher courts. Thus, reversal is generally perceived as something to be avoided (Choi, Gulati, and Posner 2012). As noted by some, however, higher reversal rates could be a sign of greater risk or position taking and creativity (Epstein, Landes, and Posner 2013).

Regarding the relationship of race and reversal, the literature is more silent but still provides the basis for several working hypotheses. First, the fact that black judges have been shown to vote differently than white judges (Kastellec 2013; Cox and Miles 2008) has suggestive implications: if this is the case, it would not be surprising that black lower-court judges have cases appealed at higher rates and then are also more likely to have those cases reversed—particularly if reviewed by white appeals judges without the same political or legal inclinations. Taken in tandem with implicit biases against blacks in other fields, including law and the legal system, this would suggest the first hypothesis: black judges will be more likely than white judges to be reversed on appeal. This forms the core inquiry of this study but would be a finding in contrast to some studies—for instance, that of Epstein, Landes, and Posner (2013), which finds no

relationship between the minority status of a district court judge and reversal rate.

Such a finding could have several possible explanations, however. The first potential mechanism concerns the gatekeeping issue highlighted above—the very decision to appeal. Because litigants have discretion in choosing to appeal, not all cases are appealed, and this could vary by the race of the lower-court judge in ways that complicate any findings. Here, I consider two possibilities. First, because the existing literature suggests that black judges vote differently than white judges on certain issues, and because most appellate panels are composed of judges who are all white, losing litigants in such cases may have some incentive to appeal and secure a reversal. A second possibility is grounded in the fact that practicing attorneys are known to have lower opinions of minority judges (at least at the state level; Gill, Lazos, and Waters 2011). Thus, attorneys might view those black judges' opinions with more skepticism and may be more inclined to appeal them. These two mechanisms have a clear observable implication, which is my second hypothesis: opinions written by blacks, particularly those on civil rights issues, will on average be more likely to be appealed than those written by white judges. Finding such a difference may suggest that discrepancies in reversal rates stem more from the nature of cases appealed than any bias on the part of appeals panels.

Parallel narratives would also have observable implications at the reversal stage (that is, conditional on appeal). That black judges have different, perhaps stronger or more resolute beliefs about affirmative action or voting rights (Kastellec 2013; Cox and Miles 2008) suggests that black judges would be reversed more in these issue areas, and in these areas only. Contrariwise, if black judges are reversed because they are more liberal across the board, then we would expect to see black judges being reversed more frequently across a wider swath of legal topics. If this is true, then black judges would be more likely to be reversed by more conservative three-judge appeals panels or panels with two or more Republican appointees. This would be consistent with the findings in other literature examining reversal rates—for example, Epstein, Landes, and Posner (2013), which finds that more liberal lower-court decisions are more likely to be reversed by more conservative panels. Thus, another hypothesis is that the black-judges effect should increase in issue areas involving civil rights or when appeals are heard by more conservative appeals panels. Finding any of these effects would suggest that disparate reversal rates stem from differences in voting behavior rather than other causes. I also note that such

a mechanism would suggest that more reversals, rather than a costly outcome to be avoided, may actually indicate risk taking, position taking, or creativity (as suggested by Epstein, Landes, and Posner [2013]).

There are two further explanations, both of which raise troubling normative implications under the assumption that reversal is costly. The first is that black judges could possibly bring with them different qualifications and professional experiences that result in decisions that are more likely to be overturned. In this regard, the literature is very far from agreement in terms of what constitutes judicial quality. At the same time, as Table 1 and some accounts suggest (Lott 2001; Sen 2014a), black lower-court judges are more likely than white judges to be awarded lower qualification ratings by the ABA (but see Smelcer, Steigerwalt, and Vining 2011); in addition, a lively scholarly debate has addressed the relative successes of black versus white graduates of elite law schools (Sander 2004; Ho 2005). Thus, a possibility that must be addressed seriously and with care is that systematic differences in educational opportunities or professional experiences have translated into some black judges arriving to the bench with qualitatively different experiences than white judges and that these differences translate into higher reversal rates. Whether this means that differently qualified judges write opinions reflecting these different experiences (or qualifications) or that appeals panels leverage this lack of prestige against judges with nontraditional professional backgrounds is extremely difficult to suss out using these kinds of quantitative data. Nonetheless, for the sake of ruling out alternate mechanisms, I address this explanatory hypothesis: any gap between black and white judges' reversal rates should attenuate when we compare judges with comparable professional experiences and ABA ratings. Finding such attenuation would suggest a gap driven not by bias but rather by possible differences in professional and educational preparedness.

The last possibility is that of implicit bias on the part of appeals panels. To a large extent, a case for implicit bias using observational data is circumstantial, taking the form of a stubborn difference that persists despite controls and robustness checks. Nonetheless, as noted, a growing literature suggests that implicit bias against blacks persists (and can be measured) in a wide variety of comparable instances. Here, beliefs about blacks' views could also shape biases. For example, black judges could simply be perceived to be more liberal (Scherer and Curry 2010). This in turn could contribute to a biased perception—whether substantiated or not—that black judges decide cases more liberally, which thus drives

Table 1. Distribution of American Bar Association Qualification Ratings for US District Court Judges: Johnson through Obama Administrations

	Not Qualified	Qualified	Well Qualified	Exceptionally Well Qualified	*N*
All	.01	.43	.54	.02	1,653
Whites	.01	.41	.56	.03	1,388
Blacks	.01	.57	.41	.00	147
Hispanics	.02	.61	.38	.00	104
Women	.00	.49	.51	.00	43

up their reversal rates. Another possibility is that appeals courts view black judges to be less qualified simply because of their race; that is, black judges might be perceived to produce opinions of poorer quality, despite a lack of substantive evidence. Ultimately, given how strongly implicit bias has been measured in other areas, and given (despite scholars' best efforts) continued unexplained gaps between blacks and whites in fields like education (Fryer and Levitt 2004b) and health care (Jha et al. 2005), this is a possibility that must be considered.

3. US DISTRICT JUDGE DATA

My data are from the two lower tiers of the federal judiciary—the US district courts and the US courts of appeals. District judges decide cases alone, which makes it easier to determine the impact of a particular judge's race on appeal and reversal; by contrast, appeals judges nearly always hear cases in panels of three. Also important is that appeals judges have met most of their lower-court counterparts and will therefore be aware of their basic demographic characteristics.[2]

To examine how characteristics of lower-court judges affect case outcomes, I look to data from the Federal Judicial Center (FJC), which makes public key characteristics of all federal judges.[3] For each of the 1,653 judges confirmed from the Johnson through Obama administra-

2. This assumption is borne out by the fact that higher- and lower-court judges interact personally (by virtue of frequently having offices in the same building) and professionally (by participating in judicial conferences and meetings). Excluding the jurisdiction least likely to meet this assumption—the 9th Circuit—does not meaningfully alter the results.

3. Federal Judicial Center, Biographical Directory of Federal Judges, 1789–Present (http://www.fjc.gov/history/home.nsf/page/judges.html).

tions (as of July 2012), I coded the judge's race or ethnicity, age at confirmation, gender, law school attended, and geographic location (Table 2). I used automated coding to further assess whether each nominee had previously been a former law clerk, a US attorney or assistant US attorney, a solicitor general or deputy or assistant solicitor general, a state judge (either a state supreme court or state lower-court judge), a former federal judge (for example, a magistrate, territorial, or bankruptcy judge), a full-time law professor or law school dean, an attorney in private practice, or a public defender. I also noted each judge's ABA qualification rating (historically a 4-point scale from "not qualified" to "exceptionally well qualified," with "exceptionally well qualified" dropped in 1989), which could reflect qualitative information not captured by the quantitative data (Table 1).

To measure partisanship, I recorded for each judge the identity of the appointing president and his or her JCS score, which relies on some combination of the common-space score of the appointing president or of the home-state senators (Boyd 2010; Giles, Hettinger, and Peppers 2001; Epstein et al. 2007; Poole 1998). I further coded the law school attended by using the 2001 *U.S. News and World Report* rankings and dividing them into rank cohorts: elite law schools in the top 14; schools ranked 15–25, 26–50, 51–76, or 76–100; and schools outside of the top 100. This is a rough measure for judges attending law schools in the 1970s and 1980s; however, an ameliorating factor is that the top tier's composition has never changed. A summary of these statistics is reported in Table 3.

Professional and educational characteristics are only half of the story. To assess the influence of these characteristics on reversal, I also examined case outcomes data via two distinct data sets. First, I used automated coding to collect judge-level reversal statistics reported by Westlaw in its Judicial Reversal Reports. Included in these reports are, for both published and unpublished cases, the total number of cases for which the district judge wrote an opinion, the total number of cases that were appealed for that judge, and how many of these appealed cases were affirmed or reversed (this is measured as a dichotomous variable—that is, the case was upheld or it was not). I used this information to create a data set that includes for each district judge his or her complete reversal rate from January 2000 to July 2012. This final data set includes reversal rates for 1,054 district judges, of whom 945 are white and 109 are black. The distribution of judges' reversal rates is displayed in Figure 1, where

Table 2. Racial, Ethnic, and Gender Distribution of Judicial Nominees by President: Johnson through Obama Administrations

President	Whites	Blacks	Hispanics	Women	*N*
Barack Obama	.73	.17	.11	.48	110
George W. Bush	.82	.07	.11	.21	261
William J. Clinton	.76	.18	.06	.29	305
George H. W. Bush	.89	.07	.04	.20	148
Ronald Reagan	.93	.02	.05	.08	290
Jimmy Carter	.78	.14	.07	.15	196
Gerald Ford	.91	.06	.02	.02	49
Richard M. Nixon	.96	.03	.01	.01	178
Lyndon B. Johnson	.92	.05	.03	.02	116

Table 3. Demographics of US District Court Nominees Named after 1960

			Blacks		
	All	Whites	All	Democrats	Republicans
Average age at investiture	50.06	50.44	48.55	48.16	49.49
Female	.17	.15	.27	.29	.21
Nominated by Democrat	.44	.40	.71	1.00	.00
Top-14 law school	.30	.30	.28	.31	.21
Private law school	.52	.51	.67	.69	.63
Law clerk	.21	.22	.14	.16	.09
Law professor	.06	.05	.12	.13	.07
Private practice	.92	.94	.76	.81	.65
US attorney	.09	.09	.03	.03	.05
Assistant US attorney	.20	.19	.29	.25	.40
Justice Department lawyer	.05	.05	.07	.08	.07
Public defender	.04	.03	.10	.12	.02
US magistrate judge	.09	.08	.10	.09	.12
US bankruptcy judge	.01	.01	.04	.05	.02
State judge	.41	.38	.55	.52	.63
N	1,653	1,388	147	104	43

the dotted vertical lines indicate means.[4] The figures include judges who had very few cases appealed (an issue I address in Section 5.1), which results in some judges reporting 0 or 100 percent of cases reversed.

4. One consideration is that Westlaw reports slightly higher reversal rates than are reported, for example, by the Administrative Office of the Courts. One reason might be that the Administrative Office includes all proceedings (such as motions) in its final count, while the Westlaw data include only those cases for which there was a written opinion. This could have the effect of including more important cases in the sample, which could lead to more variance in reversal rates than would otherwise be the case.

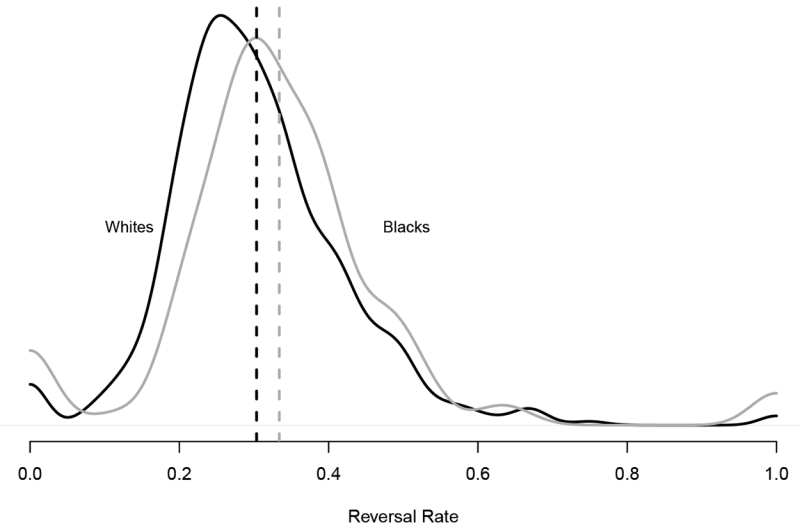

Figure 1. Reversal rates for white and black US district judges: cases appealed January 2000 to July 2012.

Because the identity of the appeals panel and other case characteristics could influence the probability of reversal (Epstein, Landes, and Posner 2013), and because the Judicial Reversal Reports by their aggregate nature do not contain this information, I examine an extant data set of appealed cases collected by Songer (2007) and by Kuersten and Haire (2011). These data include randomly selected published appeals cases decided between 1925 and 2002. In the analyses that follow, I use subsets of these data from more contemporary periods, which leaves me with 1,722 cases decided between 1996 and 2002 (or 7,279 cases decided between 1960 and 2002, for one model specification in Table 5). These data also include information on the three judges hearing the appeal, its substantive legal issue area (for instance, civil rights, criminal law, economic activity, or labor), and whether the lower-court opinion was upheld or reversed. I note that, although perfect overlap with the Westlaw data is preferable, the Kuersten and Haire (2011) data only extend through 2002. In addition, as noted by others, these data include only published cases, which could skew the sample (Epstein, Landes, and Posner 2013).[5]

5. Including unpublished cases, as I do with the Westlaw judge-level data, not only has the effect of reducing possible bias stemming from the decision to publish but also likely has the effect of presenting a more conservative overall estimate of the black-judges effect, as unpublished cases have been shown to display less variance (Keele et al. 2009). I

Here, a fact helpful to identification is that incoming cases in district courts and the appeals stage are assigned to judges (or panels) on a fairly random basis. Although the randomization can be informal, this long-standing practice makes it impermissible for federal judges to request to hear particular kinds of cases. Thus, conditional on jurisdiction, cases heard by black lower-court judges should on average be similar to those heard by white judges (that is, there should be balance in case characteristics between cases heard by black judges and those heard by white judges), and appeals panels hearing cases decided by black judges should on average be similar to appeals panels hearing cases decided by white judges (that is, potentially biased judges cannot request to hear cases decided by black lower-court judges). I present results below that suggest that the randomization is effective. I also control for issue area and other case attributes in the case-level analysis; the substantive results are unaffected.

4. RACE AS A PREDICTOR OF REVERSALS

I now turn to the key question: whether black judges are overturned more or less than their white colleagues. I perform this analysis twice, once looking at the new data on judges' overall reversal rates and again looking at case-level data from Kuersten and Haire (2011). When examining the judges' reversal rates, which include all published and unpublished cases appealed from January 2000 to July 2012, an important consideration is that the number of cases a judge hears varies by jurisdiction and length of service. For example, a judge retiring in 2001 will have fewer cases included than a judge serving the entirety of the period studied. An ordinary least squares (OLS) specification with the reversal rate as the outcome would therefore violate basic OLS assumptions: the variance of the outcome would clearly vary according to whether the judge had one case appealed or 180. For the reversal data, I therefore take a weighted least squares approach by weighting each judge by the square root of the number of cases he or she had appealed (Lewis and Linzer 2005). This is similar to what has been done in other studies of reversal (for example, Epstein, Landes, and Posner 2013).

For the analyses looking at the case-level data from Kuersten and

also note the possibility that some of the case-level data may be miscoded, as noted by Epstein, Landes, and Posner (2013).

Haire (2011), I use a logit specification, with the outcome variable being whether a case was reversed (one) or upheld (zero). I also include judge-specific random effects to account for the fact that one judge might hear multiple cases (and observations are therefore not independent). In both, to guard against the possibility that the results could be model dependent, I fit a variety of models, including dummies for appointing president, the district court where the judge sits, the circuit hearing the appeal, and several demographic characteristics. In addition, to ensure that the results are not being driven by secular changes in reversal rates, I include year and year-squared controls. Finally, I present results that rely on matching observations (Boyd, Epstein, and Martin 2010; Ho et al. 2007), which ensures that the results are not driven by a lack of common support in the data. Because these are largely consistent with the parametric results, they are presented in the Appendix.

Tables 4 and 5 present the core results. In Table 4 (judge-level reversal rates), the effect for black district judge is always positive and statistically significant, ranging in magnitude from 2 to 3 percent. Substantively, this means that black judges have a reversal rate that is between 2 and 3 percentage points higher than that of whites. This difference persists after taking into account structural characteristics that could explain discrepancies in reversal rates, such as black and white judges being appointed by different presidents (models 2–7), living in different jurisdictions (models 3–7), being of different ages on confirmation (models 5–7),[6] and having in their ranks different proportions of men and women (model 5 and, showing that an interaction between race and gender is insignificant, models 6 and 7). I also test whether the effect increases significantly when the appeals court is located in the South by including dummy variables for the Richmond-based 4th Circuit, the New Orleans–based 5th Circuit, and the Atlanta-based 11th Circuit. I find no significant difference in the effect across southern and nonsouthern jurisdictions (model 8). I also include controls for confirmation year (models 5–7), which addresses the concern that the race effect might be picking up secular trends in reversal rates.

The key results are substantiated by even stronger findings based on the case-level data, which are presented in Table 5. (Again, these are only with respect to a random subset of published cases, which could be a biased subsample of all cases.) Opinions written by black lower-court

6. Age in these analyses is included as a normalized term.

Table 4. US District Court Judges' Reversal Rates for Cases Decided 2000–2012: Weighted Ordinary Least Squares Regression Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Black district judge	.03**	.03**	.03**	.02**	.02**	.02**	.02+
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Female district judge					.001	.0002	.0001
					(.01)	(.01)	(.01)
Age					.01**	.01**	.01**
					(.003)	(.003)	(.003)
Commission year					.01	.01	−.01
					(.28)	(.28)	(.29)
Commission year^2					−.00	−.00	.00
					(.0001)	(.0001)	(.0001)
Female × black district judge						.002	.01
						(.02)	(.02)
South							.003
							(.01)
South × black district judge							.02
							(.02)
Constant	.31**	.32**	.31**	.48**	−8.75	−9.99	13.60
	(.003)	(.02)	(.02)	(.02)	(283.40)	(283.71)	(284.43)
President dummies	No	Yes	Yes	Yes	Yes	Yes	Yes
Circuit dummies	No	No	Yes	No	No	No	No
District dummies	No	No	No	Yes	Yes	Yes	Yes
Adjusted R^2	.01	.01	.31	.48	.49	.49	.49

Note. Judge-level reversal rates are continuous between 0 and 1. Standard errors are in parentheses. N = 1,054.

+ $p < .10$.

** $p < .01$.

judges are more likely to be reversed on appeal, and the effect is robust to the inclusion of district-level dummy variables and dummies for the appointing president.[7] In terms of predicted probabilities, Table 5 suggests even stronger effects: an opinion written by a black judge may have on average an approximately 10-percentage-point greater chance of being reversed than an opinion written by a white judge (depending on the model). In addition, I investigate whether the effect is an artifact of the time period in question by examining the spread of data available from Songer (2007) and Kuersten and Haire (2011), dating back to 1961, when the first black judge was nominated (model 5). The effect appears regardless of whether I examine cases only from 1996 forward or going back to 1960. In addition, an interaction between the race of the lower-court judge and the time period (for example, pre- or post-1996) is not significant (not shown), which suggests that I cannot rule out that the effect has either strengthened or attenuated.

These are substantively meaningful differences: considering that the average black district judge has approximately 196 cases appealed in a 12-year period, this 10-percentage-point gap results in a difference between black and white judges of approximately 2,800 cases. That is, if blacks were reversed at whites' comparably lower rates, approximately 2,800 black-authored decisions would have been upheld over the period from 2000 to 2012 instead of having been reversed. (Contrariwise, if whites were reversed at blacks' higher rates, approximately 7,500 white-authored decisions would have been reversed on appeal instead of having been upheld.) At the individual level, these results mean that each black judge on average had approximately 20 more cases reversed (out of an average of 196 cases appealed) than if he or she had been white.

Three points are worth further mention. First, I note that the substantive interpretations do not change depending on whether I include district court dummies (Table 4, models 4–7) or circuit dummies (model 3). However, the explanatory power from the model increases swiftly when I include any kind of control for district. (The overall R^2-value in Table 4 moves from around .01 to around .5; models 4–6.) I include dummy variables for district in most of the remaining analyses. This has the additional benefits of controlling for fluctuations in case dockets among districts (for instance, the Southern District of California versus the District

7. The results are also robust to the inclusion of age and gender and, in results not shown here, to the inclusion of the number of years on the bench (amount of experience) the judge had before hearing the case in question.

Table 5. Appeals Court Case-Level Reversal Rates: Logit Regression Results

	Model 1	Model 2	Model 3	Model 4	Model 5
Black district judge	.49**	.40*	.49*	.51*	.24*
	(.18)	(.19)	(.21)	(.21)	(.11)
Female district judge				.22	.17+
				(.18)	(.11)
Age				−.01	.05+
				(.07)	(.03)
Case year				−43.15	3.28*
				(67.93)	(1.42)
Case year^2				.01	−.001*
				(.02)	(.0004)
Constant	−1.15**	−1.06**	−.82**	43,099.92	−3,243.15*
	(.06)	(.13)	(.28)	(67,892.18)	(1,414.47)
President dummies	No	Yes	Yes	Yes	Yes
District court dummies	No	No	Yes	Yes	Yes
N	1,718	1,718	1,718	1,718	7,279
Log likelihood	−962.47	−958.94	−899.77	−898.28	−4,233.36

Note. The outcome variable is whether a case was reversed (one) or upheld (zero). Data are a randomly selected subset of published cases (Songer 2007; Kuersten and Haire 2011) from 1996–2002 (models 1–4) and 1960–2002 (model 5). All models include judge-specific random effects. Standard errors are in parentheses.

+ $p < .10$.

* $p < .05$.

** $p < .01$.

of Alaska) and of making safer the assumption that cases are randomly assigned (as cases are usually assigned randomly, but only within district). Second, the results are significant regardless of whether I look at overall reversal rates (Table 4) or randomly selected case-level data (on published cases, Table 5). In the analyses that follow, I primarily use the overall reversal rates, using the case-level data to analyze instances when the composition of the appeals panel is thought to play a salient role. I do so both because the overall reversal rate data are more conservative and, more important, because they include the universe of appealed cases, not just a random subset of published cases. Last, these core results are consistent with the additional matching analyses presented in the Appendix.

These results provide support for the first hypothesis, that black judges will be more likely to be overturned on appeal. I turn now to exploring possible reasons behind this difference, beginning with the theory that appeals courts hear cases that vary in number and type according to the race of the lower-court judge.

5. POTENTIAL EXPLANATIONS

5.1. Mechanism 1: Differences in Cases Appealed

The first explanation I consider is that cases authored by black lower-court judges will perhaps be more likely to be appealed as a result of differences between black and white voting or some kind of bias against black judges on the part of legal practitioners (as suggested by the state-courts literature; see, for example, Gill, Lazos, and Waters 2011). Although I control for the number of cases appealed (via weighting in Table 4), it is possible that a persistent difference in both the number and the type of cases appealed could be skewing the results. That is, black judges' having more cases appealed may mean that more of them will be reversed; likewise, if more civil rights cases are appealed from black judges, and if decisions in civil rights cases are more likely to be reversed, this would create the impression of race-based reversal when none in fact exists.

I first examine the number of cases appealed and whether they vary by the race of the lower-court judge. For these analyses, the outcome variable is the number of cases appealed from each district judge from 2000 to 2012.[8] To take into account the fact that some judges produce more

8. I note that these figures do not take into account the number or type of cases in which the judge presided over a settlement. As some scholarship shows (for example,

or fewer authored opinions than average, I also include the number of opinions each judge produced during this time period as a control variable (Table 6, models 2–5). Further, to take into account the variation in norms for appeal across jurisdictions, I also include specifications that include dummy variables for jurisdiction in models 3–5 (in the form of district court, although dummy variables for appeals court result in identical inferences).

Surprisingly, Table 6 shows that I cannot rule out that there is no difference between black and white judges in terms of rates of appeal: under all model specifications, the substantive difference between blacks and whites is negligible and never statistically significant. There are two further items of note. First, as would be expected, the total number of cases each judge produces is itself predictive of the number of cases appealed—that is, judges who write more opinions have more of them appealed. Second, what explanatory power there is in the model comes not from the race of the lower-court judge but from the addition of dummies for the district court (models 3, 4, and 5), the inclusion of which causes the R^2-value to increase substantially. For the purposes of this inquiry, the analysis is unable to rule out that there is no relationship between the race of the lower-court judge and rates of appeal.

However, the other concern is that the number of cases might not vary but perhaps the kind of case does. Here the concern is that black judges' rulings on racially salient issues may be more (or less) likely to be appealed. To address this issue, I supplement Table 6 by using data from Kuersten and Haire (2011) to analyze whether case dockets involving published appeals vary within circuit according to the race of the lower-court judge. I use a Fisher's exact test to test for a relationship between district judge race and the issue areas coded by the Kuersten and Haire (2011) data, conditional on the 12 federal appeals circuits. The Fisher's exact test operates by comparing the observed contingency table (here, for each circuit) to all possible contingency tables with the same marginal counts. It is useful in this case because it does not rely on large sample approximations, and many circuits had few cases in specific issue areas.

The categories I test are the eight issue areas coded by Kuersten and Haire (2011): criminal, civil rights, First Amendment, due process, privacy, labor relations, economic activity, and miscellaneous (and also a "not ascertained" category). The results, by jurisdiction, are presented in

Boyd 2013), this could vary according to the identity of the lower-court judge. This does have the potential to bias the results, although I do not see evidence of it here.

Table 6. Number of Cases Appealed: Ordinary Least Squares Regression Results

	Model 1	Model 2	Model 3	Model 4	Model 5
Black district judge	8.25	9.58	9.05	10.08	−5.29
	(16.49)	(15.26)	(12.98)	(11.36)	(10.96)
Total number heard		.12**	.14**	.10**	.09**
		(.01)	(.01)	(.01)	(.01)
Female district judge					−11.77
					(8.97)
Age					−24.43**
					(3.67)
Commission year					2,503.98**
					(368.88)
Commission year^2					−.63**
					(.09)
Constant	188.05**	125.09**	122.09**	20.51	−2,491,447.00**
	(5.30)	(6.81)	(23.95)	(23.95)	(367,739.40)
District court dummies	No	No	Yes	Yes	Yes
President dummies	No	No	No	Yes	Yes
R^2	.0002	.14	.47	.62	.65
Adjusted R^2	−.001	.14	.42	.58	.61

Note. The outcome variable is the total number of cases appealed from each judge between 2000 and 2012. Standard errors are in parentheses. N = 1,054.

** $p < .01$.

Table 7. Difference between Black and White Judges' Cases across Legal Issue Areas: Fisher's Exact Tests

Circuit	*p*-Value
DC	.18
1st	.21
2nd	.10
3rd	.99
4th	.19
5th	.15
6th	.59
7th	.43
8th	.68
9th	.13
10th	.71
11th	.12

Note. Tests use cases from Kuersten and Haire (2011). No results are significant.

Table 7. Across all of the jurisdictions, I cannot reject the null hypothesis that there is no relationship between lower-court judge race and the type of case heard on appeal. Taken together, this leads me to reject the second explanation. There is no support for concluding that attorneys either appeal opinions written by black judges at higher rates or choose to appeal different kinds of cases depending on the race of the lower-court judge. Thus, to the extent that black judges are more likely to be reversed on appeal, it is unlikely to be because of the number and nature of the cases being appealed.

5.2. Mechanism 2: Possible Differences in Quality

The FJC data show that black and white judges differ on average in terms of some types of previous employment, qualification scores, and law schools attended (Tables 1 and 3). For example, fewer black judges have private practice experience compared to white judges (76 percent versus 94 percent), while more have state judge experience (55 percent versus 38 percent). Black judges are less likely to receive high qualification ratings from the ABA; 41 percent of them receive a high "well qualified" rating compared to 56 percent of white judges. (Discrepancies in qualification ratings are explored in Lott [2001] and Sen [2014a]; Smelcer, Steigerwalt, and Vining [2011] provide evidence to the contrary.) However,

there is disagreement in the literature about whether objective criteria can ever determine what makes a good judge, whether such criteria are useful predictors of reversal, and whether some factors (for example, public defender experience) could cut both ways. Nonetheless, addressing these issues is essential to possibly understanding why reversal rates for black and white judges differ, and I do so by including both objective measures of experience and education and subjective measures such as ABA ratings.

As in Table 4, the outcome variable for this analysis is the judge's reversal rate from 2000 to 2012 across all cases (published and unpublished). I again include dummy variables for the identity of the appointing president and the district court. The results are displayed in Table 8. In model 1, I include the ABA ratings as dummy variables. In model 2, I include various educational variables, including rank cohort of law school attended. (The excluded category is attending a top-14 law school, considered the elite group; a large number of judges attended one of these schools, particularly Harvard [121] and Yale [54] law schools.) Model 3 includes professional experience, such as whether the judge had been in private practice or worked as a law clerk. Model 4 includes whether the judge had served in a judicial capacity before, perhaps as a federal magistrate or state judge (in either a state supreme court or state lower court), and model 5 includes additional controls for commission year.

Despite the inclusion of both the subjective and objective criteria, the inferences do not change: black district judges are still more likely to be overturned on appeal than are white judges—with an increase in reversal rate of approximately 2 percentage points. Perhaps surprising is the fact that relatively few of the educational measures and professional experiences predict reversal: there are few differences between those who went to high-ranked law schools and those who did not, those who were US attorneys and those who were not, and so forth. The only exception, perhaps, are former state judges (who are more likely to be reversed) and former law clerks (slightly less likely). The models are by observational standards fairly predictive, with R^2-values close to .50. However (as before), most of the explanatory power comes not from the professional or educational variables but from the dummy variables for appointing president and for jurisdiction (model 1).

Although the gap between black and white judges does not attenuate with the inclusion of these professional and educational characteristics, I must consider whether other unmeasurable or qualitative traits are driving the results. Some literature suggests that even black graduates of

high-performing law schools do not perform at the level of their white peers (Sander 2004; but see Ho [2005] for a rebuttal). For this analysis, some traits that are not captured in the FJC data are LSAT score, law review membership, law school rank, membership in the Order of the Coif, bar passage, or writing ability. I do not address this debate on black achievement specifically, but there are several reasons why it is less of a concern with this analysis. The first is that the results obtained via matching are not only consistent but actually fairly robust to potential omitted-variable bias (see the tests presented in the Appendix). The second is that few of the prestige-oriented variables that are included in the federal data are predictive of reversal rates: being a US attorney, graduating from a top law school, or even having served as a full-time law professor or dean does little to predict judges' reversal rates. Being a law clerk is only fleetingly predictive, and its effect is quite minuscule compared to the variance explained by simply conditioning on jurisdiction (or even compared to the effect of a lower-court judge being black). An appropriate supposition is that including similar variables that indicate prestige variables (such as law review membership) would result in similar nonsignificance.

Third, although I do not have access to data like class rank or bar passage (which are not made public by the FJC), I do condition on attributes clearly predicated on those marks of success: few advance to legal clerkships, law professorships, or US attorney positions without having achieved some combination of high class rank (grade point average), law review membership, and bar passage. Presumably, controlling for these professional experiences also controls to some extent for the unrecorded traits. Last, I also condition on ABA ratings—which purport specifically to assess a judicial nominee's professional competence, a qualitative assessment based on quality of legal reasoning, class rank, law review membership, and bar passage. Thus, these are characteristics that would likely be reflected in a candidate's ABA rating, and the results are robust to the inclusion of this variable. Taken together, these considerations lead me to reject this possible mechanism as an exclusive explanation for the findings.

5.3. Mechanism 3: More Liberal Voting

The results suggest that the racial gap persists despite controlling for differences in background—including both objective experiences and subjective qualification ratings. Perhaps a more likely possibility is that black judges are more liberal across the board than are comparable white

Table 8. Qualifications and US District Court Judges' Reversal Rates for Cases Decided 2000–2012: Weighted Ordinary Least Squares Regression Results

	Model 1	Model 2	Model 3	Model 4	Model 5
Black district judge	.02**	.02**	.02**	.02*	.02*
	(.01)	(.01)	(.01)	(.01)	(.01)
Female district judge	.001	.002	.003	.002	.001
	(.01)	(.01)	(.01)	(.01)	(.01)
Age	.01**	.01**	.01**	.01**	.01**
	(.003)	(.003)	(.003)	(.003)	(.003)
Law school ranked 15–25		.01	.005	.005	.003
		(.01)	(.01)	(.01)	(.01)
Law school ranked 26–50		.001	.0000	−.0005	.0003
		(.01)	(.01)	(.01)	(.01)
Law school ranked 51–75		.02	.01	.01	.01
		(.01)	(.01)	(.01)	(.01)
Law school ranked 76–100		−.01	−.02+	−.02+	−.02+
		(.01)	(.01)	(.01)	(.01)
Law school ranked 101+		.0003	−.0000	−.001	−.001
		(.01)	(.01)	(.01)	(.01)
Law clerk		−.01*	−.01+	−.01+	−.01
		(.01)	(.01)	(.01)	(.01)
Law professor			−.003	−.002	−.004
			(.01)	(.01)	(.01)
Private practice			.01	.01	.01
			(.01)	(.01)	(.01)
US attorney			−.01	−.01	−.01
			(.01)	(.01)	(.01)

Assistant US attorney			.0001	.001	.001
			(.01)	(.01)	(.01)
Justice Department lawyer			−.02	−.02	−.02
			(.01)	(.01)	(.01)
Public defender			.02+	.02+	.02+
			(.01)	(.01)	(.01)
Federal magistrate				.01	.01
				(.01)	(.01)
Federal bankruptcy judge				.01	.01
				(.02)	(.02)
State judge				.01*	.01*
				(.01)	(.01)
Commission year					.02
					(.28)
Commission year^2					−.0000
					(.0001)
Constant	.45**	.46**	.46**	.45**	−16.35
	(.03)	(.03)	(.03)	(.03)	(284.03)
R^2	.54	.54	.55	.55	.55
Adjusted R^2	.48	.49	.49	.49	.49

Note. Judge-level reversal rates are continuous between 0 and 1. All models include American Bar Association rating, president, and district court dummies. Standard errors are in parentheses. N = 1,054.

+ $p < .10$.

* $p < .05$.

** $p < .01$.

judges, even among those appointed by the same president. For example, some literature suggests that presidents who appoint minorities take the opportunity to appoint more ideologically driven individuals than they would otherwise (Asmussen 2011). For blacks, this would bring to the bench more left-leaning (or right-leaning, in the case of Republican presidents) black candidates, who would then be overruled more by moderate appeals panels across all kinds of legal issue areas. I note some evidence of this in Figure 2, which shows the JCS scores of black and white judges by party (Boyd 2010). Black judges have more left-leaning JCS scores than white judges, which in turn raises the possibility that they write opinions that are more liberal and hence reversed at higher rates by more centrist appeals panels. Again, this could be the case despite appointment by the same president.

Testing District Judge Ideology Directly. To more directly analyze the role of lower-court-judge ideology (or at least as directly as possible, given current ideological measures), Table 9 explores reversal rates with four model specifications: district judge party and race interacted, district judge JCS score and race interacted,[9] appointment by a Democrat, and appointment by a Republican.

To be clear, the possibility exists that minority judges are still more liberal than white judges appointed by the same president. However, the analysis does not rule out that there are no differences between Democratic and Republican appointments in terms of reversals among black judges (model 1); an interaction of race of the district judge and the party of appointing president is not significant under any model specifications. In addition, the interaction of JCS score and race of the lower-court judge is also not significant (model 2).[10] However, Table 9 does suggest that the effect appears to be driven by Democrat-appointed judges. Among Democrat-appointed judges, the reversal rate increases by approximately 3 percentage points for black judges compared to whites, and this difference is statistically significant; the effect is robust to the inclusion of the variables associated with prior experience and also to the inclusion

9. Note that there are many judges for whom current JCS scores are not available; thus, the sample size decreases markedly between model 1 and model 2.

10. This is not surprising, as JCS scores capture the ideology of either the appointing president or the senior home-state senator (or some combination of the two home-state senators), depending on the partisan alignment (Boyd 2010). The lack of precision in estimating lower-court ideology is an issue not just for the present study but for others trying to control for lower-court judges' beliefs.

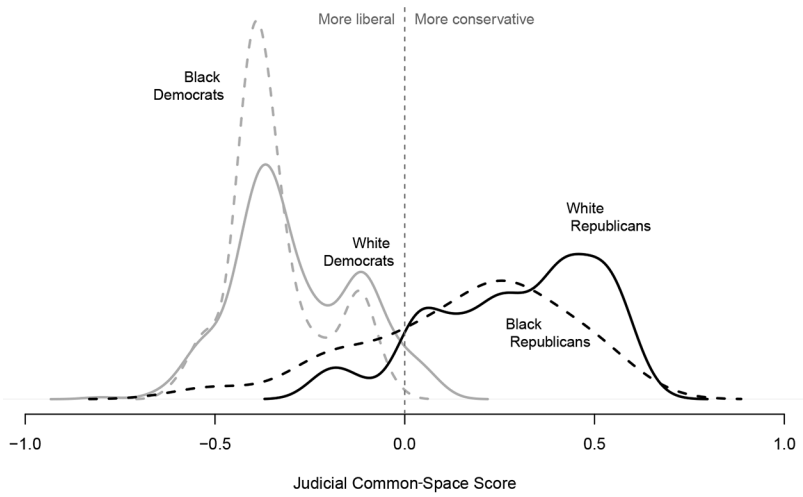

Figure 2. Judicial common-space scores of black and white district judges by party

of dummy variables for appointing president and district court. Among Republican judges, however, the effect is close to 1 percentage point and never significant. Nor is the gap driven by systematic differences between black Democrats and Republicans—for example, in terms of their previous professional experience: the models all include controls for professional experience and ABA ratings, and, as Table 3 demonstrates, black Democrats and black Republicans do not differ in ways that would suggest that black Republicans are somehow stronger or more experienced judicial candidates. (If anything, a higher share of black Democrats attended elite top-14 law schools and served as law clerks.)

Testing the Relationship to Appeals Panel Ideology. These results are consistent with the theory that black judges are simply more liberal than other judges—that is, compared to white Democrats, black judges are more liberal and therefore more likely to be reversed. (The results are not consistent with a parallel story, one in which black Democrats are more liberal and black Republicans more conservative—that is, minority candidates are more "extreme." If this story were true, then the effect would be seen regardless of party, and not for Democrats only.) Unfortunately, existing measures of lower-court ideology do not allow me to rule out this possibility conclusively.

I therefore turn to some indirect tests. I examine the appeals case-level data from Kuersten and Haire (2011) analyzed in Table 5. If the effect is

Table 9. US District Court Judges' Reversal Rates for Cases Appealed 2000–2012: Weighted Ordinary Least Squares Regression Results

	Model 1	Model 2	Democrats Only	Republicans Only
Black district judge	.02*	.02+	.03*	.01
	(.01)	(.01)	(.01)	(.01)
Republican district judge	−.07+			
	(.04)			
JCS		−.01		
		(.02)		
Republican × black district judge	−.01			
	(.02)			
JCS × black district judge		−.01		
		(.03)		
Commission year	.02	.38	−.14	.29
	(.29)	(.80)	(.93)	(.29)
Commission year^2	−.0000	−.0001	.0000	−.0001
	(.0001)	(.0002)	(.0002)	(.0001)
Constant	−19.36	−378.25	140.33	−282.28
	(284.31)	(791.23)	(923.12)	(293.49)
N	1,054	648	457	597
Adjusted R^2	.49	.50	.40	.57

Note. Judge-level reversal rates are continuous between 0 and 1. All models include American Bar Association rating, president, and district court dummies and controls for education and experience. Standard errors are in parentheses. JCS = judicial common-space score.

+ $p < .10$.

* $p < .05$.

driven by black lower-court judges' being more liberal than whites, then I would expect not only that black Democrats would be more likely than white Democrats to be reversed (as in Table 8) but also that the difference between black and white judges in terms of reversal rates would increase when black judges' opinions are reviewed by conservative three-judge appeals panels. That is, the gap between black and white judges should increase with more conservative reviewing courts, even conditional on a Democratic lower-court judge.

I check this by examining the interaction between the race of the lower-court judge and the composition of the appeals panel. I conduct three analyses, interacting the race of the lower-court judge (black or white) with the mean JCS score of the appeals panel, whether the appeals panel had two or more Republican appointments (that is, whether it had a Republican majority), and whether the appeals panel's decision was coded by Songer et al. as having ruled in a conservative direction. I include both all lower-court judges (models 1–3) and, for ease of interpretation, Democrats only (models 4–6). All of the models include dummies for the district court and for the appointing president in addition to controls for case year. If black judges' (or black Democrats') more liberal voting is driving their increased reversal, then I would expect a positive relationship in the interaction terms. That is, if black judges are voting in a more liberal direction than their white counterparts, then Republican appeals panels (or panels ruling in a more conservative direction) would be more likely to reverse them.

However, as Table 10 demonstrates, the interactions of the black-judges variable and the median ideology, partisanship, and directionality of appeals panel ruling are never significant, even when I restrict the sample to those for whom the effects are the strongest (black and white Democrats, models 4–6); that is, I cannot rule out that the black-judges effect does not vary according to the ideology or partisanship of the appeals panel (models 2 and 5), or even according to how the appeals panel rules on cases (models 3 and 6). If anything, the negative coefficients on some of the interaction terms (for example, in models 1, 2, 4, and 5) suggest that the effect decreases when the three-judge panel becomes more conservative. The explanation behind this is unclear, and none of the interactions are significant. However, it does provide additional evidence against the theory that black judges' increased reversal rate is driven by their being more liberal across the board: I cannot rule out that the increased

Table 10. Case-Level Reversal and Composition of the Appeals Panel, Case-by-Case Basis: Logit Regression Results

	All Judges			Democrats Only		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Black district judge	1.23**	1.24**	.27	1.62**	1.53*	.39
	(.36)	(.47)	(.34)	(.49)	(.63)	(.42)
Mean panel ideology (conservative)	−.14			.35		
	(.40)			(.71)		
Republican panel		−.11			−.17	
		(.22)			(.41)	
Conservative decision			−1.58**			−1.40**
			(.14)			(.25)
Black district judge × mean ideology	−.44			−1.06		
	(1.00)			(1.25)		
Black district judge × Republican panel		−.11			−.17	
		(.56)			(.74)	
Black district judge × conservative decision			.64			.76
			(.43)			(.54)
Case year	72.14	−11.98	−16.00	−106.84	−86.33	−49.80
	(118.92)	(101.13)	(72.89)	(195.26)	(168.87)	(122.79)
Case year^2	−.02	.003	.004	.03	.02	.01
	(.03)	(.03)	(.02)	(.05)	(.04)	(.03)
Constant	−72,097.20	11,938.10	15,992.03	106,853.50	86,370.12	49,848.34
	(118,853.30)	(101,069.60)	(72,848.71)	(195,157.30)	(168,782.70)	(122,726.10)
N	669	934	1,639	272	368	627
Log likelihood	−304.26	−425.33	−786.84	−124.58	−163.33	−302.24

Note. The outcome variable is whether a case was reversed (one) or upheld (zero). The data are a randomly selected subset of published cases (Songer 2007; Kuersten and Haire 2011) from 1996–2002. All models include president and district court dummies and judge-specific random effects. Standard errors are in parentheses.

* $p < .05$.

** $p < .01$.

reversal of black judges does not vary according to the composition or ruling of the appeals panel.

5.4. Mechanism 4: Voting on Certain Issues

A remaining possibility is that black judges vote differently than white judges but that they do so only with regard to cases having a significant racial, ethnic, or civil rights dimension. This could include substantive issue areas involving affirmative action and civil rights (Kastellec 2013; Cox and Miles 2008) or, possibly, criminal cases involving race-based defenses by black defendants. We might expect black judges to be overruled most frequently in these areas, where their views might differ the most from whites. Alternatively, we might expect black judges to be upheld more in these areas, with appeals judges being deferential to black judges on racially sensitive cases. Either scenario predicts that the black-judges effect would vary significantly between possibly racially salient areas (such as civil rights law) and others.

To test this possibility, I include in the case-level analyses a dummy variable for civil-rights-related cases, as coded by Kuersten and Haire (2011), interacting it with the race of the lower-court judge (Table 11). (In models not shown, I also control for all eight of Kuersten and Haire's issue areas; the inferences are not affected.) For the issue-area explanation to hold sway, the racial gap should differ across the areas identified by the judicial politics literature as being particularly racially salient; that is, I should see a significant relationship in the interaction of lower-court black authorship and civil rights issues. The results show, however, that the effect of black authorship on a case's probability of being upheld varies little by civil rights issue area: there is no difference in blacks' reversal rates across civil rights cases and non-civil-rights cases (model 2). Thus, the data provide no evidence for the proposition that black judges are being overturned at greater or lesser rates (compared to white judges) within certain legal categorizations.

5.5. Mechanism 5: Race as Signal

An explanation that must be considered is that appeals panels somehow implicitly rely on the race of the lower-court judge in reaching decisions. On the one hand, this explanation has the deepest and most troubling normative implications and challenges the fairness and race neutrality of the judiciary. On the other, such a finding would perhaps be unsurpris-

Table 11. Case-Level Reversal by Issue Area, Case-by-Case Basis: Logit Regression Results

	Model 1	Model 2	Model 3	Model 4	Model 5
Black district judge	.50*	.51*	1.18**	1.15**	1.29**
	(.21)	(.23)	(.29)	(.33)	(.30)
Civil rights case	−.27+	−.27			
	(.16)	(.18)			
Black district judge × civil rights case		−.06			
		(.47)			
One black on panel			−.39	−.53+	
			(.28)	(.31)	
Two blacks on panel			−.09	.87	
			(.72)	(.78)	
Majority black panel					1.03
					(.77)
Black district judge × one black on panel				.80	
				(.70)	

Black district judge × two blacks on panel				−20.11	
				(4,554.49)	
Black district judge × majority black panel					−20.25
					(4,575.08)
Case year	−40.23	−39.77	−5.64	−3.30	
	(67.93)	(67.97)	(101.12)	(101.77)	
Case year^2	.01	.01	.001	.001	
	(.02)	(.02)	(.03)	(.03)	
Constant	40,182.47	39,720.22	5,595.49	3,250.84	−.93*
	(67,899.46)	(67,935.23)	(101,065.00)	(101,711.60)	(.38)
N	1,717	1,717	934	934	934
Log likelihood	−897.47	−897.46	−424.52	−420.53	−422.56

Note. The outcome variable is whether case was reversed (one) or upheld (zero). The data are a randomly selected subset of published cases (Songer 2007; Kuersten and Haire 2011) fom 1996–2002. All models include president and district court dummies and judge-specific random effects. Standard errors are in parentheses.

$^{+}$ $p < .10$.

* $p < .05$.

** $p < .01$.

ing, as studies have teased out implicit biases against racial minorities in prominent economic, social, and political settings.

Here I consider one possible manifestation by testing whether having a black presence on an appeals panel attenuates the effect. This could occur because white judges become more sensitive to any possible discriminatory tendencies or because black judges raise possible concerns about bias. (On this point, Kastellec [2013] provides evidence that having a black judge on an appeals panel changes the way in which the panel votes on affirmative action issues.) To test this theory, I evaluated how the racial gap varies across different racial compositions of appeals panels: zero, one, or two black judges on the three-judge higher-court panels.[11] As before, I use a mixed-effects logit model, with an interaction of the race of the lower-court judge and the number of black circuit judges hearing the appeal (model 4). I also include an interaction of the race of the lower-court judge and whether blacks constituted a majority of the panel (model 6). No cases in the data were heard by an all-black three-judge panel.

Results from these analyses are presented in Table 11, models 4–6. Because of the low numbers of black judges on appeals courts, and because of the fact that these judges very rarely sit together, it is impossible to distinguish how the effect varies across panels involving zero, one, or two black judges (models 4 and 5). Although the black-judges effect appears to attenuate when black judges constitute a majority on the appeals panel, the interaction is not at all significant. I also note that this analysis does not rule out the possibility that black appeals judges are more liberal than white judges, an implication that would also explain these patterns on more ideological grounds.

6. CONCLUSION

The results show that discrepancies exist in how appeals courts review cases, with black judges being up to 10 percentage points more likely to be reversed than whites. This racial gap is robust and persists once I control for possible proxies for judicial qualifications—for example, quality of legal education, age, professional experience, and ABA ratings. Con-

11. Note that there are very few cases involving two or more black judges on the same appeals panel. For example, in the Songer et al. data for cases decided between 1960 and 2002, 88 percent had no black judges on the three-judge appeals panel, 11 percent had one black judge, and just under 1 percent had two.

trols for the partisanship of the lower-court judge and for the partisanship and racial composition of the reviewing appeals panel do not affect the results. The discrepancy in reversal rates between black and white judges does not appear to vary across issue area or across jurisdictions. And this gap translates into meaningful legal outcomes. Close to 3,000 federal court decisions would have been upheld if black judges were overturned at whites' lower rates. At the individual level, black judges on average have up to 20 more cases reversed than do similar white judges, out of an average of 196 cases appealed.

The reasons behind this persistent difference are not straightforward. Although having blacks on the reviewing panel appears to attenuate the effect, there are too few black appeals court judges to make meaningful inferences. The difference appears not to be driven by black judges voting differently on certain cases. At best, there is suggestive evidence that black judges are more liberal than otherwise similar whites, but the fact that I cannot rule out that black judges' increased reversal does not vary across conservative and liberal appeals panels casts doubt on this being the sole explanation. A more likely explanation is that this discrepancy is driven by a variety of factors—perhaps it is due to black judges being more liberal in ways unmeasured by extant ideology measures (accompanied or perhaps buttressed by the perception of black judges being more liberal; Scherer and Curry 2010), but it may also be due to implicit biases on the part of higher courts (perhaps based on the perception that black judges are less qualified). On this point, better measures of lower-court judicial ideology—including possible text-based measures—would go far in adjudicating between an ideologically based explanation and other possible mechanisms. This presents a clear path for future research.

A point worth emphasizing is that the gap between black and white judges attenuates at times but never fully disappears. The implications for this particular gap are striking, regardless of the reason. Since Kennedy, American presidents have actively sought to appoint judges of color—not just blacks, but also Hispanics and Asian Americans—to the nation's highest courts. At the state and international levels, too, efforts are underway to increase the proportion of judges from underrepresented communities. The racial gap demonstrated here, however, calls into question whether the mere appointment of these individuals is enough. After all, if certain judges are being systematically overturned more often, then this raises questions about their long-term impact on the law, legal precedent, and the legal system.

The results presented in this paper actually represent the tip of the iceberg in exploring the components of judicial evaluation and its relationship to descriptive representation, a topic previously unexplored in the judicial politics literature. I have touched on just one ascriptive characteristic: the race (black or not) of lower-court judges. Similar effects may exist for other racial or ethnic groups (Asian Americans, Hispanics), religious groups (Jews, Catholics), and genders—not to mention multiple combinations of these identities. In addition, if heuristics or personal familiarity play a role in how appeals panels reach decisions, then we may find different rates of overturning between judges who attended the same law school or are otherwise familiar or friendly; that is, a personal connection may strengthen a bond that makes reversal less likely. Further research should help clarify the extent to which these and other attributes might play a role in appellate review.

In addition, this is a study that relies on a quantitative analysis of aggregated data. Still remaining for future work is a closer, qualitative look at the opinions authored by both black and white lower-court and appeals judges. Do black judges use different legal reasoning or articulate legal principles in a different way? Do black judges rely on particular arguments in defining their opinions? Does the language used by appeals panels differ according to the identity of the legal actors involved? The analyses presented here suggest that there could be something qualitatively different about those opinions written by black judges, as well as some qualitative differences in how appeals panels review cases decided by black judges. Given the results of this analysis, a qualitative examination of these issues would further shed light on why black judges are more likely to be overruled and why this racial gap is so persistent.

APPENDIX: MATCHING ANALYSES

Because black and white judges differ in age, previous employment, partisanship, and geographic dispersion, and because different cases arise in different jurisdictions, simple comparisons may mask a lack of overlap in the data. To account for differences, I present additional results using matching (Boyd, Epstein, and Martin 2010; Ho et al. 2007). Matching operates by comparing reversal rates among judges who are identical across key characteristics. Thus, a black judge sitting in the 8th Circuit who graduated from a second-tier law school with previous experience working in private practice will be matched to a white judge also sitting in the 8th Circuit with a similar profile.

This approach offers advantages and disadvantages. First, matching is an effective preprocessing step that reduces dependence on modeling assumptions (Ho et al. 2007). Second, and relatedly, matching effectively tests all possible ways that variables could interact with each other. A drawback is, however, that observations are dropped, which results in inferences that are based on only a subset of the original population. For the core results presented, this does not appear to be a problem: sufficient observations remain after matching to make statistically significant inferences, and the matched sample is by no means an anomalous subset of the entire universe of judges. The results obtained by matching are consistent with the results obtained via parametric methods, presented in the main text.

A1. Matching Methodology

To implement the matching, I use coarsened exact matching (Iacus, King, and Porro 2009, 2011), which allows exact matching on key variables and coarsening and then approximate matching on the three variables that are continuous (discussed below). Coarsened exact matching has the advantage of allowing for this approximation to be as close as needed to remove biases. I also have the advantage of matching exactly on a large portion of the variables measuring judicial qualifications.[12]

Once the judges were matched, I took the difference in means in reversal rates, obtained via simple linear regression. In the results presented, I match on the same key variables analyzed above. These include whether the judge is male or female, a Republican or Democratic appointee, a former federal magistrate or bankruptcy judge, a former attorney in practice practice, and a graduate of a top-14 law school, as well as his or her corresponding appeals court.[13] I further coarsen an additional set of variables using specific cut points. These are the judge's birth year, number of years either on a federal bench (for example, as a magistrate judge) or in private practice, and JCS score (Boyd 2010).

A summary of some of the judge characteristics after matching on the judge-level data is given in Table A1. This matched sample of judges is, as expected, slightly different than the original prematched sample (Table 3) but by no means anomalous. Overall, more judges in the matched sample were appointed by Democrats (specifically, by President Bill Clinton) and had experience in private practice. Slightly fewer of them had experience as a federal judge (for example, as a magistrate or bankruptcy judge), and, on average, they had slightly less trial

12. Using different matching estimators (nearest-neighbor matching and propensity score matching) yielded similar substantive results, as did estimating the effect without discarding any treated units (that is, black judges). I present the results from coarsened exact matching, as it bounds the maximal amount of imbalance through the choice of coarsenings.

13. Matching on district court, which would be ideal, is not possible because of the sample size.

Table A1. Demographics of Matched District Court Judges Compared to Entire Population of District Court Judges

	All Whites	All Blacks	Matched Whites	Matched Blacks
Female	.15	.27	.28	.28
Appointed by Democrat	.40	.71	.68	.68
Top-14 law school	.30	.28	.23	.26
Former federal judge	.09	.14	.07	.07
Private practice	.94	.76	.90	.90
Average commission year	1987.49	1991.95	1995.83	1995.75
Trial years	16.75	10.08	13.68	12.28
N	1,388	147	172	72

experience before nomination. There are also more female judges in the matched sample.

A2. Matching Results

I run matching analyses twice. The first analysis is on the judge-level data presented in Table A2, while the second is on the Songer et al. case-level data from 1960 to 2002. The judge-level postmatching results are estimated using a weighted OLS specification (with the weights coming from the number of cases appealed), while the case-level postmatching results are calculated using a mixed-effect logit model (because of the binary nature of the outcome variable) with judge-specific random effects. The results are, however, substantially similar. After matching, reversal rates of black judges are approximately 3–4 percentage points higher than those of comparable white judges. This is the case after matching for all of the characteristics discussed above and, in the case of the judge-level reversal rates, after taking into account variable rates of appeal.

A3. Sensitivity to Omitted Variables

In addition, these results allow me to check the possibility that the core findings could be attributable to omitted-variable bias—for example, that there is some characteristic in the population of black judges that is not fully captured by the covariates used in the matching. Noting that the sample size here is fairly small (one reason why matching is a useful complement, rather than a substitute, for parametric methods), I employ a method of sensitivity analysis developed by Holland (1986) and implemented by Keele (2010). The methodology allows me to put bounds on how large some characteristic would have to be in order to render the postmatched results insignificant. That is, the methodology allows me to estimate how many times more likely it would have to be that black judges have some characteristic in order to render the results no longer significant—for example, how much more likely they are to write riskier (perhaps more liberal) opinions.

Table A2. Reversal Rate Results after Matching

	Overall (1)	Case Level (2)
Black district judge	.037*	.161+
	(.015)	(.084)
Constant	.293**	.291**
	(.008)	(.056)
Sensitivity upper bound	2.65	3.10
N	184	275

Note. Model 1 presents results from the judge-level data (with reversal rate as the outcome variable) and model 2 presents mixed logit results from the case-level data (with individual case reversal as the outcome variable) using weighted ordinary least squares and mixed logit analysis, respectively, for the reversal rate results. Both match on birth year, age, party, previous judicial experience, previous trial experience, private practice experience, attendance of top law school, JCS score, and circuit. Standard errors are in parentheses

+ $p < .10$.

* $p < .05$.

** $p < .01$.

Results from these sensitivity tests show that black judges would have to be around three times more likely to have some characteristic than white judges in order for the results to be called into question. For example, it could be that black judges are three times more likely to write sharply worded politically oriented opinions than are white judges. I do note that such a trait would have to be present despite controlling via matching for law school attended, JCS score, years of experience, and other characteristics. I also note that there is no clear answer as to what range of sensitivity is acceptable for observational studies; however, these bounds are quite robust compared to the existing literature (for example, Keele 2010). This gives some assurance that unobserved variables are not the exclusive drivers of these results.

REFERENCES

Ashenfelter, Orley, Theodore Eisenberg, and Stewart J. Schwab. 1995. Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes. *Journal of Legal Studies* 24:257–81.

Asmussen, Nicole. 2011. Nontraditional Judicial Nominees: President's Delight and Senators' Dismay? *Legislative Studies Quarterly* 36:591–619.

Bagenstost, Samuel R. 2006. The Structural Turn and the Limits of Antidiscrimi-

nation Law. *California Law Review* 1:5–10.

Baldez, Lisa, Lee Epstein, and Andrew Martin. 2006. Does the U.S. Constitution Need an ERA? *Journal of Legal Studies* 35:243–83.

Banks, R. Richard, Jennifer L. Eberhardt, and Lee Ross. 2006. Discrimination and Implicit Bias in a Racially Unequal Society. *California Law Review* 94:1169–90.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review* 94:991–1013.

Bielby, William T., and James N. Baron. 1986. Men and Women at Work: Sex Segregation and Statistical Discrimination. *American Journal of Sociology* 91:759–99.

Boyd, Christina L. 2010. Federal District Court Judge Ideology Data. Stata 10 format. University of Georgia, Department of Political Science, Athens. http://cLboyd.net/ideology.html.

———. 2013. She'll Settle It? *Journal of Law and Courts* 1:193–219.

Boyd, Christina L., Lee Epstein, and Andrew D. Martin. 2010. Untangling the Causal Effects of Sex on Judging. *American Journal of Political Science* 54:389–411.

Butler, Daniel M., and David E. Broockman. 2011. Do Politicians Racially Discriminate against Constituents? A Field Experiment on State Legislators. *American Journal of Political Science* 55:463–77.

Castilla, Emilio J. 2008. Gender, Race, and Meritocracy in Organizational Careers. *American Journal of Sociology* 113:1479–1526.

Choi, Stephen J., Mitu Gulati, and Eric A. Posner. 2012. What Do Federal District Judges Want? An Analysis of Publications, Citations, and Reversals. *Journal of Law, Economics, and Organization* 28:518–49.

Cox, Adam B., and Thomas J. Miles. 2008. Judging the Voting Rights Act. *Columbia Law Review* 108:1–54.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice*. Cambridge, MA: Harvard University Press.

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland. 2007. The Judicial Common Space. *Journal of Law, Economics, and Organization* 23:303–25.

Fernandez, Roberto M., Emilio J. Castilla, and Paul Moore. 2000. Social Capital at Work: Networks and Employment at a Phone Center. *American Journal of Sociology* 105:1288–1356.

Fryer, Roland G., and Steven D. Levitt. 2004a. The Causes and Consequences of Distinctively Black Names. *Quarterly Journal of Economics* 119:767–805.

———. 2004b. Understanding the Black-White Test Score Gap in the First Two Years of School. *Review of Economics and Statistics* 86:447–64.

Giles, Micheal W., Virginia A. Hettinger, and Todd Peppers. 2001. Picking Fed-

eral Judges: A Note on Policy and Partisan Selection Agendas. *Political Research Quarterly* 54:623–41.

Gill, Rebecca D., Sylvia R. Lazos, and Mallory M. Waters. 2011. Are Judicial Performance Evaluations Fair to Women and Minorities? A Cautionary Tale from Clark County, Nevada. *Law and Society Review* 45:731–59.

Ginther, Donna K., Walter T. Schaffer, Joshua Schnell, Beth Masimore, Faye Liu, Laurel L. Haak, and Raynard Kington. 2011. Race, Ethnicity, and NIH Research Awards. *Science* 333:1015–19.

Gottschall, Jon. 1983. Carter's Judicial Appointments: The Influence of Affirmative Action and Merit Selection on Voting on the U.S. Courts of Appeals. *Judicature* 67:165–74.

Green, Alexander R., Dana R. Carney, Daniel J. Pallin, Long H. Ngo, Kristal L. Raymond, Lisa I. Iezzoni, and Mahzarin R. Banaji. 2007. Implicit Bias among Physicians and Its Prediction of Thrombolysis Decisions for Black and White Patients. *Journal of General Internal Medicine* 22:1231–38.

Greenwald, Anthony G., and Linda Hamilton Krieger. 2006. Implicit Bias: Scientific Foundations. *California Law Review* 94:945–67.

Ho, Daniel E. 2005. Why Affirmative Action Does Not Cause Black Students to Fail the Bar. *Yale Law Journal* 114:1997–2005.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis* 15:199–236.

Holland, Paul W. 1986. Statistics and Causal Inference. *Journal of the American Statistical Association* 81:945–60.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2009. CEM: Software for Coarsened Exact Matching. *Journal of Statistical Software* 30(9):1–27.

———. 2011. Multivariate Matching Methods That Are Monotonic Imbalance Bounding. *Journal of the American Statistical Association* 106:345–61.

Jet. 1993. First Black Federal Judge, James B. Parsons, 81, Dies. July 5, p. 4.

Jha, Ashish K., Elliott S. Fisher, Zhonghe Li, E. John Orav, and Arnold M. Epstein. 2005. Racial Trends in the Use of Major Procedures among the Elderly. *New England Journal of Medicine* 353:683–91.

Kang, Jerry. 2004. Trojan Horses of Race. *Harvard Law Review* 118:1489–1593.

Kastellec, Jonathan P. 2013. Racial Diversity and Judicial Influence on Appellate Courts. *American Journal of Political Science* 57:167–83.

Keele, Denise M., Robert W. Malmsheimer, Donald W. Floyd, and Lianjun Zhang. 2009. An Analysis of Ideological Effects in Published versus Unpublished Judicial Opinions. *Journal of Empirical Legal Studies* 6:213–39.

Keele, Luke. 2010. An Overview of rbounds: An R Package for Rosenbaum Bounds Sensitivity Analysis with Matched Data. White paper. Pennsylvania State University, Department of Political Science, University Park.

Krieger, Nancy, Dana Carney, Katie Lancaster, Pamela D. Waterman, Anna Kosheleva, and Mahzarin Banaji. 2010. Combining Explicit and Implicit Mea-

sures of Racial Discrimination in Health Research. *American Journal of Public Health* 100:1485–92.

Krislov, Samuel. 1974. *Representative Bureaucracy*. Englewood Cliffs, NJ: Prentice-Hall.

Kuersten, Ashlyn K., and Susan B. Haire. 2011. Update to the Appeals Courts Database (1997–2002). University of South Carolina, Judicial Research Initiative, Columbia. http://artsandsciences.sc.edu/poli/juri/appct.htm.

Kulik, Carol T., Elissa L. Perry, and Molly B. Pepper. 2003. Here Comes the Judge: The Influence of Judge Personal Characteristics on Federal Sexual Harassment Case Outcomes. *Law and Human Behavior* 27:69–86.

Lewis, Jeffrey B., and Drew A. Linzer. 2005. Estimating Regression Models in Which the Dependent Variable Is Based on Estimates. *Political Analysis* 13:345–64.

Lott, John R., Jr. 2001. The American Bar Association, Judicial Ratings, and Political Bias. *Journal of Law and Politics* 17:41–61.

Manning, Kenneth L., Bruce A. Carroll, and Robert A. Carp. 2004. Does Age Matter? Judicial Decision Making in Age Discrimination Cases. *Social Science Quarterly* 85:1–18.

Martin, Elaine, and Barry Pyle. 2000. Gender, Race, and Partisanship on the Michigan Supreme Court. *Albany Law Review* 63:1205–36.

Pelander, A. John. 1998. Judicial Performance Review in Arizona: Goals, Practical Effects and Concerns. *Arizona State Law Journal* 30:643–760.

Peresie, Jennifer L. 2005. Female Judges Matter: Gender and Collegial Decision Making in the Federal Appellate Courts. *Yale Law Journal* 114:1759–92.

Pinello, Daniel R. 2003. *Gay Rights and American Law*. New York: Cambridge University Press.

Pitkin, Hanna F. 1967. *The Concept of Representation*. Berkeley: University of California Press.

Poole, Keith T. 1998. Recovering a Basic Space from a Set of Issue Scales. *American Journal of Political Science* 42:954–93.

Sander, Richard H. 2004. A Systemic Analysis of Affirmative Action in American Law Schools. *Stanford Law Review* 57:367–483.

Scherer, Nancy. 2004. Blacks on the Bench. *Political Science Quarterly* 119:655–75.

Scherer, Nancy, and Brett Curry. 2010. Does Descriptive Race Representation Enhance Institutional Legitimacy? The Case of the U.S. Courts. *Journal of Politics* 72:90–104.

Segal, Jeffrey, and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. New York: Cambridge University Press.

Segal, Jennifer A. 2000. Representative Decision Making on the Federal Bench: Clinton's District Court Appointees. *Political Research Quarterly* 53:137–50.

Sen, Maya. 2014a. How Judicial Qualification Ratings May Disadvantage Mi-

nority and Female Candidates. *Journal of Law and Courts* 2:33–65.

———. 2014b. How Minority Judicial Candidates Have Changed, but the ABA Ratings Gap Has Not. *Judicature* 98(4):46–53.

Smelcer, Susan N., Amy Steigerwalt, and Richard L. Vining, Jr. 2011. Bias and the Bar: Evaluating the ABA Ratings of Federal Judicial Nominees. *Political Research Quarterly* 65:827–40.

Songer, Donald B. 2007. The United States Courts of Appeals Database, 1925–1996. University of South Carolina, Judicial Research Initiative, Columbia. http://artsandsciences.sc.edu/poli/juri/appct.htm.

Spohn, Cassia. 1990. The Sentencing Decisions of Black and White Judges: Expected and Unexpected Similarities. *Law and Society Review* 24:1197–1216.

Walker, Thomas G., and Deborah J. Barrow. 1985. The Diversification of the Federal Bench: Policy and Process Ramifications. *Journal of Politics* 47:596–617.

Welch, Susan, Michael Combs, and John Gruhl. 1988. Do Black Judges Make a Difference? *American Journal of Political Science* 32:126–36.

Yinger, John. 1986. Measuring Racial Discrimination with Fair Housing Audits: Caught in the Act. *American Economic Review* 76:881–93.