Is Justice Really Blind? Race and Reversal in U.S. Courts[*†]

Maya Sen[‡]

September 12, 2014

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, a gap in reversal of up to ten percentage points. This gap is robust and persists after taking into account previous professional and judicial experience, educational backgrounds, qualification ratings assigned by the American Bar Association, and differences in appellate panel composition. In total, it suggests 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 judiciary more reflective of the general population, racial disparities within the legal system appear to persist.

[*]A previous version of this paper was the recipient of the 2012 Best Graduate Student Paper Prize, awarded by 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, Kevin Quinn, Joseph Ura, and participants at the University of Chicago Conference on Rational Choice and Decisionmaking for helpful feedback. Thanks also to the Federal Judicial Center and to the Harvard Law Library and Westlaw staff.

[‡]Assistant Professor, Harvard Kennedy School, 79 John F. Kennedy Street, Cambridge, MA 02138 (maya_sen@hks.harvard.edu, http://mayasen.org).

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 John F. Kennedy, and over the course of the call, Kennedy asked Parsons if he would accept a federal judgeship at the U.S. 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.'" The significance of this conversation – an otherwise routine exchange between a President and a potential judicial nominee – was that Parsons was black, and his investiture made him the first African American appointed to the U.S. District Courts.

Thanks to jurists like James 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, judge qualifications (including ratings awarded by the American Bar Association), experi-

ence, and jurisdiction, I find that cases decided by African-American lower court judges are up to 10 percentage points more likely to be overturned than are cases written by similar white judges. This gap is significant, robust, and appears 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. 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. Neither are the results driven by distinctive voting by African Americans 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 factors, possibly differences in lower-court judge ideology, but also the possibility of implicit biases by 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 Parts 2 and 3. In Part 4, I discuss the data, which are two newly collected

data sets and one existing data set on U.S. district judges. Part 5 presents the core results: black judges are indeed more likely to be reversed than white judges, and the difference is robust. In Parts 7–10, 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) possible differences in professional experience, (3) differences in ideology, (4) differences in voting patterns in certain issue areas, and (5) implicit racial bias by higher courts. I conclude in Part 11 with a discussion of the limitations and implications of this research. Additional results from matching analyses are presented in the Appendix.

2 Race and Judicial Decision-Making

Ever since 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 important viewpoints that might otherwise be unshared. Descriptive representation can therefore often (although not always does) 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 that they do so in the context of substantively salient issues. For example, for the minority judges who are the focus of this study,[1] Kastellec (2013) finds that black judges are more likely to vote in favor of affirmative

[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 – e.g., sex discrimination (Boyd, Epstein and Martin, 2010; Baldez, Epstein and Martin, 2006; Peresie, 2005; Massie, Johnson and Gubala, 2002; Segal and Spaeth, 2002; Crowe, 1999). Others have found

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 difference in voting in related civil rights areas. For example, Pinello (2003) finds that black judges are more likely to side with LGBT claimants than white judges, and Martin and Pyle (1999) 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 white judges (but see Spohn (1990), which 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. African Americans 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, 2014*b*). In addition, a number of black judges have historically come from the trenches of the civil rights movement (e.g., Thurgood Marshall, Constance Baker Motley, Matthew Perry), or possibly 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

little or no effects 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.

for different political views (by way of judicial common space 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 is 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 descriptive representation's impact: after all, if these judges have diminished impact due to 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 African Americans 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, 2004*a*), 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 of 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). With regard to minority judges, Gill, Lazos and Waters (2011) have found

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. However, perhaps the only measure of judicial "quality" comes in the form of ratings awarded by the American Bar Association (ABA); here, black judges have been shown to receive lower ratings by some studies (Lott, 2001; Sen, 2014*a*), but not others (Smelcer, Steigerwalt and Jr, 2011). In addition, within public opinion, Scherer and Curry (2010) finds that many perceive black judges to be more liberal.

3 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 (e.g., Epstein, Landes and Posner, 2013; Choi, Gulati and Posner, 2012). Reversal may result in increased workload as judges have to re-visit cases, forcing them to allocate scarce resources 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 assumed 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 black judges vote differently than white judges, 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 African Americans in other fields, and also within law and the legal system, this would suggest the first hypothesis: *African-American judges will be more likely to be reversed on appeal than white judges.* This forms the core inquiry of this study, but would be a finding in contrast to some studies, e.g., Epstein, Landes and Posner (2013), which find no relationship between minority status of 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 comprised of all-white panels, losing litigants in such cases may have some incentives 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 African Americans, 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, rather than any kind of bias by appeals panels.

Parallel narratives would also have observable implications at the reversal stage (i.e., 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 African American 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 3-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), who find 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 (1) issue areas involving civil rights and/or (2) 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 activity 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, 2014*a*), African-American lower-court judges are more likely than white judges to be awarded lower qualification ratings from the American Bar Association (see Smelcer, Steigerwalt and Jr (2011) for no findings on this point); in addition, a lively scholarly debate has addressed the relative successes of African-American 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

	Not Qualified	Qualified	Well Qualified	Exceptionally Well Qualified	N
All	0.01	0.43	0.54	0.02	1653
Whites	0.01	0.41	0.56	0.03	1388
Blacks	0.01	0.57	0.41	0.00	147
Hispanics	0.02	0.61	0.38	0.00	104
Women	0.00	0.49	0.51	0.00	43

Table 1: Distribution of ABA Qualification Ratings for U.S. District Court judges (Johnson to Obama Administrations).

or professional experiences have translated into some African American 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 non-traditional professional backgrounds is extremely difficult to suss out using this kind 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 Bar Association 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 the possibility of implicit bias by appeals panels. To a large extent, a case for implicit bias using observational data is circumstantial, a stubborn difference that persists despite controls and robustness checks. Nonetheless, as noted, a growing literature suggests that implicit bias against African Americans persists (and can be measured) within a wide variety of comparable instances, including not just law (Greenwald and Krieger, 2006; Banks, Eberhardt and Ross, 2006; Bagenstost, 2006; Kang, 2004) and politics (Butler and Broockman, 2011), but also employment (Bertrand and Mullainathan, 2004; Fryer and Levitt, 2004*a*), academia (Ginther et al., 2011), public health (Green et al., 2007), etc. Here, beliefs about blacks' views could also shape biased views. 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, thus driving up their reversal rates. Another possibility is that appeals courts view black judges as less qualified simply because of their race; that is, black judges could be perceived to produce opinions of poorer quality, despite no 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., 2004*b*) and health care (Jha et al., 2005), this is a possibility that must be considered.

4 U.S. District Judge Data

The data come from the two lower tiers of the federal judiciary – the U.S. District Courts and the U.S. 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 demographics.[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 administrations (as of July of 2012) I coded each judge's (1) race or ethnicity, (2) age at confirmation, (3) gender, (4) law school attended, and (5) geographic location (Table 2). I used

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

[3]http://www.fjc.gov/history/home.nsf/page/judges.html.

President	Whites	African Americans	Hispanics	Women	*N*
Barack Obama	0.73	0.17	0.11	0.48	110
George W. Bush	0.82	0.07	0.11	0.21	261
William J. Clinton	0.76	0.18	0.06	0.29	305
George H.W. Bush	0.89	0.07	0.04	0.2	148
Ronald Reagan	0.93	0.02	0.05	0.08	290
Jimmy Carter	0.78	0.14	0.07	0.15	196
Gerald Ford	0.91	0.06	0.02	0.02	49
Richard M. Nixon	0.96	0.03	0.01	0.01	178
Lyndon B. Johnson	0.92	0.05	0.03	0.02	116

Table 2: Racial/ethnic and gender distribution of judicial nominees by President (Johnson through Obama administrations).

automated coding to further assess whether each nominee had previously been (6) a former law clerk, (7) a U.S. Attorney or Assistant U.S. Attorney, (8) a Solicitor General or Deputy or Assistant Solicitor General, (9) a state judge (either a state supreme court or state lower court judge), (10) a former federal judge (e.g., magistrate, territorial, or bankruptcy judge), (11) a full-time law professor or law school dean, (12) an attorney in private practice, or (13) a public defender. I also noted each judge's American Bar Association qualification rating (historically a four-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).

For partisanship, I recorded for each judge the identity of the appointing President as well as his or her judicial common space score, which relies on some combination of the common space score of the appointing President or of the home-state Senators (Boyd, 2011; 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 & World Report* rankings and dividing them into rank cohorts: (1) elite law schools in the "Top 14," (2) schools ranked #15-25, (3) #26-50, (3) #51-76, (4) #76-100, and (5) outside of the top 100. This is a rough measure for judges attending law schools in the 1970s and 80s; however, an assuaging factor is that

	All	Whites	Blacks	Black Democrats	Black Republicans
Ave Age at Investiture	50.06	50.44	48.55	48.16	49.49
Female	0.17	0.15	0.27	0.29	0.21
Nominated by Democrat	0.44	0.40	0.71	1.00	0.00
Top 14 Law School	0.30	0.30	0.28	0.31	0.21
Private Law School	0.52	0.51	0.67	0.69	0.63
Law Clerk	0.21	0.22	0.14	0.16	0.09
Law Professor	0.06	0.05	0.12	0.13	0.07
Private Practice	0.92	0.94	0.76	0.81	0.65
U.S. Attorney	0.09	0.09	0.03	0.03	0.05
Assistant U.S. Attorney	0.20	0.19	0.29	0.25	0.40
Justice Dept Lawyer	0.05	0.05	0.07	0.08	0.07
Public Defender	0.04	0.03	0.10	0.12	0.02
U.S. Magistrate Judge	0.09	0.08	0.10	0.09	0.12
U.S. Bankruptcy Judge	0.01	0.01	0.04	0.05	0.02
State Judge	0.41	0.38	0.55	0.52	0.63
N	1653	1388	147	104	43

Table 3: Demographics of U.S. District Court nominees named after 1960.

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, a commercial legal data base, in its "Judicial Reversal Reports." Included in these reports are (1) the total number of cases, both published and unpublished, for which the district judge wrote an opinion, (2) the total number of cases, published and unpublished, that were appealed for that judge, and (3) how many of these appealed cases, again published and unpublished, were affirmed or reversed. (This is measured as a dichotomous variable, i.e., the case was upheld or it was not.) I used this information to create a data set that includes for all district judges his or her complete reversal rate from 2000 to July of 2012. This final data set includes reversal rates for 1,054 district judges, of whom 945 are white and 109 are African American. The distribution of judges' reversal rates

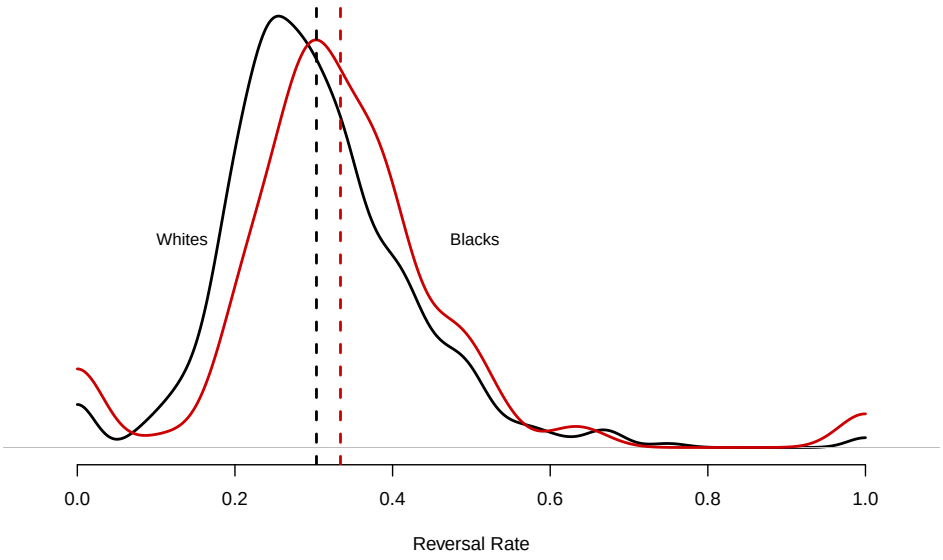

Figure 1: Reversal rates for white and black U.S. District Judges (cases appealed 2000 to July 2012). Dotted vertical lines denote white and black means.

is displayed in Figure 2.[4] The figures include judges who had very few cases appealed (an issue I address in the methodological discussion, below), resulting in some judges reporting 0 or 100 percent of cases reversed.

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 Motion Reports by their aggregate nature do not contain this information, I examine an extant data set of appealed cases collected by Songer 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 subset these data to more contemporary periods, leaving me with 1,722 cases decided between 1996 and 2002 (or 7,289 cases decided between 1960 and

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

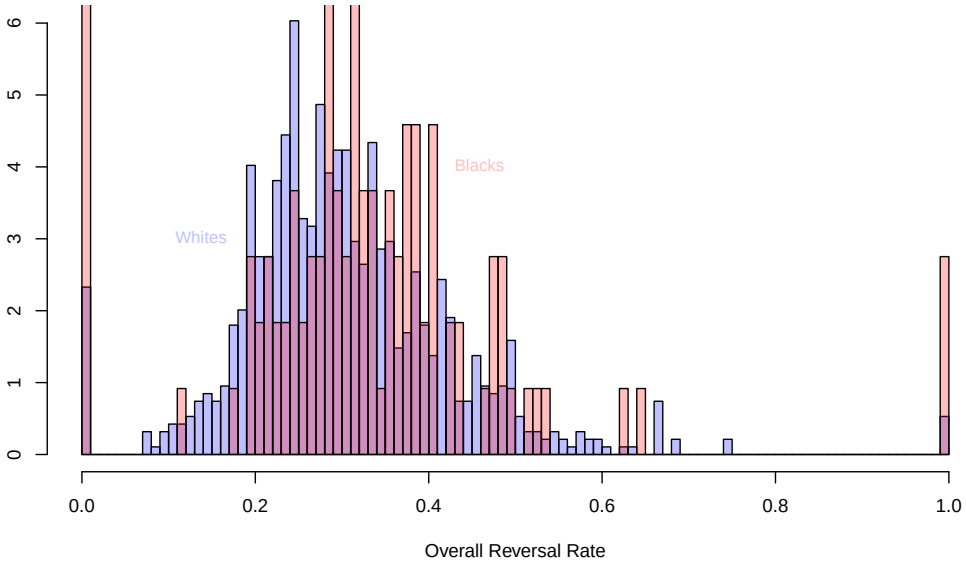

Figure 2: Reversal rates for white and black U.S. District Judges (cases appealed 2000 to July 2012).

2002, for at least one model specification in Table **??**). These data also include information on the three judges hearing the appeal, its substantive legal issue area (e.g., "Civil Rights," "Criminal Law," "Economic Activity," "Labor"), and whether the lower-court opinion was (1) upheld or (2) reversed. I note that, although perfect overlap with the Westlaw data is preferable, the Kuersten and Haire (2011) data only go through 2002. In addition, as noted by others, these data also include only published cases, which could skew the sample (Epstein, Landes and Posner, 2013).[5]

Here, a fact helpful to identification is that incoming cases in (1) district courts and in (2) the appeals stage are assigned to judges (or panels) on a fairly random basis. Although the randomization can be informal, this longstanding practice makes it impermissible for

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

federal judges to request to hear particular kinds of cases. Thus, conditional on jurisdiction, (1) cases heard by black lower-court judges should on average be similar to those heard by white judges (i.e., there should be balance in case characteristics between cases heard by black judges versus those heard by white judges) and (2) appeals panels hearing cases written by black judges should on average be similar to appeals panels hearing cases written by white judges (that is, potentially biased judges cannot request to hear cases decided by lower-court black judges). I present results that suggest that the randomization is working below. I also control for issue area and other case attributes in the case-level analysis; the substantive results are unaffected.

5 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 do 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 2000-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 versus a judge serving the entirety of 2000-2012. An ordinary least squares 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 1 case appealed or 180. For the reversal data, I therefore take a weighted least squares (WLS) approach by weighing 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 (e.g., Epstein, Landes and Posner, 2013).

For the analyses looking at the case-level Kuersten and Haire (2011) data, I use a logit

specification, with the outcome variable being whether a case was reversed (1) or upheld (0). Here, 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 guard 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 assures that the results are not driven by a lack of common support in the data. Because these are largely consistent with parametric results, these are presented in the Appendix.

Tables 4 (judge-level reversal rates) and 5 (case-by-case reversals) present the core results. Focusing on Table 4, the effect for black district judge is always positive and statistically significant, ranging in magnitude from 2% to 3%. Substantively this means that black judges have a reversal rate that is between 2 and 3 percentage points higher 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 (Table 4 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, in 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 non-Southern jurisdictions (Models 8). I also include include controls for confirmation year (Models 5–7), which addresses the

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

	Judge-level Reversal Rates (Continuous 0 to 1)						
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**
Black District Judge	0.03***	0.03***	0.03***	0.02***	0.02***	0.02***	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female District Judge					0.001	0.0002	0.0001
					(0.01)	(0.01)	(0.01)
Age					0.01***	0.01***	0.01***
					(0.003)	(0.003)	(0.003)
Commission Year					0.01	0.01	−0.01
					(0.28)	(0.28)	(0.29)
Commission Year^2					−0.0000	−0.0000	0.0000
					(0.0001)	(0.0001)	(0.0001)
Female * Black District Judge						0.002	0.01
						(0.02)	(0.02)
South							0.003
							(0.01)
South * Black District Judge							0.02
							(0.02)
Constant	0.31***	0.32***	0.31***	0.48***	−8.75	−9.99	13.60
	(0.003)	(0.02)	(0.02)	(0.02)	(283.40)	(283.71)	(284.43)
President Dummies		✓	✓	✓	✓	✓	✓
Circuit Dummies			✓				
District Dummies				✓	✓	✓	✓
N	1054	1054	1054	1054	1054	1054	1054
Adj. R-squared	0.01	0.01	0.31	0.48	0.49	0.49	0.49

***p < .01; **p < .05; *p < .1

Table 4: Weighted OLS regression results, U.S. District Court judges' reversal rates (continuos between 0 and 1) for cases decided between 2000-2012.

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 respective to a random subset of published cases, which could be a biased subsample of all cases.) Cases written by black lower-court judges are more likely to be reversed on appeal, and the effect is robust to the inclusion of district-level dummies and dummy variables for the appointing President.[7] In terms of predicted probabilities, Table 5 actually suggests even stronger effects: a case written by a black judge on average may have an approximately 10 percentage point greater

[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 (experience) the judge had before hearing the case in question.

chance of being reversed than a case written by a white judge (depending on the model). In addition, I compare whether the effect is an artifact of the time period in question by examining the spread of data available from Songer et al, dating back to the era when the first black judge was nominated in 1961 (Model 5). The effect appears regardless of whether we examine only cases from 1996 moving forward *or* going back to 1960. In addition, an interaction between race of lower court judge and time period in question (for example, pre- or post-1996) is not significant (not shown), suggesting that we 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 cases 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 cases would have been reversed on appeal instead of having been upheld.) At the individual level, these results mean that each African American judge on average had approximately 20 additional cases (out of an average of 196 cases appealed) reversed 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 we include District Court dummies (Table 4, Models 4–7) or Circuit dummies (Table 4, Model 3). However, the explanatory power from the model increases swiftly when we include any kind of control for district. (Table 4's overall R^2 moves from around 0.01 to around 0.5, Models 4–6.) I move forward in most of the remaining analyses by including dummy variables for district. This has the additional benefits of (1) controlling for fluctuations in case dockets among districts (e.g., the Southern District of California versus the District for Alaska), and also of (2) allowing the assumption that cases

	Case-level Reversal (0 or 1)				
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Black District Judge	0.49***	0.40**	0.49**	0.51**	0.24**
	(0.18)	(0.19)	(0.21)	(0.21)	(0.11)
Female District Judge				0.22	0.17*
				(0.18)	(0.11)
Age				−0.01	0.05*
				(0.07)	(0.03)
Case Year				−43.15	3.28**
				(67.93)	(1.42)
Case Year^2				0.01	−0.001**
				(0.02)	(0.0004)
Constant	−1.15***	−1.06***	−0.82***	43099.92	−3243.15**
	(0.06)	(0.13)	(0.28)	(67892.18)	(1414.47)
Judge-Specific Random Effects	✓	✓	✓	✓	✓
President Dummies		✓	✓	✓	✓
District Court Dummies			✓	✓	✓
N	1718	1718	1718	1718	7279
Log Likelihood	−962.47	−958.94	−899.77	−898.28	−4233.36

***p < .01; **p < .05; *p < .1

Table 5: Logit regression regression results, case-by-case basis. Outcome variable is whether case reversed (1) or upheld (0). Data are randomly selected subset of published cases (from Songer et al), 1996-2002 (Models 1–4) and 1960–2002 (Model 5). Judge random effects included.

are randomly assigned to be made more safely (as cases are usually assigned randomly, but only within district). Second, the results are significant regardless of whether we 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 because the overall reversal rate data is both more conservative and, more importantly, because it includes the universe of appealed cases, not just a random subset of published cases. Lastly, these core results are consistent with additional matching analyses, which are 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 are hearing cases that vary in number and type according to the race of the lower-court judge.

6 Mechanism 1: Differences in Cases Appealed

The first explanation I consider is that cases authored by black lower-court judges will be more likely to be appealed, perhaps as a result of differences in black-white voting, or from some kind of bias against black judges by legal practitioners (as suggested by the state-courts literature, e.g., 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, having more cases appealed from black judges may mean that more of them will be reversed; likewise, if more civil rights cases are appealed from black judges, and if 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 race of lower-court judge. For these analyses, the outcome variable is the number of cases each district judge had appealed 2000-2012.[8] To take into account the fact that some judges produce more or fewer authored opinions, 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 that norms about appeal vary across jurisdictions, I also include specifications that include dummy variables for jurisdiction in Models 3–5 (in the form of district, although dummy variables for appeals court result in identical inferences).

Surprisingly, Table 6 shows that we cannot rule out that there is no difference between

[8]I note that these figures do not take into account the number of type of cases a judge had settled. As some scholarship has shown (e.g., 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.

	Number of Cases Appealed				
	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		0.12***	0.14***	0.10***	0.09***
		(0.01)	(0.01)	(0.01)	(0.01)
Female District Judge					−11.77
					(8.97)
Age					−24.43***
					(3.67)
Commission Year					2503.98***
					(368.88)
Commission Year^2					−0.63***
					(0.09)
Constant	188.05***	125.09***	122.09***	20.51	−2491447.00***
	(5.30)	(6.81)	(23.95)	(23.95)	(367739.40)
District Court Dummies			✓	✓	✓
President Dummies				✓	✓
N	1054	1054	1054	1054	1054
R-squared	0.0002	0.14	0.47	0.62	0.65
Adj. R-squared	−0.001	0.14	0.42	0.58	0.61

***p < .01; **p < .05; *p < .1

Table 6: OLS regression results. Outcome variable is total number of cases appealed from each judge (between 2000 and 2012).

black and white judges in terms of rates of appeal: under all model specifications, the substantive difference between black 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 we have 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 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.

Circuit	p-value	Significant?
DC	0.18	No
1st	0.21	No
2nd	0.10	No
3rd	0.99	No
4th	0.19	No
5th	0.15	No
6th	0.59	No
7th	0.43	No
8th	0.68	No
9th	0.13	No
10th	0.71	No
11th	0.12	No

Table 7: Fisher's exact tests of difference between black and white judges' cases across legal issue area, using cases from Kuersten and Haire (2011).

However, the other concern is perhaps that the number of cases might not vary, but the *kind* of case does. Here, the particular concern that black judges' rulings on racially salient judges 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. Specifically, 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 seven 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 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 that attorneys either (1) appeal from opinions written by black judges at higher rates, or (2) 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 due to the number and nature of the kinds of cases being appealed.

7 Mechanism 2: Possible Differences in "Quality"

The FJC data show that black and white judges differ on average in terms of some 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% compared to 94%), while more have state judge experience (55% compared to 38%). Black judges are also less likely to receive higher qualification ratings from the American Bar Association; 41% of them receive a high "Well Qualified" rating compared to 56% of white judges. (Discrepancies in qualification ratings are explored in Lott (2001) and Sen (2014*a*), with Smelcer, Steigerwalt and Jr (2011) providing evidence to the contrary.) However, there is disagreement in the literature on whether objective criteria can ever determine what makes a "good" judge, whether such criteria are useful predictors of reversal, and whether some factors (e.g., 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 also subjective measures such as ABA ratings.

As in Table 4, the outcome variable for this analysis is the judge's reversal rate from 2000-2012, across all cases (published and unpublished). I again include dummy variables for the (1) identity of the appointing President and (2) 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 ("T14") law school, considered the elite group; a large number of judges attended one of these schools, particularly Harvard (121 judges) and Yale Law Schools (54 judges).) Model 3 includes professional experience, such as whether the judge had been in private practice or was a law clerk. Model 4 includes whether the judge had served in a judicial capacity before, perhaps as a federal magistrate or state judge (either state supreme or state lower court judge), 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 their reversal rate of approximately two 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 didn't, those who were U.S. attorneys and those who weren't, etc. The only exception, perhaps, are former state judges (who are more likely to be reversed) and former law clerks (slightly less). The models are by observational standards fairly predictive, with R^2 values close to 0.50. However (as before), most of the explanatory power comes not from the professional or educational variables, but rather 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, we must consider whether other unmeasurable or qualitative traits are driving the results. Here, 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

	Judge-level Reversal Rates (Continuous 0 to 1)				
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Black District Judge	0.02***	0.02***	0.02***	0.02**	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female District Judge	0.001	0.002	0.003	0.002	0.001
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Law School Ranked 15-25		0.01	0.005	0.005	0.003
		(0.01)	(0.01)	(0.01)	(0.01)
Law School Ranked 26-50		0.001	0.0000	−0.0005	0.0003
		(0.01)	(0.01)	(0.01)	(0.01)
Law School Ranked 51-75		0.02	0.01	0.01	0.01
		(0.01)	(0.01)	(0.01)	(0.01)
Law School Ranked 76-100		−0.01	−0.02*	−0.02*	−0.02*
		(0.01)	(0.01)	(0.01)	(0.01)
Law School Ranked 101+		0.0003	−0.0000	−0.001	−0.001
		(0.01)	(0.01)	(0.01)	(0.01)
Law Clerk		−0.01**	−0.01*	−0.01*	−0.01
		(0.01)	(0.01)	(0.01)	(0.01)
Law Professor			−0.003	−0.002	−0.004
			(0.01)	(0.01)	(0.01)
Private Practice			0.01	0.01	0.01
			(0.01)	(0.01)	(0.01)
U.S. Attorney			−0.01	−0.01	−0.01
			(0.01)	(0.01)	(0.01)
Assistant U.S. Attorney			0.0001	0.001	0.001
			(0.01)	(0.01)	(0.01)
Justice Department			−0.02	−0.02	−0.02
			(0.01)	(0.01)	(0.01)
Public Defender			0.02*	0.02*	0.02*
			(0.01)	(0.01)	(0.01)
Federal Magistrate				0.01	0.01
				(0.01)	(0.01)
Federal Bankruptcy Judge				0.01	0.01
				(0.02)	(0.02)
State Judge				0.01**	0.01**
				(0.01)	(0.01)
Commission Year					0.02
					(0.28)
Commission Year^2					−0.0000
					(0.0001)
Constant	0.53***	0.54***	0.54***	0.53***	−16.27
	(0.04)	(0.04)	(0.04)	(0.04)	(284.03)
ABA Ratings Dummies	✓	✓	✓	✓	✓
President Dummies	✓	✓	✓	✓	✓
District Court Dummies	✓	✓	✓	✓	✓
N	1054	1054	1054	1054	1054
R-squared	0.54	0.54	0.55	0.55	0.55
Adj. R-squared	0.48	0.49	0.49	0.49	0.49

***p < .01; **p < .05; *p < .1

Table 8: Weighted OLS regression results, U.S. District Court judges' reversal rates (continuous between 0 and 1) for cases decided between 2000-2012.

rank, 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 this 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 (with these tests presented in the Appendix). The second is that few of the prestige-oriented variables that are included in the federal data are actually predictive at all of reversal rates: being a U.S. Attorney, graduating from a top law school, or even having served as a full-time law professor or dean do little to predict judges' reversal rates. Being a law clerk is only fleetingly so, 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 prestige-oriented variables (e.g., law review membership) would result in similar non-significance.

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 U.S. Attorney positions without having achieved some combination of high class rank (GPA), law review membership, and bar passage. Presumably controlling for these professional experiences also controls to some extent for these unrecorded traits. Lastly, 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 via a candidate's ABA rating, and the results are robust to the inclusion of this variable. Taken together, this leads me to reject this possible mechanism as an exclusive explanation behind the findings.

8 Mechanism 3: More Liberal Voting

The results suggest that the racial gap persists despite controlling for differences in backgrounds – including both objective experiences and subjective qualification ratings. Perhaps

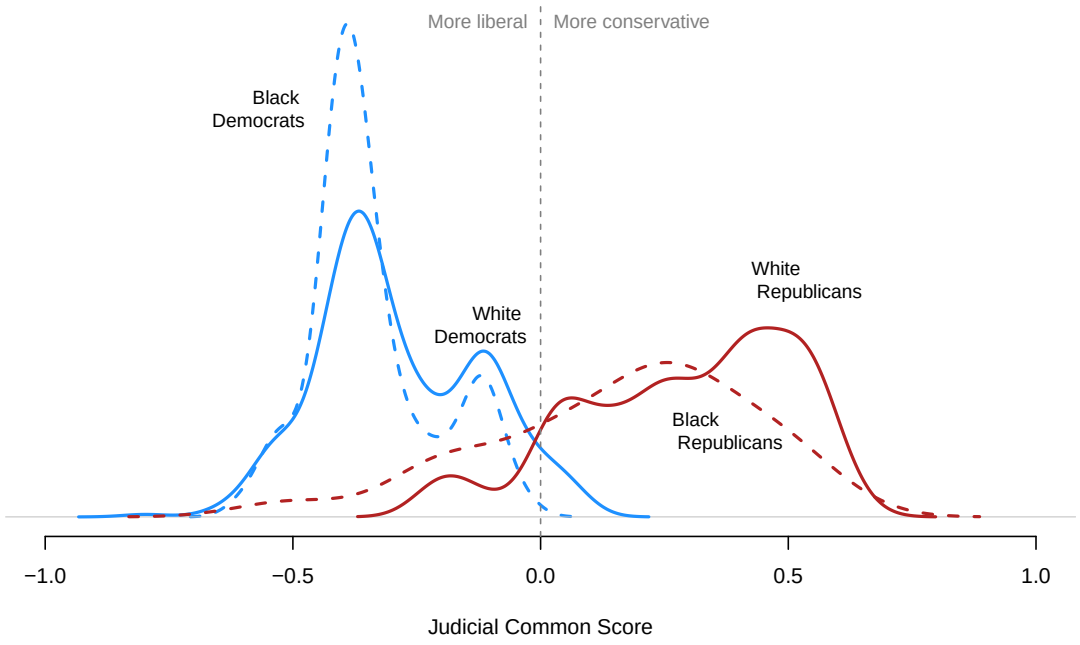

Figure 3: Judicial common space scores of black and white district judges, by race and party

a more likely possibility is that black judges are more liberal across the board than are comparable white 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 African Americans, 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 3, which shows the judicial common space scores of black and white judges by party (Boyd, 2011). Black judges have more left-leaning judicial common space scores, which in turn raises the possibility that they write opinions that are more liberal and hence reversed at higher rates by more centrists appeals panels. Again, this could be the case despite appointment by the same President.

	Judge-level Reversal Rates (Continuous 0 to 1)			
	Model 1	**Model 2**	**Democrats Only**	**Republicans Only**
Black District Judge	0.02**	0.02*	0.03**	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Republican District Judge	−0.07*			
	(0.04)			
Judicial Common Space Score (JCS)		−0.01		
		(0.02)		
Republican * Black District Judge	−0.01			
	(0.02)			
JCS * Black District Judge		−0.01		
		(0.03)		
Commission Year	0.02	0.38	−0.14	0.29
	(0.29)	(0.80)	(0.93)	(0.29)
Commission Year^2	−0.0000	−0.0001	0.0000	−0.0001
	(0.0001)	(0.0002)	(0.0002)	(0.0001)
Constant	−19.28	−378.25	140.33	−282.28
	(284.31)	(791.23)	(923.12)	(293.49)
ABA Ratings Dummies	✓	✓	✓	✓
President Dummies	✓	✓	✓	✓
District Court Dummies	✓	✓	✓	✓
Educational Controls	✓	✓	✓	✓
Professional Controls	✓	✓	✓	✓
N	1054	648	457	597
Adj. R-squared	0.49	0.50	0.40	0.57

***p < .01; **p < .05; *p < .1

Table 9: Weighted OLS regression results, U.S. District Court judges' reversal rates (continuous between 0 and 1) for cases appealed between 2000-2012.

Testing district judge ideology directly. To more closely analyze the role of lower-court ideology (or at least as directly as possible, given current ideological measures), Table 9 explores reversal rates with four model specifications: (1) district judge party and race interacted, (2) district judge judicial common space score and race interacted,[9] and separately for (3) district judges appointed by a Democrat, and for (4) district judges appointed by a Republican.

To be clear, the possibility exists that minority judges appointed by the same President are still more liberal. However, the analysis does not rule out that there are no differences

[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 1.

between Democrat and Republican appointments in terms of reversals among black judges (Model 1); an interaction between race of district judge and party of appointing President is not significant under any model specifications. In addition, the judicial common space score interacted with the race of the lower-court judge is also not significant (Model 2).[10] However, what Table 9 does suggest is that the effect appears driven by Democratic appointed judges. Among Democratic 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 for the variables associated with prior experience and also to the inclusion of dummy variables for appointing President and district court. Among Republican judges, however, the effect is close to 1 percentage point and never significant. Neither 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 "T14" law schools and served as law clerks.)

Testing 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, African American judges are more liberal and therefore more likely to be reversed. (The results are actually not consistent with a parallel story, one in which black Democrats are more liberal, and black Republicans more conservative – i.e., minority candidates are more "extreme." If this story was true, then we would see the effect regardless of party, and not for Democrats only.) Unfortunately, existing measures of lower-court

[10]This is not surprising, as judicial common space scores capture either the ideology of the appointing President or the senior home-state Senator (or some combination of the two home-state Senators), depending on the partisan alignment (Boyd, 2011). 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 judges' beliefs.

	Case-level Reversal					
	All Judges			Democrats Only		
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Black District Judge	1.23***	1.24***	0.27	1.62***	1.53**	0.39
	(0.36)	(0.47)	(0.34)	(0.49)	(0.63)	(0.42)
Meal Panel Ideology (Conservative)	−0.14			0.35		
	(0.40)			(0.71)		
Republican Panel		−0.11			−0.17	
		(0.22)			(0.41)	
Conservative Decision			−1.58***			−1.40***
			(0.14)			(0.25)
Black District Judge * Mean Ideology	−0.44			−1.06		
	(1.00)			(1.25)		
Black District Judge * Republican Panel		−0.11			−0.17	
		(0.56)			(0.74)	
Black District Judge * Conservative Decision			0.64			0.76
			(0.43)			(0.54)
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 Squared	−0.02	0.003	0.004	0.03	0.02	0.01
	(0.03)	(0.03)	(0.02)	(0.05)	(0.04)	(0.03)
Constant	−72097.20	11938.10	15992.03	106853.50	86370.12	49848.34
	(118853.30)	(101069.60)	(72848.71)	(195157.30)	(168782.70)	(122726.10)
District Court Dummies	✓	✓	✓	✓	✓	✓
President Dummies	✓	✓	✓	✓	✓	✓
N	669	934	1639	272	368	627
Log Likelihood	−304.26	−425.33	−786.84	−124.58	−163.33	−302.24

***p < .01; **p < .05; *p < .1

Table 10: Logit regression regression results, case-by-case basis. Outcome variable is whether case reversed (1) or upheld (0). Data are randomly selected subset of published cases (Songer et al), 1996-2002. Judge random effects included.

ideology do not allows us to rule out this possibility conclusively.

I therefore turn to some indirect tests. Specifically, I examine the appeals case-level data from Kuersten and Haire (2011) analyzed in Table 5. If the effect is driven by black lower-court judges being more liberal than whites, then we would expect that (1) not only would black Democrats be more likely than white Democrats to be reversed (as we see in Table 8), but also that (2) the difference between black and white judges in terms of reversal rates increases when black judges' opinions are reviewed by conservative 3-judge appeals panels. That is, the gap between black and white judges should grow 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 (1) the mean JCS score of the appeals panel, (2) whether the appeals panel had two or more Republican appointments (e.g., was a Republican majority), and (3) 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 also, for ease of interpretation, Democrats only (Models 4–6). All of the models further 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 we would expect a positive relationship in the interaction terms. That is, if black judges are voting in a more liberal direction compared to their white counterparts, then Republican appeals panels (or panels ruling in a more conservative direction) would be more likely to be reversing 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 we restrict the sample to those for whom the effects are the strongest (black and whiteDemocrats, Models 4–6); that is, we 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* with when the 3-judge panel becomes more conservative. The explanation behind this is unclear, and none of the interactions is significant. However, it does provide additional evidence against the theory that the fact that black judges' increased reversal is being driven by these black judges being more liberal across the board: we cannot rule out that the increased reversal of black judges does not vary according to the composition or ruling of the appeals panel.

9 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 may expect black judges to be overruled most frequently in these areas, where their views might differ the most from whites. Or, we may expect that black judges may 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 (e.g., 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 seven of Kuersten and Haire et al'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, we would see a significant relationship in the interaction of black lower-court authorship and civil rights issues. The results show, however, that the effect of black authorship on a case's probability of being upheld actually varies little by civil rights issue area: there is no difference in blacks' reversal rates across civil rights cases and non-civil rights cases (Table 11, 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 different legal categorizations.

	Case-level Reversal (0 or 1)				
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Black District Judge	0.50**	0.51**	1.18***	1.15***	1.29***
	(0.21)	(0.23)	(0.29)	(0.33)	(0.30)
Civil Rights Case	−0.27*	−0.27			
	(0.16)	(0.18)			
Black District Judge * Civil Rights Case		−0.06			
		(0.47)			
1 Black on Panel			−0.39	−0.53*	
			(0.28)	(0.31)	
2 Blacks on Panel			−0.09	0.87	
			(0.72)	(0.78)	
Majority Black Panel					1.03
					(0.77)
Black District Judge * 1 Black on Panel				0.80	
				(0.70)	
Black District Judge * 2 Blacks on Panel				−20.11	
				(4554.49)	
Black District Judge * Majority Black Panel					−20.25
					(4575.08)
Case Year	−40.23	−39.77	−5.64	−3.30	
	(67.93)	(67.97)	(101.12)	(101.77)	
Case Year^2	0.01	0.01	0.001	0.001	
	(0.02)	(0.02)	(0.03)	(0.03)	
Constant	40182.47	39720.22	5595.49	3250.84	−0.93**
	(67899.46)	(67935.23)	(101065.00)	(101711.60)	(0.38)
District Court Dummies	✓	✓	✓	✓	✓
President Dummies	✓	✓	✓	✓	✓
N	1717	1717	934	934	934
Log Likelihood	−897.47	−897.46	−424.52	−420.53	−422.56

***p < .01; **p < .05; *p < .1

Table 11: Logit regression regression results, case-by-case basis. Outcome variable is whether case reversed (1) or upheld (0). Data are randomly selected subset of published cases (Songer et al), 1996-2002. Judge random effects included.

10 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 implication and challenges the fairness and race neutrality of the judiciary. On the other, such a finding would perhaps be unsurprising, as studies have teased out implicit biases against racial minorities in prominent economic,

social, and political settings.

Here, I consider one possible manifestation – whether having a black presence on an appeals panel attenuates the effect. This could happen as white judges become more sensitive to any possible discriminatory tendencies, or as black judges raise possible concerns about bias. (On this point, Kastellec (2013) provides evidence that having a black judge on an appeals panel will change the way that the panel votes on affirmative action issues.) To test this theory, I evaluated how the racial gap varies across different racial appeals panel compositions – zero, one, or two black judges on the three-judge higher-court panels.[11] As before, I use mixed-effect logit model, with an interaction between race of the lower-court judge and the number of black circuit judges hearing the appeal (Model 4). I also include an interaction between race of lower-court judge and whether African Americans constituted a majority of the panel (two or more, Model 6).No cases in the data were heard by an all-black 3-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 on the appeals (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 doesn't 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.

[11]Note that there are very few cases involving two or more African Americans on the same appeals panel. For example, in the Songer et al data for cases decided between 1960 and 2002, 88% had no black judges on the 3-judge appeals panel, 11% had one black judge, and just under 1% had two.

11 Conclusion

The results show that discrepancies exist in how appeals courts review cases, with black judges being up to ten percentage points more likely to be reversed than whites. This racial gap is robust and persists once we control for possible proxies for judicial "qualifications" – e.g., quality of legal education, age, professional experience, and American Bar Association ratings. Controls for the partisanship of the lower-court judge, as well as 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 judge does not appear to vary across issue area and across jurisdictions. And this gap translates into meaningful legal outcomes. Close to 3,000 federal court cases would still be on the books if black judges were overturned at whites' lower rates. At the individual level, African American judges on average have up to 20 additional 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, we have suggestive evidence that black judges are more liberal than otherwise similar whites, but the fact that we can't 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 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 also perhaps due to implicit biases by 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 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 John F. Kennedy, American Presidents have actively sought to appoint judges of color – not just African Americans, but also Hispanics, and Asian Americans – to the nation's highest courts. At the state and international level, too, efforts are underway to increase the proportion of judges from under-represented 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 touched upon just one singular ascriptive characteristic: the race (black or not) of lower court judges. Whether a judge is African American is, however, just one facet of judicial identity, and we may think that the 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 we think that heuristics or personal familiarity may play a role in how appeals panels reach decisions, then maybe we would find different rates of overturning between judges who attended the same law school or are otherwise knowledgeable or friendly – that is, that a personal connection strengthens 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 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 into these issues would further shed light into why black judges are more likely to be overruled, and why this racial gap is so persistent.

12 Appendix: Matching Analyses

Because black and white judges differ in their 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 within 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 within 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, resulting 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 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.

Matching Methodology. To implement the matching, I use coarsened exact matching (Iacus, King and Porro, 2011, 2009), which allows exact matching on key variables and coarsening and then matching approximately 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 (1) male or female, (2) a Republican or Democratic appointee, (3) a former federal magistrate or bankruptcy judge, (4) a former attorney in practice practice, (5) a graduate of a Top Tier ("T14") law school, and (6) his or her corresponding appeals court.[13] I further coarsen an additional set variables using specific cut points. These are (7) birth year, (8) number of years either on a federal bench (e.g., as a magistrate judge) or in private practice, and (9) the judge's common space score (Boyd, 2011).

A summary of some of the judge characteristics post-matching on the judge-level data is given by Table 12. This matched sample of judges is, as expected, slightly different than the

[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 (i.e., black judges). I present the results from coarsened exact matching, as it bounds the maximal amount of imbalance through the choice of coarsenings (Iacus, King and Porro, 2011, 2009).

[13]Matching on district court, which would be ideal, is not possible due to the sample size.

	All Whites	All Blacks	Matched Whites	Matched Blacks
Female	0.15	0.27	0.28	0.28
Appointed by Democrat	0.40	0.71	0.68	0.68
Top 14 Law School	0.30	0.28	0.23	0.26
Former Federal Judge	0.09	0.14	0.07	0.07
Private Practice	0.94	0.76	0.90	0.90
Ave Commission Year	1987.49	1991.95	1995.83	1995.75
Trial Years	16.75	10.08	13.68	12.28
N	1388	147	172	72

Table 12: Demographics of matched district court judges compared to the entire population of district court judges.

original pre-matched sample (Table 3) but by no means anomalous. Overall, more judges in the matched sample were appointed by Democrats (specifically 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 experience before nomination. We also have more female judges in the matched sample.

Matching Results. I run matching analyses twice. The first is on the judge-level data presented in Columns 1 and 2 of Table 13, while the second is on the Songer et al case-level data from 1960 to 2002. The judge-level post-matching results are estimated using a weighted OLS (with the weights coming from the number of cases appealed), while the case-level post-matching results are calculated using a mixed-effect logit (because of the binary nature of the outcome variable), with judge-specific random effects. The results are, however, substantially similar. Post matching, reversal rates of black judges are approximately 3 to 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, taking into account variable rates of appeal.

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 – e.g., that

	Overall Reversal Rate (WOLS)	Case Level Reversal (Mixed Logit)
Black District Judge	0.037**	0.161*
	(0.015)	(0.084)
Constant	0.293***	0.291***
	(0.008)	(0.056)
N	184	275

*p < .1; **p < .05; ***p < .01

Table 13: Post-matching results. Model 1 presents results form the judge-level data (with reversal rate as the outcome variable); Model 2 presents results from the case-level data (with individual case reversal as the outcome variable). Both match on birth year, age, party, previous judicial experience, previous trial experience, private practice experience, top law school attended, judicial common space score, and Circuit.

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 post-matched results insignificant. That is, the methodology allows me to estimate how many times more likely black judges would have to have some characteristic in order to render the results no longer significant – e.g., how much more often they are to write riskier (perhaps more liberal) opinions.

Results from these sensitivity tests are presented in Table 14. What these results show is that black judges would have to have some characteristic around three times more 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, common space score, years of experience,

	Overall Reversal Rate	Case Level Reversal
Sensitivity Upper Bound	2.65	3.10
N	184	275

Table 14: Sensitivity analyses results. Column 1 presents results form the judge-level data (with reversal rate as the outcome variable); Column 2 presents results from the case-level data, 1960-2002 (with individual case reversal as the outcome variable). Both results obtained after matching on birth year, age, party, previous judicial experience, previous trial experience, private practice experience, top law school attended, judicial common space score, and Circuit.

etc. 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 (e.g. Keele, 2010). This gives some assurance that unobserved variables are not the exclusive drivers of these results.

References

Ashenfelter, O., T. Eisenberg and S.J. Schwab. 1995. "Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes." *The Journal of Legal Studies* 24(2):257–281. 4

Asmussen, Nicole. 2011. "Nontraditional Judicial Nominees: President's Delight and Senators' Dismay?" *Legislative Studies Quarterly* 36(4):591–619. 27

Bagenstost, Samuel R. 2006. "The Structural Turn and the Limits of Antidiscrimination Law." *California Law Review* 1:5–10. 5, 9

Baldez, Lisa, Lee Epstein and Andrew Martin. 2006. "Does the U.S. Constitution Need an ERA?" *Journal of Legal Studies* 35(1):243–283. 3

Banks, R. Richard, Jennifer L. Eberhardt and Lee Ross. 2006. "Discrimination and Implicit Bias in a Racially Unequal Society." *California Law Review* 94:1169. 5, 9

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(4):991–1013. 5, 9

Bielby, W.T. and J.N. Baron. 1986. "Men and Women at Work: Sex Segregation and Statistical Discrimination." *American Journal of Sociology* pp. 759–799. 5

Boyd, Christina L. 2011. "Federal District Court Judge Ideology Data." University of Georgia.
URL: *http://cLboyd.net/ideology.html* 11, 27, 29, 38

Boyd, Christina L. 2013. "She'll Settle It?" *Journal of Law and Courts* 1(2):193—219. 20

Boyd, Christina L., Lee Epstein and Andrew D. Martin. 2010. "Untangling the Causal Effects of Sex on Judging." *American Journal of Political Science* 54(2):389–411. 3, 16, 37

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(3):463–477. 5, 9

Castilla, E.J. 2008. "Gender, Race, and Meritocracy in Organizational Careers." *American Journal of Sociology* 113(6):1479–1526. 5

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(3):518–549. 6

Cox, Adam B. and Thomas J. Miles. 2008. "Judging the Voting Rights Act." *Columbia Law Review* 108(1):1–54. 2, 4, 6, 8, 32

Crowe, Nancy E. 1999. "The Effects of Judges' Sex and Race on Judicial Decision Making on the U.S. Courts of Appeals, 1981-1996.". 3

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, and Organization* 23(2):303–325. 11

Epstein, Lee, William M. Landes and Richard A. Posner. 2013. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice*. Harvard University Press. 6, 7, 8, 13, 14, 15

Fernandez, R.M., E.J. Castilla and P. Moore. 2000. "Social Capital at Work: Networks and Employment at a Phone Center." *American Journal of Sociology* pp. 1288–1356. 5

Fryer, Roland G. and Steven D. Levitt. 2004*a*. "The Causes and Consequences of Distinctively Black Names." *Quarterly Journal of Economics* 119(3):767–805. 5, 9

Fryer, Roland G. and Steven D. Levitt. 2004*b*. "Understanding the Bblack-White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 86(2):447–464. 10

Giles, Micheal W., Virginia A. Hettinger and Todd Peppers. 2001. "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas." *Political Research Quarterly* 54(3):623–641. 11

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(3):731–759. 5, 7, 20

Ginther, D.K., W.T. Schaffer, J. Schnell, B. Masimore, F. Liu, L.L. Haak and R. Kington. 2011. "Race, Ethnicity, and NIH Research Awards." *Science* 333(6045):1015–1019. 5, 9

Gottschall, J. 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. 4

Green, A.R., D.R. Carney, D.J. Pallin, L.H. Ngo, K.L. Raymond, L.I. Iezzoni and M.R. Banaji. 2007. "Implicit Bias Among Physicians and its Prediction of Thrombolysis Decisions for Black and White patients." *Journal of General Internal Medicine* 22(9):1231–1238. 9

Greenwald, A.G. and L.H. Krieger. 2006. "Implicit Bias: Scientific Foundations." *California Law Review* 94(4):945–967. 5, 9

Ho, Daniel E. 2005. "Why Affirmative Action Does Not Cause Black Students to Fail the Bar." *Yale Law Journal* 114(8):1997–2005. 8, 24

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* . 16, 37

Holland, Paul W. 1986. "Statistics and causal inference." *Journal of the American Statistical Association* 81(396):945–960. 40

Iacus, S., G. King and G. Porro. 2009. "CEM: Software for Coarsened Exact Matching.". 38

Iacus, S.M., G. King and G. Porro. 2011. "Multivariate Matching Methods That are Monotonic Imbalance Bounding." *Journal of the American Statistical Association* . 38

Jha, A.K., E.S. Fisher, Z. Li, E.J. Orav and A.M. Epstein. 2005. "Racial Trends in the Use of Major Procedures Among the Elderly." *New England Journal of Medicine* 353(7):683–691. 10

Kang, Jerry. 2004. "Trojan Horses of Race." *Harvard Law Review* 118:1489. 5, 9

Kastellec, Jonathan P. 2013. "Racial Diversity and Judicial Influence on Appellate Courts." *American Journal of Political Science* 57(1):167–183. 2, 3, 6, 8, 32, 34

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(1):213–239. 14

Keele, Luke. 2010. "An Overview of `rbounds`: An `R` package for Rosenbaum Bounds Sensitivity Analysis with Matched Data." *White Paper. Columbus, OH* pp. 1–15. 40, 41

Krieger, N., D. Carney, K. Lancaster, P.D. Waterman, A. Kosheleva and M. Banaji. 2010. "Combining explicit and implicit measures of racial discrimination in health research." *American Journal of Public Health* 100(8):1485. 5

Krislov, Samuel. 1974. *Representative Bureaucracy*. Prentice-Hall. 3

Kuersten, Ashlyn K. and Susan B. Haire. 2011. "Update to the Appeals Courts Database (1997–2002)." *available at: http://www.wmich.edu/nsf-coa* . 2, 13, 14, 15, 22, 30, 32

Kulik, C.T., E.L. Perry and M.B. Pepper. 2003. "Here comes the judge: The influence of judge personal characteristics on federal sexual harassment case outcomes." *Law and Human Behavior* 27(1):69–86. 4

Lewis, Jeffrey B. and Drew A. Linzer. 2005. "Estimating Regression Models in which the Dependent Variable is Based on Estimates." *Political Analysis* 13(4):345–364. 15

Lott, John L. 2001. "The American Bar Association, Judicial Ratings, and Political Bias." *Journal of Law & Politics* 17:41. 6, 8, 23

Manning, K.L., B.A. Carroll and R.A. Carp. 2004. "Does Age Matter? Judicial Decision Making in Age Discrimination Cases." *Social Science Quarterly* 85(1):1–18. 4

Martin, Elaine and Barry Pyle. 1999. "Gender, Race, and Partisanship on the Michigan Supreme Court." *Albany Law Review* 63:1205. 4

Massie, T., S.W. Johnson and S.M. Gubala. 2002. The Impact of Gender and Race in the Decisions of Judges on the United States Courts of Appeals. In *Annual meeting of Midwest Political Science Association, Chicago IL*. 3

Pelander, A. John. 1998. "Judicial Performance Review in Arizona: Goals, Practical Effects and Concerns." *Arizona State Law Journal* 30:643. 5

Peresie, Jennifer L. 2005. "Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts." *Yale Law Journal* 114(7):1759–1892. 3

Pinello, Daniel R. 2003. *Gay Rights and American Law*. Cambridge University Press. 4

Pitkin, Hanna F. 1967. *The Concept of Representation*. University of California Press. 3

Poole, K.T. 1998. "Recovering a Basic Space from a Set of Issue Scales." *American Journal of Political Science* pp. 954–993. 11

Sander, Richard H. 2004. "A Systemic Analysis of Affirmative Action in American Law Schools." *Stanford Law Review* pp. 367–483. 8, 24

Scherer, Nancy. 2004. "Blacks on the Bench." *Political Science Quarterly* 119(4):655–675. 4

Scherer, Nancy and Brett Curry. 2010. "Does Descriptive Race Representation Enhance Institutional Legitimacy? The Case of the U.S. Courts." *Journal of Politics* 72(1):90–104. 3, 6, 10, 35

Segal, Jeffrey and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. Cambridge University Press. 3

Segal, Jennifer A. 2000. "Representative Decision Making on the Federal Bench: Clinton's District Court Appointees." *Political Research Quarterly* 53(1):137. 4

Sen, Maya. 2014*a*. "How Judicial Qualification Ratings May Disadvantage Minority and Female Candidates." *Journal of Law and Courts* 1(2):33–65. 6, 8, 23

Sen, Maya. 2014*b*. "How Minority Judicial Candidates Have Changed, but the ABA Ratings Gap Has Not." In Press, *Judicature*.
URL: *http://j.mp/1odyDyl* 4

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* . 6, 8, 23

Songer, Donald B. 2007. "The United States Courts of Appeals Database, 1925–1996." *available at: http://artsandsciences.sc.edu/poli/juri/appct.htm* . 2

Spohn, Cassia. 1990. "The Sentencing Decisions of Black and White Judges: Expected and Unexpected Similarities." *Law & Society Review* 24(5):1197–1216. 4

Walker, Thomas G. and Deborah J. Barrow. 1985. "The Diversification of the Federal Bench: Policy and Process Ramifications." *The Journal of Politics* 47(02):596–617. 4

Welch, S., M. Combs and J. Gruhl. 1988. "Do Black Judges Make a Difference?" *American Journal of Political Science* 32(1):126–136. 4

Yinger, John. 1986. "Measuring Racial Discrimination with Fair Housing Audits: Caught in the Act." *The American Economic Review* 76(5):881–893. 5