

Racial Diversity and Judicial Influence on Appellate Courts

Jonathan P. Kastellec Princeton University

This article evaluates the substantive consequences of judicial diversity on the U.S. Courts of Appeals. Due to the small percentage of racial minorities on the federal bench, the key question in evaluating these consequences is not whether minority judges vote differently from nonminority judges, but whether their presence on appellate courts influences their colleagues and affects case outcomes. Using matching methods, I show that black judges are significantly more likely than nonblack judges to support affirmative action programs. This individual-level difference translates into a substantial causal effect of adding a black judge to an otherwise all-nonblack panel. Randomly assigning a black counterjudge—a black judge sitting with two nonblack judges—to a three-judge panel of the Courts of Appeals nearly ensures that the panel will vote in favor of an affirmative action program. These results have important implications for assessing the relationship between diversity and representation on federal courts.

What are the consequences of judicial diversity? While white males still occupy a majority of judgeships in American courts, the increasing number of female and minority judges makes this question a central one in the study of judicial politics. Given the diversity of the citizenry subject to the power of the judiciary in the United States, whether the presence of women and racial minorities on the bench influences legal outcomes and the development of the law are central questions when evaluating the representativeness of courts and their legitimacy.

Since the diversification of the federal bench began three decades ago, numerous scholars have explored its consequences by asking whether women and racial minorities tend to vote differently from white male judges in several areas of the law. While results have varied, on the whole these literatures suggest that in issue areas salient to either women or minorities, such as civil rights cases, female and minority judges tend to vote more liberally than white male judges. These studies demonstrate that female and minority judges, on average, bring a different judicial perspective to the bench.

While the study of individual differences in judicial voting across race and gender is certainly worthwhile, the consequences of such differences are mitigated by the current demographic distributions on U.S. courts, which are still largely stocked by white males. This mitigation is particularly acute with respect to both racial minorities as a group and appellate courts as institutions. Despite the commitment of recent presidents to appoint more minorities to the federal bench, minority judges today occupy less than 20% of the active seats on the Courts of Appeals, the level of appellate courts below the U.S. Supreme Court in the federal system. Because appellate courts are multimember courts, with cases decided by panels of judges, individual differences in voting may not necessarily lead to any differences in *case outcomes*, due to the fact that a minority judge is likely to be outnumbered on any given panel. Thus, whether judicial diversity has large-scale consequences depends on whether it leads to differences not just in individual voting by judges but also to differences in case outcomes, which is what litigants care about and what shapes the development of legal doctrine in a system of *stare decisis*.

Jonathan P. Kastellec is Assistant Professor, Department of Politics, 39 Corwin Hall, Princeton University, Princeton, NJ 08544 (jkastell@princeton.edu).

I thank the following people for helpful comments and suggestions: Deborah Beim, Christina Boyd, Charles Cameron, Tom Clark, Jeronimo Cortina, Paul Frymer, Jeffrey Lax, Jee-Kwang Park, Maya Sen, and Christopher Wlezien, along with seminar participants at Columbia University. I also thank Herschel Nachlis for excellent research assistance and the Mamdouha S. Bobst Center for Peace and Justice at Princeton University for research support. Earlier versions were presented at the 2011 meeting of the Midwest Political Science Association and the 2011 Conference on Empirical Legal Studies. Replication materials can be found on the Dataverse Network at <http://hdl.handle.net/1902.1/17948>.

American Journal of Political Science, Vol. 00, No. 0, xxx 2012, Pp. 1–17

© 2012, Midwest Political Science Association

DOI: 10.1111/j.1540-5907.2012.00618.x

While a number of recent studies have demonstrated the existence of gender-based “panel effects” on the Courts of Appeals—that is, that the presence of a single female judge on a three-judge panel increases the likelihood that male judges on the panel will vote liberally—less attention has been paid to the possibility of race-based panel effects. This comparative neglect is unfortunate, given that the small proportion of racial minorities on the Courts of Appeals (relative to the proportion of women) means that the existence of race-based panel effects is even more important for assessing the substantive implications of judicial diversity.

In this article, I evaluate the substantive consequences of judicial diversity by examining both whether black judges vote differently than nonblack judges in affirmative action cases *and* whether black judges affect case outcomes by influencing the voting of nonblack judges in these cases. I first discuss why such influence is important for assessing the quality of substantive representation on multimember courts. Then, drawing on literatures from political science, economics, social psychology, and organizational science, I advance three possible mechanisms by which black judges could influence their nonblack colleagues. Using matching methods, I show that African American judges are significantly more likely than *comparable* nonblack judges to support affirmative action programs. More importantly, this individual-level difference translates into a substantial causal effect of adding a black judge to an otherwise all-nonblack panel: the presence of a black judge increases the probability that a nonblack judge will rule in favor of an affirmative action program by about 20 percentage points. In fact, the random assignment of a black *counterjudge*—a black judge sitting with two nonblack judges—to a three-judge panel of the Courts of Appeals nearly ensures that the panel will vote in favor of an affirmative action program. It is thus clear that the substantive consequences of racial diversity on the Courts of Appeals are quite large.

Judicial Diversity, Substantive Representation, and Appellate Courts

Social scientists have long been concerned with the political implications of gender and racial diversity and the incorporation of racial minorities into the political process (Key 1949; Myrdal 1944). A key question here is how best to achieve the representation of minority interests, and, in consideration of this, a large literature has flowed from Pitkin’s (1967) important distinction between descriptive and substantive representation. As summarized

by Farhang and Wawro (2004, 301), “A political body or institution is descriptively representative by literally resembling or reflecting the constituent elements of the community that it governs. In contrast, substantive representation is concerned with what the representative actually does on behalf of the interests of the group he or she is associated with.” The distinction between these two types of representation has been most thoroughly explored in the legislative arena, particularly with respect to the question of whether the use of majority-minority districts in the U.S. House of Representatives serves the interests of black Americans (Swain 1993).

Since the diversification of the federal judiciary began in earnest in the late 1970s, following President Carter’s efforts to recruit more women and minorities to the bench, political scientists and legal scholars have addressed the rise in diversity from the perspectives of both descriptive and substantive representation. Studying the former with respect to the judiciary is straightforward. While until the 1970s white males comprised more than 95% of the federal bench, the last three decades have seen a significant rise in descriptive representation. In fact, as of 2010, the percentage of African Americans on the Courts of Appeals (11%) nearly matched the percentage of blacks in the U.S. population (12.6%, based on the 2010 U.S. Census). Hispanics comprise 13% of the population; 7% of judges on the Courts of Appeals are Hispanic.¹ Women, however, as of 2010 comprised only 30% of seats on the circuit courts.² There is, of course, variation in descriptive representation across time and different courts, which scholars have examined by studying which institutional and political factors predict more female and minority appointments (see, e.g., Hurwitz and Lanier 2008). There is also evidence that increased descriptive representation may enhance the legitimacy of courts among traditionally underrepresented groups (Scherer and Curry 2010).

Studying substantive representation in judicial institutions is less straightforward—particularly on federal courts, given that federal judges, who are unelected and serve with life tenure, are not directly accountable to the public and are not generally expected to follow

¹ Data on the distribution of judges is compiled from *History of the Federal Judiciary* (2011). In this article, I use the term “Hispanic” to refer to persons in the United States who can trace their ancestry to the Spanish-speaking regions of Latin America and the Caribbean. The only other minority group with representation on the Courts of Appeals is Asians—there have been four Asian judges in the history of the Courts of Appeals, one of whom was sitting in 2010.

² Note the categories are not mutually exclusive: a Hispanic female judge, for example, is counted as both Hispanic and female. See the online appendix for more details.

public opinion. It is well established, however, that experience and ideology can influence judges in some cases, and it is thus natural to ask whether judges with different backgrounds might “represent” citizens from similar backgrounds by siding with their interests in judicial decisions. Thus, most scholars have conceptualized substantive representation on courts by focusing on judicial votes and the extent to which female and minority judges tend to vote differently from nonblack male judges. Focusing on race, it is likely that black judges and nonblack judges will approach the bench with different experiences and views related to race-related legal issues (Washington 1994). Harry Edwards, a black judge on the D.C. Circuit, has stated that it is “inevitable that judges’ different professional and life experiences have some bearing on how they confront various problems that come before them” (2002). He continues: “Because of the long history of racial discrimination and segregation in American society, it is safe to assume that a disproportionate number of blacks grow up with a heightened awareness of the problems that pertain to these areas of the law” (328). In 2009, Sonia Sotomayor’s confirmation put this question front-and-center in discussions of the American judiciary, as she found herself defending comments about her “hope that a wise Latina woman with the richness of her experiences would more often than not reach a better conclusion than a white male who hasn’t lived that life” (Sotomayor 2002, 92).

Does the weight of evidence support the conclusion that a difference in judicial voting across groups exists? Yes—in the areas of the law where we would expect the backgrounds of judges to potentially influence their decision making.³ At the district court level, Ashenfelter, Eisenberg, and Schweb (1995), for example, find no differences between minority and nonminority judges in a vast set of civil rights and prisoners cases, while Chew and Kelly (2006) find significant differences in racial harassment cases. Similarly, in their comprehensive study of individual gender effects across 13 areas of the law, Boyd, Epstein, and Martin (2010) only find significant differences in how men and women vote in sex discrimination cases.⁴ Thus, to oversimplify, women and minorities tend to be more liberal than white males, but

they are exceptionally liberal (above and beyond what simple ideology would predict) only in legal issues where we would expect race or gender to be influential.

Because trial court judges are solitary decision makers, the question of whether female and minority judges vote differently speaks directly to the scope of substantive representation on those courts. If a black district court judge votes in favor of a black plaintiff in an employment discrimination case, where a nonblack judge hearing the same case would not have, the substantive effect of racial diversity is immediate (although the possibility of appellate reversal still exists). Thus, the question of individual vote differences is the proper analytical focus for evaluating substantive representation on trial courts.

The mapping from votes to representation on appellate courts is complicated by the fact that they are multimember courts. On the Courts of Appeals, cases are heard by panels of three judges, who decide cases by majority rule. Recognizing the institutional consequences of panel decision making, scholars in recent years have shifted from studying judicial votes on three-judge panels in isolation toward studying them in the context of a judge’s colleagues. In many areas of the law, Courts of Appeals judges tend to vote differently depending on which judges they happen to sit with in a given case. This phenomenon is broadly referred to as “panel effects,” meaning that a judge’s vote depends on the composition of the panel.

One line of inquiry has focused on *partisan-based panel effects*, showing the likelihood of a liberal vote by an individual judge decreases with every Republican judge added to a panel, and vice versa (Cross and Tiller 1998; Kastellec 2011b; Revesz 1997; Sunstein et al. 2006). More specifically, Kastellec (2011a) introduces the notion of a partisan “counterjudge”: a single judge from the opposite party of the two other judges on a panel. A counterjudge effect can be said to exist when the presence of a counterjudge influences the other two judges on the panel. More relevant to this article, a second line of inquiry has focused on *characteristic-based panel effects*. In particular, several recent studies have demonstrated the existence of gender-based panel effects by showing that adding a female judge to an otherwise all-male panel significantly increases the probability that the male judges will support a plaintiff in sex discrimination or sexual harassment cases (Boyd, Epstein, and Martin 2010; Farhang and Wawro 2004, 2010; Peresie 2005). Here it is useful to extend the concept of counterjudging to include gender or minority counterjudges: a woman sitting with two men can be described as a “female counterjudge,” while a black judge sitting with two nonblack judges can be described as a “black

³ For excellent reviews of the literature on gender differences and racial differences, see Boyd, Epstein, and Martin (2010) and Chew and Kelly (2006), respectively.

⁴ Boyd, Epstein, and Martin (2010) do not find significant differences in sexual harassment cases, a somewhat puzzling result given the nature of these cases. However, in a broader dataset of sexual harassment cases, Farhang and Wawro (2010) do find significant gender differences.

counterjudge.”⁵ Conversely, a man sitting with two females would be a “male counterjudge.” While these studies do not use this language, what they show is the existence of substantive female counterjudge effects.

As Farhang and Wawro (2004) note, the existence of such effects has significant implications for the scope of substantive representation on appellate courts. Because a female judge may cause her male colleagues to vote differently than they would in her absence, the presence of a *single* female counterjudge significantly increases the possibility of substantive representation by affecting *case outcomes*, and not just individual votes. If male judges tended not to vote in favor of women’s interests while female judges did, but no female counterjudge effects existed, then women would tend to influence case outcomes only in cases where they comprised a majority of the panel. Many more panels consist of one female judge than two.⁶ Thus, of greater interest in assessing substantive representation on appellate courts is not whether women vote differently from men, but, first, whether men alone provide substantive representation (a point I return to shortly), and second, if they do not, *whether women induce counterjudge effects by causing men to vote differently when they sit with a female counterjudge*.

While the existence of gender-based panel effects has been well documented, the possibility of race-based counterjudge effects has been explored in only one issue area; Cox and Miles (2008a, 2008b) find that the addition of an African American counterjudge to a three-judge panel increases the likelihood that a nonblack colleague will find that a state or locality violated the Voting Rights Act.⁷

⁵ The use of this term should not be taken as an assumption that the views of black and nonblack judges necessarily run counter to one another (just as Democratic and Republican judges agree on many cases). Rather, “counterjudge” is a useful and general way to describe any type of judge who differs from the other two judges on any given dimension of judicial characteristic, without implicating the actual votes of the judges. In contrast, the term “minority judge” could simultaneously describe a racial minority judge, a political minority judge, or a judge in the voting minority. “Counterjudge” avoids such confusion.

⁶ In the dataset analyzed in Farhang and Wawro (2010), 16% of panels contained at least one woman, while only 6% contained more than one (only two cases had three female judges).

⁷ This article differs from the excellent analysis in Cox and Miles in the following ways. First, by evaluating a new issue area (affirmative action), I show that race-based effects extend beyond Voting Rights cases. Second, while Cox and Miles (2008b) offer a brief discussion of the possible mechanisms driving race-based counterjudge effects, I extensively discuss three possible mechanisms, including one that is new to the literature on characteristic-based panel effects. I also conduct an indirect adjudication of the mechanisms, which is also new to the literature. Third, this article directly places the question of race-based counterjudge effects within the

Indeed, these articles comprise the only published work to examine and uncover such effects.⁸ This comparative neglect of attention on race is unfortunate, given that the prospect of individual voting differences translating into substantive differences in terms of case outcomes is even more remote, when we consider the small number of racial minorities on the Courts of Appeals and the subsequent fact that minority judges will almost always be counterjudges. Hypothetically, if nonblack judges voted monolithically for the employer in racial harassment cases, and black judges voted monolithically for the plaintiff, the viewpoint of the African American judges would only carry the day in the exceedingly rare instances where two black judges were assigned to a panel. The low number of black judges on the Courts of Appeals, however, makes this an unlikely event. In 2008, for example, out of 154 active judges on the Courts of Appeals, 14 were black. No more than two African American judges served on any one circuit. And in the dataset of affirmative action cases I analyze below, a panel comprises two black judges in just two out of 182 cases, or 1.1% of the time.

It is worth noting the assumption implicit in most studies of representation in appellate courts that descriptive representation is a *necessary* condition for substantive representation, even if not sufficient, because nonblack male judges alone cannot provide substantive representation for traditionally unrepresented goods. For example, in their study of the influence of race on voting in state supreme courts, Bonneau and Rice argue: “If there are no differences between nonblack and African American judges, then it also means that there is no substantive representation on the bench. After all, if African American judges are deciding cases the same way as nonblack judges, then neither group is representing the interests of minorities” (2009, 382). However, if nonblack judges uniformly supported the interests of minorities, as did black judges, that would result in a maximization of substantive representation, even in the absence of differences in voting between the groups. Similarly, if nonblacks acted in this manner in a judiciary without *any* minority judges, we would observe a similar maximization in the

context of the larger literature in political science on diversity and representation. Finally, this article employs matching procedures; as I discuss below, matching does not alter the general conclusions I reach, but it helps put bounds on the scope of our inferences. It also allows me to conduct a sensitivity analysis to check for “hidden bias” (see the online appendix).

⁸ In addition, an unpublished paper by Cameron and Cummings (2003) finds race-based panel effects of limited magnitude in affirmative action cases.

complete absence of descriptive representation. Thus, when evaluating voting differences among group differences, it is worth evaluating whether some type of substantive representation would be achieved in the absence of such differences.

The Possible Mechanisms of Judicial Influence

As discussed above, the mechanisms underlying the differences in individual judicial decision making across judges of different races (and gender) are fairly straightforward, as they flow directly from the well-known result that judges with different backgrounds and ideologies will tend to approach at least some cases differently. Understanding the mechanisms underlying race-based counterjudge effects, however, is more complicated. In this section, I discuss three possible mechanisms through which the presence of a black judge on a three-judge panel could cause a nonblack judge to vote differently in a case—one of which is new to the literature on panel effects.⁹ While it is not possible to completely adjudicate between the possible mechanisms, for reasons I explain below, a secondary analysis of nonblack Republican judges provides some suggestive evidence.

Deliberation. Perhaps the most intuitive mechanism of judicial influence is a deliberation or informational effect, in which black judges offer arguments or information in support of their preferred outcome in a case—information tied to their unique life or professional experiences—that cause their colleagues to think differently about a case. The role of diversity in affecting group deliberations has been explored in depth by social psychologists and organizational scientists. The effects of diversity are complex, as heterogeneity can lead to group divisiveness and hence negative group performance (Mannix and Neale 2005). However, in many situations, diversity can lead to increased information sharing and improvements in complex thinking overall (Antonio et al. 2004; Phillips and Loyd 2006), leading group members to evaluate a task differently. Thurgood Marshall, the first black Supreme Court justice, argued that his arrival on the Court led to his colleagues obtaining new information. “What do they know about Negroes? You can’t name one member of this court who knows anything about Negroes before he came to this court” (Liptak 2009).

⁹ In keeping with the theme of the article, I motivate the discussion in terms of race-based counterjudge effects. It could easily be applied to gender effects as well, however.

Unfortunately, we cannot directly observe the arguments Courts of Appeals judges make to each other before they issue their decisions. But, to see how this information effect might work in practice, consider the experiment run by Sommers (2006), which was set in a legal context. The author randomly assigned mock jurors into either all nonblack juries or juries with four nonblack and two black members, then had them engage in a mock jury deliberation based on a simulated trial in which a black defendant was charged with sexual assault. The deliberations of the members of the diverse juries were longer on average and evaluated a larger range of information. More provocatively, nonblack participants “were largely responsible for the influence of racial composition [on deliberation], as they raised more case facts, made fewer factual errors, and were more amenable to discussion of race-related issues when they were members of a diverse group” (Sommers 2006, 606). Thus, the arguments made by black judges on three-judge panels may not just present new information that otherwise would not emerge; they may systematically change the way nonblack judges evaluate a case.

Votes. Perhaps the most straightforward way in which a judge can influence her colleagues is simply through her vote. That is, the fact that a judge votes in one direction may cause one or both of her colleagues to vote in the same way. For example, a judge may disagree with her colleagues in the panel majority, but rather than cast a dissenting vote (which may be a costly activity), she decides to go along with them—thus, her vote is directly influenced by the other judges’ votes, and nothing more. Drawing on the economics literature of social interactions (Manski 1993), which studies how individuals acting in group settings are affected by their peers or neighbors, Fischman (2011) develops a model of “endogenous effects” to explain panel effects, in which one judge’s votes may be endogenous to those of his colleagues. In this account, nonblack judges would be more likely to go along with the position of a black colleague simply because the colleague votes liberally in a case, just as a teen might choose to smoke solely because his friends are smoking as well (Powell, Tauras, and Ross 2005). Such “consensus voting” likely helps explain the high rate of unanimity on the Courts of Appeals.

Presence. Both the deliberation and votes mechanisms are based on *actions* black judges undertake on a three-judge panel. There is a potential, however, for the mere *presence* of a black judge on a three-judge panel to cause nonblack judges to approach a case differently. In the social economics literature, one type of peer influence falls under the heading of “contextual,” in which an

individual's actions are influenced by the *characteristics* of other group members, such as race or gender. A presence effect is equivalent to a contextual effect. Fischman argues: "It is difficult to state a plausible theory of panel effects that relies solely on contextual effects. Such a theory would require that a judge's . . . personal characteristics would affect her colleagues' votes irrespective of the position that the judge actually took in the case" (2011, 6). While it is true that contextual and endogenous effects work in tandem, it is plausible that the characteristics of a counterjudge could influence the panel, *independent of the votes cast in the case*.

Returning to the Sommers (2006) experiment, the participants in the mock jury trial cast a preliminary vote after viewing the trial simulation; the votes were indicated on a pre-deliberation questionnaire. Thus, these votes were recorded *after* each member knew the racial composition of the jury she was sitting on, but *before* any deliberation or public voting had taken place. Nonblack jurors on mixed-race juries were significantly *less* likely to decide that the black defendant was guilty in the pre-deliberation votes compared to nonblack jurors on homogenous juries. Thus, there was a *presence* effect from adding black jurors to the group that was independent of votes or deliberation. In this account, simply sitting with a black judge on a three-judge panel causes nonblack judges to vote more liberally. To the best of my knowledge, this mechanism has not been advanced in the panel effects literature.

In fact, we have anecdotal evidence of a presence effect occurring on the U.S. Supreme Court. Reflecting on the tenure of Justice Marshall, Justice Antonin Scalia echoed Marshall's sentiment that his arrival changed the dynamics on the Court. "Marshall could be a persuasive force just by sitting there," Scalia said. "He wouldn't have to open his mouth to affect the nature of the [justices' private] conference and how seriously the [justices] would take matters of race" (Liptak 2009).

Adjudicating between mechanisms. Ideally, one could adjudicate among these mechanisms in an empirical study of race-based counterjudge effects. Unfortunately, a clean test is not possible or feasible given observational data. First, while we can observe votes and characteristics, we cannot observe the deliberation or arguments among Courts of Appeals judges. Second, and more consequentially, Manski (1993) demonstrates that it is nearly impossible to separately identify contextual and endogenous effects.¹⁰ Accordingly, in the analysis below, I focus mainly on identifying the causal *effect* of adding a black

counterjudge to a three-judge panel. I do, however, provide some suggestive evidence that votes alone do not seem to be the main mechanism underlying the effects I uncover.

Evaluating Affirmative Action Cases

To study the influence of race on the Courts of Appeals, I turn to affirmative action cases decided over the last three decades. Affirmative action has been one of the most contested areas of the law in which race is salient, as evidenced by a series of divided Supreme Court decisions in this period (Kellough 2006, chaps. 5–6). This contestation has been reflected in mass public opinion: since Presidents Kennedy and Johnson issued executive orders formally calling for "affirmative action" in hiring, affirmative action policies have largely focused on promoting the advancement of black Americans, who have tended to support affirmative action programs much more strongly than whites (Sigelman 1991; Steeh and Krysan 1996). For example, from 1985 to 2005, the percentage of black respondents who believed that blacks should be given preference in hiring and promotion due to past discrimination ranged roughly from 45% to 65%; among whites, the percentage supporting such preferences never exceeded 20% in that period (Le and Citrin 2008, 177). While federal judges are certainly not representative of the public at large, it is likely that black judges and nonblack judges will, on average, carry different views about affirmative action to the bench.

Because of the ambiguity of many of the Supreme Court's decisions, the Courts of Appeals have played a large role in the development of affirmative action law as they have sought to fill in gaps in the Supreme Court's doctrine (Bhagwai 2001–2). For example, following the Supreme Court's 1995 decision in *Adarand Constructors, Inc. v. Peña* (515 U.S. 200), in which the Court ruled that all race-conscious government programs were subject to strict scrutiny analysis, the appellate courts were left to flesh out which programs rose to a sufficiently "compelling" level in order to survive strict scrutiny (Bhagwai 2001–2, 263–70). With respect to affirmative action programs at public higher education institutions, this ambiguity was not resolved until eight years later in the Court's decisions in *Grutter v. Bollinger* (539 U.S. 306) and *Gratz v. Bollinger* (539 U.S. 244). Thus, if race-based counterjudge effects exist in these cases, they may have

¹⁰ Fischman (2011) develops separate models of contextual and endogenous effects, each of which assumes the absence of the others. While it would be possible to employ his model of endogenous

effects using the data I analyze below, the model cannot distinguish between majority acquiescence to the counterjudge or vice versa; identifying such acquiescence is the very goal of my analysis.

had a large impact on the development of affirmative action law on the Courts of Appeals.

To answer these questions, I combined existing data with original data in an attempt to collect the universe of published large-scale affirmative action cases in the Courts of Appeals from 1971 to 2008.¹¹ I began with the set of cases analyzed in Sunstein et al. (2006), which covered the years from 1980 to 2003. While Sunstein et al. examined and found partisan-based panel effects in these cases, they did not study race-based voting in any way. These cases involve constitutional challenges to race-based affirmative action programs by governments, employers, and universities. I exclude all affirmative action cases in which race is not an issue, such as cases focusing solely on gender-based affirmative action.¹² (They also do *not* involve statutory claims, such as under Title VII of the 1964 Civil Rights Act, of race-based discrimination by employers, which is a qualitatively different area of the law than affirmative action.) After reexamining all of the cases in the Sunstein et al. dataset, I backdated the dataset to 1971 and updated it through 2008. This process resulted in 182 cases heard by three-judge panels, or 546 judge-votes. In a typical case, the court is tasked with evaluating whether an affirmative action program, such as with respect to school admissions or hiring decisions, violates the 14th Amendment's equal protection clause by unfairly favoring black citizens over nonblack citizens. Following Sunstein et al., I coded a decision as "conservative" (coded 0) if a judge voted to strike down any part of an affirmative action plan as unconstitutional and "liberal" (coded 1) if a judge upheld the program in its entirety. (A complete description of the case selection procedures and variable coding can be found in the online appendix.)

I collected a battery of information on each case and each judge, including the party of the appointing president for each judge, as well as their race.¹³ Of the

546 judge-votes, 32 were cast by African American judges, or about 6%; those 32 votes were cast by 18 unique judges. In terms of panel composition, 152 of the cases (84%) were heard by all nonblack panels; 28 cases (15%) featured a single black judge, while two cases (1%) featured two black judges on the panel. Not a single case was heard by an all-black panel. All told, of the 514 votes cast by non-black judges, 58 were cast when sitting with a black colleague, while 456 were cast when sitting on all-nonblack panels. While the sample size of votes cast by black judges altogether and nonblack judges sitting with black judges is relatively low, this will lead to larger standard errors of the estimates and thus make it more difficult to uncover individual differences and panel effects that are statistically different from zero.

Figure 1A presents a model-free look at the voting rates in the affirmative action cases, broken down in a number of ways. The left plot depicts mean voting rates, at the level of the case (i.e., the panel level). The right plot depicts mean voting rates at the individual-judge level; the horizontal lines show 95% confidence intervals. Beginning with the case-level data, 59% of panels issued liberal decisions upholding affirmative action programs. Evaluating the differences in partisan-based voting by panel composition provides a useful benchmark by which to compare the race-based effects. Consistent with Sunstein and colleagues' (2006) evaluation, when examining unified panels it is clear that voting on affirmative action cases is polarized by partisan affiliation: 75% of cases heard by all-Democratic panels were decided liberally, compared to only 33% of cases heard by three Republican judges. In another indication of polarization, 18% of cases featured a dissent, a rate much higher than is usual on the Courts of Appeals. Thus, it is clear that affirmative action is a more politically charged area of the law than most; this should be recognized when evaluating the generalizability of the results to other race-related issues. At the same time, the high political salience of affirmative action cases, along with the high rate of dissents, makes it more difficult to uncover race-based counterjudge effects. The existence of these effects is dependent on a high degree of "consensus" voting on three-judge panels, which contributes to the "norm of unanimity" on the Courts of Appeals (Fischman 2010). In other words, if judges are frequently dissenting, the likelihood of them being influenced by their colleagues in a set of cases decreases, making it less likely that panel effects exist in those cases.

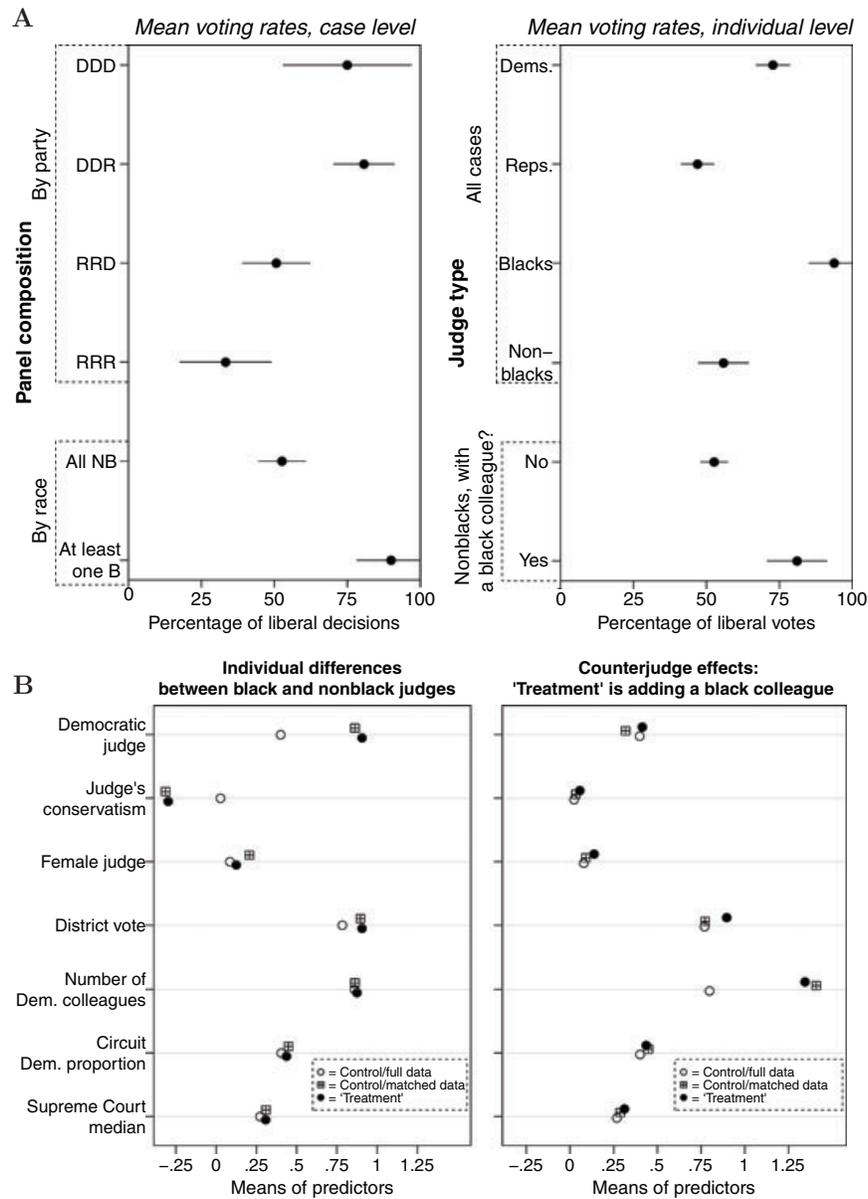
judges were not statistically more likely to vote liberally than white judges, nor did they cause white judges to vote any differently. As noted above, I pool all nonblack judges together in the analyses below.

¹¹ While unpublished (or "nonprecedential") cases have become widely available via Lexis or Westlaw for decisions made in the last decade or so, it is practically infeasible to collect unpublished cases from prior decades.

¹² Many cases involve affirmative action programs that involve both race and gender—all such cases are included in the analysis.

¹³ One possibility is that all minority judges, not just black judges, might be more friendly to affirmative action programs than white judges. The only other minority group with significant numbers in the federal judiciary is Hispanic; public support for affirmative action among Hispanics has been much more uneven than among black Americans (Le and Citrin 2008, 179–81). Still, it is worth examining whether Hispanic judges are more supportive of affirmative action programs, and, if they are, whether they induce panel effects. For all the analyses that appear below, I replicated each, using Hispanic judges instead of African American judges. Hispanic

FIGURE 1 Voting Rates and Balance Statistics



Note: (A): Mean voting rates in affirmative action cases, at the case level and individual-judge level. Horizontal lines depict 95% confidence intervals. The left plot breaks down voting at the case level by partisan panel composition and racial panel composition. Voting is polarized in cases heard by uniform Democratic or Republican panels, with the former nearly 40 percentage points more likely to issue a pro-affirmative action decision. The bottom of the plot shows enormous differences between all-nonblack panels (pooling Democratic and Republican judges together) and a panel with at least one black judge. While the former rule in favor of affirmative action plans about 53% of the time, the latter do so in 90% of cases. The right plot depicts voting at the individual-judge level. Black judges are much more likely to vote liberally than nonblack judges, and nonblack judges are more likely to do so when they sit with a black colleague. (B): Balance in predictors of judicial voting, before and after matching. The open white circles depict the means of the control units in the full, unmatched data; the checkered squares depict means in the control units of the matched data (from the nearest-neighbor matching); and the solid circles depict the means of the treatment units. Because the matching procedures retain all treatment observations, their means are identical across the full and matched data. Balance on circuit indicators is not displayed. While most predictors show good balance in the full data, black judges are much more likely to be Democrats than nonblack judges.

The left plot in Figure 1A also reveals that partisan-based panel effects are less pronounced than the level of polarized voting among unified panels and only occur among Republican-majority panels: Republican-majority panels with one Democrat were more likely to issue liberal decisions compared to unified Republican panels, but there were no meaningful differences among unified Democratic panels and those with two Democrats and a single Republican. In contrast, the bottom of the plot shows enormous differences between all-nonblack panels and a panel with at least one black judge. While the former rule in favor of affirmative action plans about 53% of the time, the latter do so in 90% of cases. This is a striking difference and suggests the existence of large race-based counterjudge effects.

The right plot in Figure 1A switches the perspective to the level of individual votes. The top set of points depicts mean voting rates in all cases, broken down separately by party and race. Unsurprisingly, Democratic judges are much more likely than Republican judges to vote in favor of affirmative action (73% vs. 47%). But the difference between nonblack judges and African American judges is even more pronounced. While nonblack judges cast pro-affirmative and anti-affirmative action votes at about equal rates, *almost every vote* (94%) a black judge casts is in favor of the affirmative action plan under dispute. The bottom of the plot shifts the focus to nonblack judges only, which is the proper unit of analysis when evaluating race-based counterjudge effects. The last two points show mean voting rates by nonblack judges, depending on whether they sit with a black counterjudge or not. Similar to the case-level results, we find that nonblack judges are much more likely to cast pro-affirmative action votes when sitting with a black judge (53% vs. 81%).

While these patterns in the data are highly suggestive, a systematic analysis is necessary in order to assess whether these differences in voting rates persist after accounting for other predictors of judicial behavior. From this point on, I use individual-judge votes as the unit of analysis. Like most studies that examine judicial behavior quantitatively, I use traditional regression analyses to estimate whether black judges vote differently from nonblack judges and whether nonblack judges support affirmative action policies more when sitting with black judges, *ceteris paribus*. In addition, following the example of Boyd, Epstein, and Martin (2010), I use matching methods to implement the potential outcomes approach to causal inference (see, e.g., Morgan and Winship 2007; Rubin 1974). As it turns out, both approaches lead to the same statistical and substantive conclusions; however, as I discuss below, the process of matching itself produces an analytical clarity that would not necessar-

ily emerge from taking the data directly to a regression model.

In the causal inference framework, we are interested in using the logic of counterfactuals to estimate causal effects, or the average effect of adding a “treatment,” compared to a controlled unit. For estimating panel effects, conceptualizing the average treatment effect as a causal effect is straightforward: the treatment effect is the *difference in voting by nonblack judges depending on whether they sit with a black judge or not*. Thus, the treatment is adding a black counterjudge to an otherwise all-nonblack panel, which occurs via a procedure that strongly resembles random assignment of judges to three-judge panels.¹⁴ As Boyd, Epstein, and Martin (2010, 396–97) note, however, conceptualizing the average treatment effect for assessing the difference in voting between nonblack judges and African American judges is less straightforward. Because race (like gender) is not manipulable, it does not make sense to think of being a black judge as a “treatment.” Still, we can use the same causal inference framework to estimate the “average individual difference” between black and nonblack judges.

Before estimating either the average individual differences between nonblack and black judges, or the average treatment effect of adding a black judge to a panel, it is necessary to ensure that observations in which the treatment is present are as similar to observations in which it is not, where similarity is based on the other covariates of interest besides the treatment itself. For both the analysis of individual differences and panel effects, I implemented propensity score matching to create matched datasets for analyses, using both nearest-neighbor matching (with replacement) and optimal matching (with a 1:1 ratio).¹⁵ Both types of matching lead to the same statistical and substantive conclusions.

Turning to predictors of judicial voting beyond race, in each case I coded the direction of the decision of the lower court or federal agency from which the case was

¹⁴ The actual procedures employed for panel assignment vary across circuits and allow for some discretion in panel selection, which mitigates against truly random selection. For instance, judges can trade places on panels in some circuits, and the original judges in a case that requires additional hearings may be selected for such subsequent hearings. In addition, Hall finds that certain circuits have not used random assignment, but instead assigned based on “the date a case was filed or other undisclosed criteria” (2010, 579). While there is no immediate reason to think these criteria would correlate with either votes or the characteristics of the judges or cases that would affect voting, an advantage of matching is that it helps account for possible nonrandom treatment assignment.

¹⁵ All matching procedures were performed using the `MatchIt` package in *R* (Ho et al. 2009). The optimal matching procedure employs Hansen’s `optmatch` package (Hansen and Klopfer 2006).

appealed. To account for the influence of the judicial hierarchy, and the interaction of panel effects and hierarchy (Kastellek 2011a; Kim 2008), I include the proportion of Democrats on the circuit in which the case was decided, as well as Bailey's (2007) estimates of the Supreme Court's ideology for each year, which run from about $-.18$ (most liberal) to $.48$ (most conservative). To measure the ideology of each judge, I assigned each the scores based on the method introduced by Giles, Hettinger, and Peppers (2001), which employs the Common Space score of the appointing president and/or the nominee's home state senator. These scores run from roughly -0.8 (most liberal) to 0.6 (most conservative); I label this predictor *judge's conservatism*. These scores help distinguish between Democratic appointees from northern and southern states; the latter were perhaps less sympathetic to affirmative action programs, particularly in earlier decades. I also include the gender and age of each judge. Finally, I matched on the year of each case and the circuit in which the case was heard.

As the literature on panel effects makes clear, it is important to account for panel composition when modeling voting behavior. Thus, even when estimating the influence of race—including when estimating individual differences between black and nonblack judges—it is important to account for the partisan composition of the panel. I include as a predictor the number of Democratic colleagues on the panel for both the analysis of individual differences and black counterjudge effects. Accounting for race-based panel effects is, of course, by construction what the latter analysis does. But accounting for them in estimating individual differences poses a difficulty. Because there are so few panels with more than one black judge, it is nearly impossible to compare a nonblack judge who sits with a single African American colleague to a black judge who does the same. Thus, while we can control for whether a nonblack judge sits with a black judge, we cannot adequately pose the counterfactual of how a black judge would vote if she sat with another black colleague. At the same time, this difficulty presents an analytic advantage in the estimation of black counterjudge effects. Because majority-black panels are so infrequent, we do not have to distinguish between cases where nonblack judges sit with a single counterjudge and cases where nonblack judges *are* the counterjudge, sitting with two black colleagues. As Kastellek (2011a) notes, the dynamics of panel effects differ greatly depending on whether a judge is a counterjudge or in the panel majority; these differences can complicate estimates of counterjudge effects.

Figure 1B presents descriptive information on some of the predictors used in the matching procedures (and in

subsequent analyses below) and depicts the level of balance in the data before and after matching (fixed effects for circuits are included in the propensity score equations, but not displayed).¹⁶ The left plot shows the results for individual differences; the right plot for the panel effects analysis. The open white circles depict the means of the control units in the full, unmatched data; the checkered squares depict means in the control units of the matched data (from the nearest-neighbor matching); and the solid circles depict the means of the treatment units. Because the matching procedures retain all treatment observations, their means are identical across the full and matched data. Thus, the difference between the circles shows how imbalanced a predictor is, while the difference between the solid circles and the squares shows how well the matching procedure corrects the imbalance.

The figure reveals that, on the whole, the quasi-random assignment of judges to panels does a good job of ensuring balance across treatment and control units for both types of analyses. However, looking at the left plot, one stark difference in the full data can be seen with *Democratic judge*: in the full data, a much higher percentage of treatment units—that is, black judges—are Democrats (90%), compared to control units (40%). Since most African American judges on the Courts of Appeals were appointed by Democratic presidents, this is not surprising.¹⁷ What this means, however, is that when we are asking whether black judges vote differently from nonblack judges, we are effectively asking whether *black judges vote more liberally than nonblack Democratic judges*. The results of the matching procedure illustrate this: 90% of control units in the matched dataset are Democratic judges. Thus, until Republican presidents appoint more black judges to the Courts of Appeals, we cannot answer the general question of whether black judges vote differently from nonblack judges. Turning to the balance statistics for the panel effects analysis, we again observe good balance on most predictors. However, the fact that most black judges are Democrats also has consequences here: when nonblack judges sit with a black colleague, in nearly every instance they are also sitting with at least one Democratic judge. Thus, while there is no imbalance

¹⁶ In performing the actual matching, I did not include every predictor in the propensity score equations. Rather, using trial and error, I selected the model that produced the best balance on all covariates, including ones that did not enter into the equations themselves.

¹⁷ While I use “judge’s conservatism” in the propensity score equations and in the models below, highlighting the lack of balance in *Democratic judge* illustrates this point more effectively. As follows from this fact, the left plot in Figure 1B also shows that the control units in the full data are more conservative than the treatment units; the matching procedure removes this imbalance.

in the partisanship of the nonblack judges, the matching process reveals that when estimating whether black counterjudges cause nonblack judges to vote more liberally in affirmative action cases, we are mainly asking whether they do so compared to the scenario when they sit with one nonblack Democratic colleague.

Results

Table 1 presents the results of several logit models: the dependent variable is whether an individual judge upheld an affirmative action program or not. The first set of three models presents the results of the analysis of individual differences between African American and nonblack judges; the second set presents the results of the counterjudge effects analysis (I discuss the last model below).¹⁸ For each set, the first column presents the results from the full data, and the next two columns present the results from the matched datasets. Thus, within each set the models are run on three different datasets: the full data, the matched data from the nearest-neighbor matching, and the matched data from the optimal matching. Each model includes fixed effects for circuits, except the models evaluating individual differences with the matched datasets (Models 2 and 3), since doing so would consume too many degrees of freedom, relative to the number of observations.

In the models evaluating individual differences, the predictor of interest is *black judge*. For each model, the coefficient on *black judge* is positive and statistically different from zero, indicating that black judges are more likely than nonblack judges to vote in favor of affirmative action programs, *ceteris paribus*. This result holds in the models from the matched datasets, despite their relatively small sample sizes. Thus, we can conclude that there is a statistically significant difference between black and nonblack judges in their individual voting in affirmative action cases.

The top row of graphs in Figure 2A depicts the substantive magnitude of this difference. The left graph shows the average predicted probability of a liberal vote, conditional on the actual values of the predictors for each observation, for nonblack judges and African American

judges. Horizontal lines depict 95% confidence intervals (which are calculated via simulation). In the full data, the average probability of voting liberally if a judge were nonblack is about 56%. In the matched datasets, this rises to about 65%, in large part because the matching process removes the more conservative, Republican judges from the data. If a judge were black, the average probability of supporting affirmative action is about 90% in all three models—meaning in only one out of every 10 votes is a black judge predicted to vote against an affirmative action program, an extremely low rate. The plot to the right depicts, for each model, the average individual difference between black and nonblack judges, which is the difference between the two estimates in the left plot. The estimated difference ranges from 23 to 30 percentage points. And while the confidence intervals are large, it is clear that the difference between how black and nonblack judges vote in affirmative action cases is substantially very large—even when we compare black judges who are similar to nonblack judges on every dimension *except* race.

Turning to the question of whether black judges induce counterjudge effects, the second set of models in Table 1 presents the results of this analysis. The predictor of interest here is *black colleague*; for each model, the coefficient on this variable is positive and statistically different from zero, indicating that nonblack judges are more likely to vote in favor of affirmative action programs when they sit with a black colleague, *ceteris paribus*. Again, the results hold in the matched datasets as well as in the full data.¹⁹

Returning to Figure 2A, the bottom row of graphs depicts the substantive magnitude of this effect. The left plot depicts the average probability that nonblack judges support an affirmative action plan, depending on whether they sit with an African American colleague or not. For the full and matched datasets, the results are nearly identical. Nonblack judges are predicted to vote liberally in about 50% of cases when they sit with two nonblack colleagues. When a black judge joins the panel, the predicted probability rises to about 80%. This means that when a nonblack judge sits with a black judge, she is nearly as likely to vote liberally as the black judge himself.

The bottom-right plot in Figure 2A depicts the difference between the two predicted probabilities for each model, which is the average treatment effect of adding a black judge to an otherwise all-nonblack panel.

¹⁸ In three of the models in Table 1—(2), (5), and (6)—complete or quasi-separation occurs for at least one of the predictors. Rather than dropping these variables from the models, which throws out valuable information and potentially biases results, I use the method introduced by Gelman, Jakulin, Pittau, and Sung Su (2008), which places a weakly informative prior distribution on each coefficient to overcome the problem of separation. These models are run using the `bayesglm` command from the *ARM* package in *R*.

¹⁹ As a robustness check, I performed a sensitivity analysis that employs Rosenbaum bounds to test for whether omitted variable bias could potentially confound the results. This analysis confirms the robustness of the results from the matched data. See the online appendix for full details.

Table 1 Logit Models of the Probability of Upholding an Affirmative Action Program

	Individual Differences			Counter-judge Effects			Nonblack Republicans
	Full data (1)	Nearest-neighbor matching (2)	Optimal matching (3)	Full data (4)	Nearest-neighbor matching (5)	Optimal matching (6)	Full data (7)
Intercept	-1.18 (1.02)	-3.14 (3.67)	-8.68* (4.39)	-1.36 (0.99)	-0.66 (2.63)	-1.17 (2.70)	-1.01 (1.45)
Black judge	2.18* (0.77)	1.95* (0.82)	3.88* (1.23)				
Black colleague				1.46* (0.40)	1.75* (0.52)	1.90* (0.49)	
Judge's conservatism (GHP score)	-1.91* (0.33)	-2.19 (1.48)	-3.45* (1.57)	-1.91* (0.33)	-0.10 (0.75)	-0.49 (0.71)	-1.73* (0.81)
Female judge	0.52 (0.40)	1.61 (1.42)	2.29 (1.28)	0.39 (0.40)	-0.49 (0.76)	-0.30 (0.75)	0.84 (0.75)
Age	0.01 (0.01)	0.01 (0.05)	0.08 (0.05)	0.01 (0.01)	0.01 (0.03)	0.02 (0.03)	0.02 (0.02)
Lower court vote	0.83* (0.27)	-0.36 (1.14)	1.84 (1.28)	0.72* (0.26)	-0.14 (0.73)	-0.39 (0.84)	0.60 (0.38)
Number Dem. colleagues	0.78* (0.17)	0.69 (0.63)	-0.63 (0.65)	0.63* (0.17)	0.56 (0.46)	0.09 (0.44)	
Dem. proportion on circuit	1.29 (0.98)	5.25 (3.61)	2.77 (4.10)	1.42 (0.97)	2.78 (2.22)	3.23 (2.19)	0.27 (1.39)
Supreme Court's conservatism	0.77 (1.05)	-1.15 (3.20)	-0.16 (5.99)	0.33 (1.06)	-4.22 (3.00)	-3.60 (3.18)	-1.25 (1.46)
One non-black Dem. colleague							0.19 (0.32)
Two non-black Dem. colleagues							1.94* (0.55)
One black Dem. colleague							3.11* (0.64)
<i>N</i>	546	61	64	514	102	116	306
% in modal category	58.1	81.2	75.0	55.8	71.6	65.5	53.2
% correctly classified	73.6	88.5	89.1	73.3	82.3	78.4	74.0
% reduction in error	37.1	36.3	56.2	39.7	37.9	37.5	44.4

Note: Standard errors in parentheses. *indicates significance at $p < 0.05$. The following models use the method advanced in Gelman et al. (2008) in order to avoid separation or quasi-separation in the logit estimations: (2), (5) and (6). In the matched analyses, the number of observations increases relative to the individual analysis due to the fact that non-black judges are being matched with other non-black judges, whereas in the individual analysis the small number of black judges are being matched with non-black judges. Each model except Models (2) and (3) includes fixed effects for circuits.

Substantively, adding a black counterjudge increases the probability that a nonblack judge will vote in favor of affirmative action by about 25 to 30 percentage points. While the confidence intervals are large, as a point of comparison, this estimate is larger than the effect that Boyd, Epstein, and Martin (2010) found from adding a female judge to an otherwise all-male panel (about 13%). In ad-

dition, this estimate is comparable to the largest partisan counterjudge effect uncovered in Kastlelec (2011a), which only occurs when both the full circuit and the Supreme Court are aligned in favor of the counterjudge and against the panel majority. Even more dramatically, returning to the data for a moment, of the 30 cases in which at least one black judge participated, 27 (90%) resulted in a decision

in favor of affirmative action. And in only *one case* did a black judge write a dissent.

Suggestive evidence on the mechanism of judicial influence. While we cannot definitely pin down the mechanism underlying the black counterjudge effect, further analysis can shed some light on this question. Consider the votes of nonblack Republicans, the judges who are least likely to uphold affirmative action programs. Figure 2B depicts the percentage of liberal votes nonblack Republicans cast, depending on the partisan and racial composition of the panel. As a baseline, when nonblack Republicans sit with two other Republicans (neither of whom is ever black in the data), they vote liberally about 36% of the time. Adding a single *nonblack* Democratic counterjudge to the panel (creating an all-nonblack panel of two Republicans and one Democrat) makes basically no difference—nonblack Republicans still uphold affirmative action plans at a similar rate. What happens if another nonblack Democratic judge joins the panel, meaning that the nonblack Republican judge is now in the partisan minority of the panel and is himself a counterjudge? Unsurprisingly, the percentage of liberal votes by nonblack Republicans increases to about 67%, a sizable difference.

Next, consider the scenario where we add a *single* black colleague (who are always Democrats in the data) to the panel.²⁰ When a nonblack Republican sits with one black judge and one nonblack Democratic judge, the nonblack Republican votes liberally 80% of the time. Similarly, when a nonblack Republican sits with one black judge and a nonblack Republican judge, the nonblack Republican votes liberally 86% of the time. Thus, *adding a single black judge to the panel* leads nonblack Republicans to vote liberally in four out of every five cases—and this result is not driven at all by the partisanship of the third judge. This rate is substantially higher than the 67% seen when a nonblack Republican sits with *two* nonblack Democrats. Thus, randomly assigning *one* black Democratic judge is more likely to result in a nonblack Republican voting to uphold an affirmative action plan than randomly assigning *two* nonblack Democratic judges, even though the nonblack Republican judge is in the partisan majority in the former case and the partisan minority in the latter.²¹

²⁰ For clarity, I dropped the single case that featured two black judges on the panel.

²¹ The results of a regression in which votes are regressed on these categories (with two nonblack Republican colleagues as the base category) are presented as Model (7) in Table 1. Due to the small sample sizes and the fact that there is little difference in voting in

The presence or deliberation mechanisms could explain this pattern. However, if nonblack Republicans were simply being influenced by liberal votes of judges in the voting minority, it would be hard to explain why they are not influenced *at all* (on average) by the votes of nonblack Democratic judges. While it is true that black Democrats vote more liberally than nonblack Democrats, that difference cannot seem to explain the huge discrepancy seen in Figure 2B. Thus, given how salient race is in these cases, it seems likely that the voting mechanism is not driving the counterjudge effects; rather, black judges are leading their nonblack colleagues to change their votes either by their presence on the panel, or by presenting information or arguments that their nonblack colleagues would not receive otherwise.

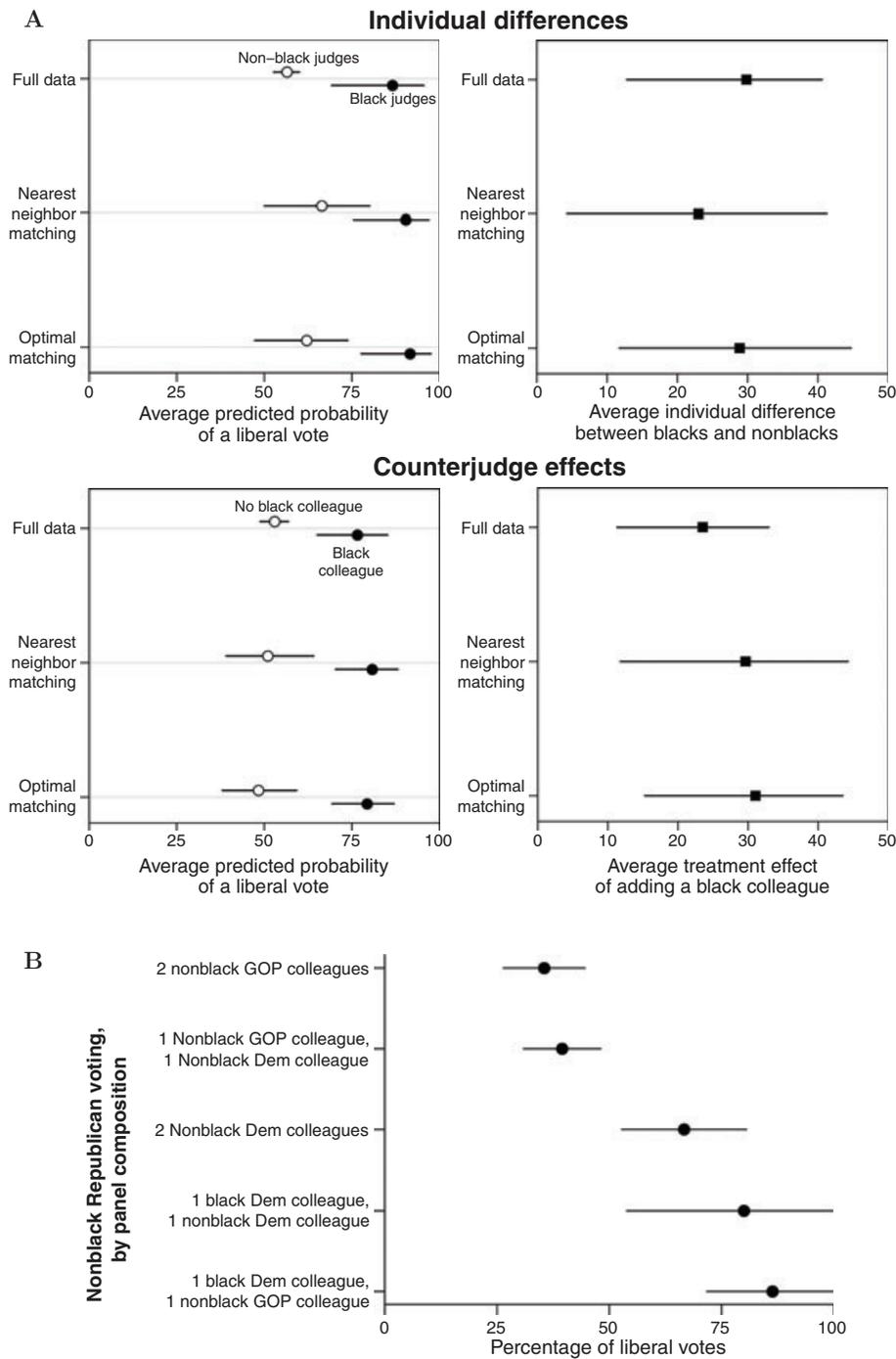
Discussion and Conclusion

This article has demonstrated that the presence of black judges has consequences far beyond what their small numbers on the Courts of Appeals would suggest. The individual differences analysis shows that black judges are much more likely than nonblack judges to support affirmative action plans and do so at a rate of about 90%. In turn, nonblack judges who sit with a black colleague uphold affirmative action plans about 80% of the time. Given majority rule on three-judge panels, these two results in tandem mean *the random assignment of a black judge to a three-judge panel in affirmative action cases nearly ensures that the panel will issue a liberal decision*. Preliminary evidence on the possible mechanisms underlying this causal effect suggests that it cannot be attributed to nonblack judges responding to the votes of black judges, but rather is due either to a presence or deliberation effect.

This study provides stark evidence that judicial diversity seems to have the effects many of its proponents intend. At the same time, the finding of large racial counterjudge effects leads to a somewhat thorny question: is the fact that the random assignment of a black judge to a three-judge panel potentially sways the outcome of a case a good thing or a bad thing? Unfortunately, politically salient cases that reach the Courts of Appeals, such as affirmation action cases, do not lend themselves to

the two “1 black Dem colleague” categories depicted in Figure 2B, I pooled the categories together into the variable labeled “One black Dem. colleague” in Table 1. While the coefficient on *One nonblack Dem. colleague* is effectively zero, the coefficients on both *One black Dem. colleague* and *Two nonblack Dem. colleagues* are positive and statistically different from zero. However, the coefficient on the former is much larger and is statistically larger than the coefficient on the latter ($p < .07$, two-tailed.)

FIGURE 2 The Substantive Significance of Counterjudge Effects



Note: (A): Predicted probabilities, average individual differences, and counterjudge effects in affirmative action case. The top graphs present the results from the analysis of individual differences, while the bottom graphs present the results from the counterjudge effects analysis. For both, the left graph shows the average predicted probability of a liberal vote, conditional on the levels of the predictors for each observation. The right graphs depict the average individual difference between black and nonblack judges and the average treatment effect of adding a black judge to an otherwise all-nonblack panel, respectively. Horizontal lines depict 95% confidence intervals. (B): The votes of nonblack Republicans, by partisan and racial panel composition. The figure presents the percentage of liberal votes by nonblack Republicans, depending on the partisan and racial composition of the panel. Adding a single black Democratic colleague to the panel leads to a larger increase in liberal voting than adding two nonblack Democratic colleagues.

measures of “correctness” (Cox and Miles 2008b, 52). Thus, we cannot say whether panels that include a black judge are more likely to reach the “right” answer. We can, however, conclude that the presence of an African American judge does change the underlying dynamics of panel decision making; as Cox and Miles (2008b, 53) note, this finding is consistent with arguments that diversity enhances the legitimacy of deliberations by increasing the range of perspectives considered by a decision-making body.

We can draw clearer conclusions with respect to the article’s implications for evaluating representation on federal courts. Returning to the question of whether non-blacks alone can provide representation for minorities, all-nonblack panels have voted for affirmative action in roughly 50% of cases. Whether this constitutes effective representation requires a thorough evaluation of the mapping between public opinion and judicial decisions. Nevertheless, it is clear that mass opinion among black Americans is much more favorable to affirmative action than the opinion of nonblack Americans, and the existence of black counterjudge effects means that more affirmative action programs survive judicial review than would otherwise. In addition, their existence suggests that the counterjudge effects that flow from descriptive representation provide an increase in substantive representation over a *baseline* offered by nonblack judges, rather than the complete absence of it. This possibility has not been explored in studies of characteristic-based panel effects to date; evaluating both these baselines and subsequent increases from gender- or race-based counterjudges effects offers scholars a way to understand representation on federal courts, and this avenue might be pursued profitably in future studies of diversity on the federal courts.

A conclusion one might draw from the existence of this baseline of substantive representation is that descriptive representation is not really necessary—or, alternatively, that it might lead to “overrepresentation” of minority interests. Here it is important to recognize that in many areas of the law the baseline will be much lower, simply because all judges are more predisposed to vote liberally in affirmative action cases compared to other issue areas.²² Given that women and (especially) minority judges have tended to be more liberal than their counterparts, the existence of characteristic-based counterjudge effects will have a much more significant impact in translating descriptive representation into substantive representation in other areas of the law, such as employment

discrimination, where most cases tend to be decided conservatively.

Similarly, while we might expect race to be most salient in affirmative action cases—and thus in this area we might expect to see the largest counterjudge effect—this does not lead to the conclusion that the substantive importance of diversity on the federal bench is narrow or constrained. First, we now have evidence of large gender- and race-based counterjudge effects in important and active areas of the law—affirmative action cases have wide-reaching consequences, while employment discrimination cases are among the most common types of lawsuits filed in federal courts each year. Second, there is little reason to think we would not see similar effects on other dimensions of “surface-level diversity,” such as age or disabilities (Harrison, Price, and Bell 1998; Phillips and Loyd 2006). Such effects would likely permeate across a wide range of cases in the federal judiciary.

In addition, moving beyond pure surface-level diversity, there also exists on the federal bench what organizational scientists call “deep-level” diversity, which includes dimensions such as cognition, tools, and ability. Such diversity has been shown to influence and improve group-level performance (Page 2007). While these qualities would be difficult to measure in observational data, it seems likely that diversity on these dimensions on multimember courts translates into differences in judicial decisions and legal outcomes. Indeed, the two most recent Supreme Court vacancies—which resulted in an increase in surface-level diversity in the appointment of two female justices—have seen calls to increase the deep-level diversity on the Court by selecting justices who are not sitting federal judges with Ivy League law degrees and based in the Northeast (Sherman 2010). As Page (2009) has argued: “A court of nine diverse people is wiser than a court of identical minds because members of a diverse court bring different ideas to bear and productively challenge one another’s interpretations of the law.” Thus, the effects of diversity on the Courts of Appeals likely extend far beyond what we can observe and immediately measure.

Incorporating these various dimensions of diversity could lead to broader understandings of substantive representation on courts and the many ways in which all facets of minority views can find themselves represented on multimember courts. For example, one limitation of existing studies is they have only examined counterjudge effects with respect to judicial votes and not to legal doctrine. While case dispositions are most important to the litigants in a given case, the legal doctrine that emerges from the decision of a three-judge panel can be critical to the development of the law in a particular area, given the

²² In the 24 issue areas studied by Sunstein et al. (2006, 151), affirmative action is among the four areas with the highest rates of liberal voting.

policymaking role of the Courts of Appeals and the fact that the vast majority of their decisions are not reviewed by the Supreme Court. While scholars have recently developed tools to measure the output of judicial opinions on the Supreme Court (Clark and Lauderdale 2010), such measures have not been applied to the Courts of Appeals. This extension could lead to the evaluation of whether counterjudge effects lead to material differences in legal doctrines, which would mean that the boost in substantive representation would occur both in instant cases and beyond, due to the development of and reliance on precedent in a common law system.

Finally, with respect to theoretical models of judicial politics in more generality, the results presented here (and in the panel effects literature more broadly) illustrate the limitation of straightforwardly applying the median voter theorem to judicial decision making. Recent work on the Supreme Court has demonstrated conclusively that judicial institutions such as opinion writing and the distinction between dispositions and rules mean that the median voter theorem does not paint the right picture of preference aggregation among the justices (Carrubba et al. 2012; Lax and Cameron 2007). Similarly, the counterjudge effects documented in this article arise from the failure of a simple median voter model on three-judge panels, as the ability of a single black counterjudge to persuade her colleagues moves judicial outcomes away from the implied median. As students of the courts continue to move beyond the importation of legislative models, it seems profitable to ask how other judicial institutions complicate the aggregation of preferences on multimember courts.

References

- Antonio, Anthony L., Mitchell J. Chang, Kenji Hakuta, David A. Kenny, Shana Levin, and Jeffrey F. Milem. 2004. "Effects of Racial Diversity on Complex Thinking in College Students." *Psychological Science* 15(8): 507–10.
- Ashenfelter, Orley, Theodore Eisenberg, and Stewart J. Schwab. 1995. "Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes." *Journal of Legal Studies* 24(2): 257–81.
- Bailey, Michael A. 2007. "Comparable Preference Estimates across Time and Institutions for the Court, Congress and Presidency." *American Journal of Political Science* 51(3): 433–48.
- Bhagwai, Ashutosh. 2001–2. "Affirmative Action and Compelling Interests: Equal Protection Jurisprudence at the Crossroads." *University of Pennsylvania Journal of Constitutional Law* 4(2): 260–80.
- Bonneau, Chris W., and Heather Marie Rice. 2009. "Impartial Judges? Race, Institutional Context and U.S. State Supreme Courts." *State Politics & Policy Quarterly* 9: 381–403.
- Boyd, Christina L., Lee Epstein, and Andrew Martin. 2010. "Untangling the Causal Effects of Sex on Judging." *American Journal of Political Science* 54(2): 389–411.
- Cameron, Charles M., and Craig Cummings. 2003. "Diversity and Judicial Decision Making: Evidence from Affirmative Action Cases in the Federal Courts of Appeals, 1971-1999." Presented at the annual meeting of the Midwest Political Science Association.
- Carrubba, Cliff, Barry Friedman, Andrew D. Martin, and Georg Vanberg. 2012. "Does the Median Justice Control the Content of Supreme Court Opinions?" *American Journal of Political Science* 56(2): 400–412.
- Chew, Pat K., and Robert E. Kelly. 2006. "Unwrapping Racial Harassment Law." *Berkeley Journal of Employment and Labor Law* 27: 49–110.
- Clark, Tom S., and Benjamin Lauderdale. 2010. "Locating Supreme Court Opinions in Doctrine Space." *American Journal of Political Science* 54(4): 871–90.
- Cox, Adam B., and Thomas J. Miles. 2008a. "Judicial Ideology and the Transformation of Voting Rights Jurisprudence." *University of Chicago Law Review* 75: 1493–1539.
- Cox, Adam B., and Thomas J. Miles. 2008b. "Judging the Voting Rights Act." *Columbia Law Review* 108(1): 1–54.
- Cross, Frank B., and Emerson H. Tiller. 1998. "Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals." *Yale Law Journal* 107: 2155–76.
- Edwards, Harry T. 2002. "Race and the Judiciary." *Yale Law & Policy Review* 20: 325–30.
- Farhang, Sean, and Gregory Wawro. 2004. "Institutional Dynamics on the U.S. Court of Appeals: Minority Representation under Panel Decision Making." *Journal of Law, Economics, and Organization* 20: 299–330.
- Farhang, Sean, and Gregory Wawro. 2010. "Indirect Influences of Gender on the U.S. Courts of Appeals: Evidence from Sexual Harassment Law." U.C. Berkeley working paper.
- Fischman, Joshua B. 2010. "Estimating Preferences of Appellate Judges: A Model of Consensus Voting." University of Virginia working paper.
- Fischman, Joshua B. 2011. "Interpreting Circuit Court Voting Patterns: A 'Social Interactions' Framework." University of Virginia working paper.
- Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau, and Yung Su. 2008. "A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models." *Annals of Applied Statistics* 2(4): 1360–83.
- 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: 623–41.
- Hall, Matthew. 2010. "Randomness Reconsidered: Modeling Random Judicial Assignment in the U.S. Courts of Appeals." *Journal of Empirical Legal Studies* 7(3): 574–89.
- Hansen, Ben B., and Stephanie Olsen Klopfer. 2006. "Optimal Full Matching and Related Designs Via Network Flows." *Journal of Computational & Graphical Statistics* 15: 609–27.
- Harrison, David A., Kenneth H. Price, and Myrtle P. Bell. 1998. "Beyond Relational Demography: Time and the Effects of Surface- and Deep-Level Diversity on Work Group Cohesion." *Academy of Management Journal* 41(1): 96–107.

- History of the Federal Judiciary*. 2011. Website of the Federal Judiciary Center. Washington, DC. <http://www.uscourts.gov/judgesAndJudgeships/BiographicalDirectoryOfJudges.aspx>.
- Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2009. "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." Harvard University Technical Report.
- Hurwitz, Mark S., and Drew Noble Lanier. 2008. "Diversity in State and Federal Appellate Courts: Change and Continuity across 20 Years." *Justice System Journal* 29(1): 47–70.
- Kastellec, Jonathan P. 2011a. "Hierarchical and Collegial Politics on the U.S. Courts of Appeals." *Journal of Politics* 73(2): 345–61.
- Kastellec, Jonathan P. 2011b. "Panel Composition and Voting on the U.S. Courts of Appeals Over Time." *Political Research Quarterly* 62(2): 377–91.
- Kellough, J. Edward. 2006. *Understanding Affirmative Action: Politics, Discrimination, and the Search for Justice*. Washington, DC: Georgetown University Press.
- Key, V. O., Jr. 1949. *Southern Politics in State and Nation*. New York: Alfred Knopf.
- Kim, Pauline T. 2008. "Deliberation and Strategy on the United States Courts of Appeals: An Empirical Exploration of Panel Effects." *University of Pennsylvania Law Review* 157: 1319–81.
- Lax, Jeffrey R., and Charles M. Cameron. 2007. "Bargaining and Opinion Assignment on the U.S. Supreme Court." *Journal of Law, Economics, and Organization* 23(2): 276–302.
- Le, Loan, and Jack Citrin. 2008. "Affirmative Action." In *Public Opinion and Constitutional Controversy*, ed. Jack Citrin, Nathaniel Persily, and Patrick J. Egan. Oxford: Oxford University Press, 162–83.
- Liptak, Adam. 2009. "The Waves Minority Judges Always Make." *New York Times*. May 31. A1.
- Mannix, Elizabeth, and Margaret A. Neale. 2005. "What Differences Make a Difference?" *Psychological Science in the Public Interest* 6: 31–55.
- Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60: 531–42.
- Morgan, Stephen L., and Christopher Winship. 2007. *Counterfactuals and Causal Inference*. Cambridge: Cambridge University Press.
- Myrdal, Gunnar. 1944. *An American Dilemma: The Negro Problem and Modern Democracy*. New York: Random House.
- Page, Scott E. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. Princeton, NJ: Princeton University Press.
- Page, Scott E. 2009. "Sotomayor's Diversity and the Supremes." The Huffington Post. Available at <http://www.huffingtonpost.com/scott-page/sotomayors-diversity-and-b.216777.html> (accessed October 17, 2011).
- Peresie, Jennifer L. 2005. "Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts." *Yale Law Journal* 114(7): 1759–90.
- Phillips, Katherine W., and Denise Lewin Loyd. 2006. "When Surface and Deep-Level Diversity Collide: The Effects on Dissenting Groups Members." *Organizational Behavior and Human Decision Processes* 99: 143–60.
- Pitkin, Hanna F. 1967. *The Concept of Representation*. Berkeley: University of California Press.
- Powell, Lisa M., John A. Tauras, and Hana Ross. 2005. "The Importance of Peer Effects, Cigarette Prices and Tobacco Control Policies for Youth Smoking Behavior." *Journal of Health Economics* 24(5): 950–68.
- Revesz, Richard L. 1997. "Environmental Regulation, Ideology and the D.C. Circuit." *Virginia Law Review* 83: 1717–72.
- Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66: 688–701.
- Scherer, Nancy, and Brent Curry. 2010. "Does Descriptive Race Representation Enhance Judicial Legitimacy? The Case of the U.S. Courts." *Journal of Politics* 72(1): 90–104.
- Sherman, Mark. 2010. "Is Supreme Court in Need of Regional Diversity?" Associated Press. Available at <http://www.msnbc.msn.com/id/36890869/ns/politics/t/supreme-court-need-regional-diversity/> (accessed October 17, 2011).
- Sigelman, Lee, and Susan Welch. 1991. *Black Americans' Views of Racial Inequality: The Dream Deferred*. Cambridge: Cambridge University Press.
- Sommers, Samuel R. 2006. "On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations." *Journal of Personality and Social Psychology* 90(4): 597–612.
- Sotomayor, Sonia. 2002. "A Latina Judge's Voice." *Berkeley La Raza Law Journal* 13: 87–93.
- Steeh, Charlotte, and Maria Krysan. 1996. "Affirmative Action and the Public, 1970-1995." *Public Opinion Quarterly* 60(1): 128–58.
- Sunstein, Cass R., David Schkade, Lisa M. Ellman, and Andres Sawicki. 2006. *Are Judges Political? An Empirical Analysis of the Federal Judiciary*. Washington, DC: Brookings Institution Press.
- Swain, Carol. 1993. *Black Faces, Black Interests: The Representation of African Americans in Congress*. Cambridge, MA: Harvard University Press.
- Washington, Linn, ed. 1994. *Black Judges on Justice*. New York: The New Press.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

1. Case selection and coding procedures
2. Robustness check: sensitivity analysis with Rosenbaum bounds
3. References

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.