Scholars have long sought to resolve whether and to what degree diversity influences the outputs of political institutions, including courts. When it comes to the judiciary, diverse judges may greatly affect outcomes. Despite this potential, no consensus exists for whether judicial diversity affects behavior in trial courts—that is, the stage where the vast majority of litigants interact with the judicial branch. This article first addresses the research design limitations in previous trial court diversity studies. It then notes that the results of this study indicate that a trial judge’s gender and race have very large effects on his or her decision making. These results have important implications for how we view diversity throughout the judiciary.

**Representation on the Courts?**

The Effects of Trial Judges’ Sex and Race

Christina L. Boyd

Representation and its connection to diversity among political elites is as important today as ever before. Whether discussing legislators, bureaucrats, judges, or other elites, how well these actors represent the citizenry’s diverse characteristics and interests—both in terms of composition and outputs—has significant implications for our political institutions’ function, membership, and legitimacy. To date, the results of diversity-in-politics research indicate, for example, that black and female state legislative representatives pursue representative policies (Bratton and Haynie 1999), female federal administrative agency bureaucrats support more female-friendly policies than their male counterparts (Dolan 2000), and that female and black appellate court judges decide diversity-specific cases differently from their white, male colleagues and affect their colleagues’ decision making (Boyd, Epstein, and Martin 2010; Kastellec 2013).

Although these insights are noteworthy, much remains unknown about the presence and degree of diversity representation within many of our political institutions today. To see this in one particular political institution, one need look no further than the behavior and words of the Obama Administration. Between taking office in 2009 and July 2016, President Obama successfully appointed more than ninety minority and 100 female federal trial court judges. Appointed to life tenure positions under Article III of the Constitution, these newly selected judges make up nearly 62 percent of Obama’s trial court selections, by far the highest percentage of diverse judges appointed by a U.S. president. Although Obama’s pride in diversifying the federal trial courts is clear, what is less clear is what impact, if any, this large number of diverse appointments is likely to have on our nation’s trial courts and the decisions made therein.

The Obama Administration argues that these new judges serve as descriptive representatives of diverse communities in the United States and that their presence encourages the public to view the courts as legitimate. As Senior Counsel to the President Chris Kang (2013) notes,
How Diversity May Affect Trial Judging

Scholars and policy makers have long sought to discover whether diverse judges, be they females or racial minorities, behave and make decisions differently than more descriptively traditional male and white judges. And with good reason. Judges, and particularly trial judges, are well positioned to affect their assigned cases and the way that they progress. This has tremendous implications in trial courts for outcomes, settlements, costs, appeals, the distribution of resources after a case, and even the decision of litigants to file their cases and seek adjudicated remedies at all. In short, if diverse trial judges behave differently from their colleagues, we should expect substantive differences in the outputs of the judiciary.

Diversity Theories and Trial Judging

With so much interest in diverse judges, numerous prominent accounts theorizing why and when female and racial minority judges will behave differently from their traditional colleagues on the bench have emerged. These include different voice, informational, representational, balance, and organizational theories. A brief discussion of each is provided below.

Established by Gilligan (1982), the different voice theory applies primarily to males and females and rests on the idea that the two sexes think, communicate, and view the world differently from one another. Gilligan argues that the male voice is distinctly masculine and committed to concepts like logic and justice. By contrast, the female voice is much more devoted to obligations, relationships, and problem solving through personal communication. Applied to judging, the different voice theory expects to see a “feminine perspective” that extends to females’ decisions as judges across “all aspects of society, whether or not they affect men and women differently” (Sherry 1986, 160).

Informational theory argues that female and black judges bring a unique knowledge base and expertise to the bench in key areas such as traditional sex and race employment discrimination issues, respectively. This expertise and information is “credible and persuasive” (Boyd, Epstein, and Martin 2010, 392) and yields from a combination of their professional background and their distinctive group-influenced experiences and challenges.
Still others believe that female and minority decision makers serve as high-profile, substantive representatives of others of their class working to advance their group’s interests through their decisions (e.g., Farhang and Wawro 2004; Kastellec 2013; Pitkin 1967). For female judges, this means that they “will seize decision-making opportunities to liberate other women” (Cook 1981, 216). For black judges, a strong race-based social identity will likely encourage strong in-group closeness and racial consciousness (McClain et al. 2009). In turn, this should affect black judge decision making on racially salient topics particularly because, as federal judge Edwards (2002, 328) notes, “it is safe to assume that a disproportionate number of blacks grow up with a heightened awareness of the problems that pertain to” racial discrimination “areas of the law.” ²

Importantly, while all three of these accounts—different voice, informational, and representational—sometimes vary in their empirical implications at the margins (see Boyd, Epstein, and Martin 2010), they converge when it comes to individual judge behavior in issue areas closely connected to the judge’s diversity characteristic(s). Indeed, all three anticipate that, with all other things being equal, female judges will be more likely to support sex discrimination claimants than male judges and black judges will be more likely to favor race discrimination or affirmative action plaintiffs than white judges.

Balance theory goes even further than these theories. It expects that for minority decision makers, their group’s historical experience with discrimination is likely to make them more sympathetic to others facing discrimination, even if that discrimination is not based on race (Bratton and Haynie 1999; Heldner 1958). As Kulik, Perry, and Pepper (2003, 71) note, since “people of color [are] at greater risk of harassment . . . [this] increased vulnerability might make members of minority groups especially sensitive to the harassment of others.” Applied to diversity judging, Crowe (1999, 115) argues that while “female judges [may] not generalize [their identification with] sex discrimination plaintiffs to other types of discrimination plaintiffs,” black judges may have a broader desire to support discrimination plaintiffs.

Of course, not all theoretical accounts expect that diverse judges should be expected to behave distinctly from white and male judges on the bench. Under organizational theory, some scholars argue that all judges—white, minority, male, and female—undergo the same professionalization and training before acquiring their judicial employment (Kritzer and Uhlman 1977). These judges’ decisions must rely on the same laws and norms, something that may be particularly constraining in trial courts. In effect, the “powerful ‘organizational’ influences of selection and socialization to the judicial role” along with “adherence to prevailing norms, practices, and precedents” can offset any pre-officeholding biological, socialization, psychological, or experiential differences of the judges (Steffensmeier and Britt 2001, 752–53). For adherents to this organizational theory, judicial diversity characteristics like sex and race are thus expected to have no systematic effects on trial judge behavior. This is likely to sit well with those, like the Obama White House, who have argued that the addition of diverse judges to the federal judiciary has nothing to do with outcomes and everything to do with descriptive and symbolic representation—the belief that “democratic institutions in heterogeneous societies should reflect the make-up of society” (Cameron and Cummings 2003, 28).

One important assumption of the above discussion of diversity judging theory is that judges hold the discretion and opportunities necessary to behave in ways to implement their preferred outcomes. In other words, female and black judges, should they be so inclined, have enough decision-making flexibility to allow them to make decisions that favor sex and race discrimination plaintiffs. For appellate court judging, there is little question that this is present (P. M. Collins and Martinek 2011; Hettinger, Lindquist, and Martinek 2006; Klein 2002; Scott 2006; Wahlbeck 1997). However, when it comes to trial court judges, much debate and discussion exists on this subject. Many trial court scholars hold firmly to the belief that trial judges are primarily administrators, appliers of law, and organizational bureaucrats, spending their time in rote activities like holding hearings and supervising jury trials. Scholars often fail to find evidence that trial judges’ identities—including things like their ideological preferences—affect outcomes (e.g., Dumas and Haynie 2012; Keele et al. 2009). If this view of trial judging is correct and generalizable, when applied to diversity judging, it could very well mean that even if diverse trial judges wish to favor diverse plaintiffs, their hands are often tied. Empirically, it would be nearly impossible to disentangle these trial judges from those
who behave according to organizational theory's expectations (i.e., who, even unconstrained, do not behave differently from their traditional colleagues).

The more likely scenario, however, is that trial judges hold numerous discretion-filled behavioral opportunities—that is, decisions rivaling those made by appellate court brethren in most cases—but these opportunities do not come to pass as often as they might. With careful isolation of district judging roles, it should therefore be possible to focus solely on the parts of trial judging that leave room for judges’ identities and preferences to matter. Which trial judging jobs provide the most discretionary behavior opportunities? The most obvious examples are decisions made during bench trials and on dispositive motions. In both instances, the judge serves as the fact finder and the law applier and also holds real potential to interpret the law. By comparison, while serving as a case manager (e.g., holding status and settlement conferences), supervising jury trials, or deciding whether to transfer a case, trial judges hold less decision-making discretion that could affect outcomes toward their preferred direction (e.g., Williams and George 2013). This is not to say, of course, that trial judges have no discretion in these parts of their jobs, but the important point is that the discretion that is present is, at best, indirectly linked to the case's outcome and being able to drive that outcome toward favoring one party over another.

The Empirical Evidence So Far

Moving from diversity theory to empirics, which side of the theoretical divide receives the bulk of the support? When it comes to trial courts, little consensus exists. For judge race, nearly half of the projects’ results indicate no statistically significant impact from a trial judge's race on his or her decision-making behavior (e.g., Ashenfelter, Eisenberg, and Schwab 1995; Kulik, Perry, and Pepper 2003; Segal 2000; Spohn 1990; Walker and Barrow 1985). Another set of these studies find that minority trial judges are more likely to support liberal positions than their white colleagues (e.g., Chew and Kelley 2009; Schanzenbach 2005; Sisk, Heise, and Morriss 1998; Weinberg and Nielsen 2012; Welch, Combs, and Gruhl 1988). Only a small handful of judge-race trial judging studies find that minority judges behave contrary to all expectations by regularly supporting a conservative position more frequently than white judges, with, for example, both Steffensmeier and Britt (2001) and Uhlman (1978) finding that black judges have harsher sentencing behavior than white judges.

For judge sex, the research findings on trial judge behavior are perhaps even more convoluted. The majority of these studies find no evidence that female trial judges behave differently than male judges across a range of issue areas, from sentencing to sexual harassment claims and employment discrimination disputes (e.g., Ashenfelter, Eisenberg, and Schwab 1995; Chew and Kelley 2009; Gruhl, Spohn, and Welch 1981; Kulik, Perry, and Pepper 2003; Sisk, Heise, and Morriss 1998; Weinberg and Nielsen 2012). A smaller, but nontrivial, set of studies have found that female judges tend to behave more conservatively than their male counterparts in the trial courts. This includes being less likely to support women's issues like personal liberties, maternity rights, reproductive freedom, custody, and equal employment than male judges (Segal 2000; Walker and Barrow 1985) and being more likely to harshly sentence criminal defendants (P. M. Collins, Manning, and Carp 2010; Spohn 1990). Another small set of studies, however, have found that female judges behave more liberally than their male trial court colleagues (e.g., Johnson 2014; Manning 2004).

Research Design and Data

Looking for Diversity Effects in Trial Courts

As the above discussion indicates, when it comes to what the current literature has to say about diversity and judicial decision making in trial courts, uncertainty reigns. Before embarking on yet another study of diversity in trial judging and simply adding to the confusion, it is necessary to address the unique considerations for diversity trial court judging that present empirical challenges and may share responsibility for the state of the literature. The three key factors are as follows: the unit of analysis for judge decision making, the large number of frivolous trial court cases, and narrow diversity-centric issue area data. For a study on the subject to yield reliable results, these features must be taken into account in the research design.

First, trial courts are the courts of first instance, where litigation begins, develops a factual record, provides
the initial application of law to those facts, and permits the regular and often years’ long interaction between trial judges and case participants. As a product of this design, and quite distinctly from appellate courts, trial court judges have numerous opportunities to make important decisions throughout a case. Of note, most civil cases develop through motions practice, with the judge’s decision to grant or deny the motion having critical implications for how the case will proceed (Hoffman, Izenman, and Lidicker 2007; Kim et al. 2009). It is these decisions on significant motions throughout a case rather than simply a trial judge’s case-terminating opinion, should one exist, that present the most fruitful opportunity to examine trial court judicial behavior. Take, for example, a sex discrimination case being litigated in the trial court where the case ultimately terminates via a pro-defendant jury verdict. An outcome-only or opinion-only focus would not study this case, because no judge decision on the verdict exists. However, a motions focus may reveal that prior to the trial, the judge denied the defendant’s motion to dismiss and motion for summary judgment. The judicial decisions on these motions are pro-discrimination claimant rulings in the case, albeit not ones that serve to conclude the litigation or that necessarily lead to published judicial opinions.

Second, many trial court cases, particularly in civil litigation, involve weak, potentially frivolous claims. As Epstein, Landes, and Posner (2013, 232) put it, “many cases are filed by pro se or emotional litigants and by litigants represented by inept or inexperienced lawyers.” Most of these cases will be dismissed or abandoned no matter who the judge is. Any trial court dataset that includes a high proportion of these inherently weak cases is therefore likely biased against finding judge-specific effects due to the swamping effect of these “easy” cases and the motions practice that lies within them. Such concerns may be particularly salient in diversity-connected issue areas like employment discrimination where many worry about high numbers of frivolous lawsuits (Nielsen and Nelson 2005).

Finally, the above-discussed theories predicting diversity effects on judging are generally issue area specific. As such, any test of them needs narrowly defined data that fit squarely within or outside of those issue areas. This means that trial court datasets focused on broad issue areas like “civil rights” or “civil liberties” may not be precise enough to detect differences in underlying judging behavior because the cases include not only those that fall within traditional women’s issues and minority issues but also those extending well beyond (e.g., assault, housing discrimination, immigration).

Data

This project uses the Kim, Schlanger, and Martin (2013) newly collected dataset of EEOC-brought federal trial court employment discrimination lawsuits. The Kim, Schlanger, and Martin data sample more than 2,300 EEOC cases filed between 1997 and 2006. They permit the above concerns inherent in most trial court-diversity judging research designs to be directly addressed.

First, the EEOC dataset is motion centered rather than opinion or case outcome-centric. Coded from case docket sheets and documents, the data include every dispositive and substantive motion that takes place within a case’s district court life. Each observation in the resulting data is a judge’s decision on the motion. This project focuses on a subset of the data which excludes motions that are not directly related to the case’s potential disposition or its merits like discovery motions, motions by a private complainant to intervene in the EEOC’s lawsuit, and motions to relate or consolidate multiple lawsuits together. This study also excludes motions filed with the consent of both parties because these motions strip trial judges of decision-making discretion. The data also exclude observations in which no judge decision was recorded, something that occurs when the case settles prior to the motion decision, and those in which the deciding judge was a magistrate judge or Article I district judge rather than an Article III district judge.

Second, the Kim, Schlanger, and Martin (2013) dataset’s inclusion of only those cases brought into the federal district courts by the EEOC has important implications for the type of cases that are in the dataset. The EEOC has statutory authority through the 1972 Equal Employment Opportunity Act to sue private employers to enforce Title VII of the Civil Rights Act of 1964 (which bans employment discrimination on the basis of sex, race, religion, and national origin), the Age Discrimination in Employment Act, and the Americans with Disabilities Act. Prior to bringing federal lawsuits to enforce these statutes, the EEOC conducts internal administrative review of discrimination complaints. For many complaints, the EEOC chooses to not file a federal lawsuit. The U.S. EEOC
(2014) describes its decision whether to file a federal district court lawsuit as follows:

When deciding whether to file a lawsuit, EEOC will consider several factors, including the seriousness of the violation, the type of legal issues in the case, and the wider impact the lawsuit could have on EEOC efforts to combat workplace discrimination.

In other words, the EEOC's decision to file a federal lawsuit is not random, but rather one that is based on their extensive internal investigation indicating the high quality and importance of pursuing a particular case into federal court (Hirsh 2008). As an illustration of the EEOC's significant filtering, in fiscal year 2001, the EEOC reported that 80,840 individual charges were filed with its office. During the same year, the EEOC filed just 428 federal district courts enforcement lawsuits. For purposes of this current project, the presence of this EEOC filtering means that, unlike most civil litigation, the likelihood of weak, frivolous cases littering the data and drowning any judge identity effects is low.

Third, the EEOC dataset codes the types of alleged discrimination present in a case. This permits the examination of a subset of employment discrimination cases that is narrowly focused on the diversity characteristics of interest here (sex and race). This is a particularly important point for a study of diversity in judging for two reasons. First, many of the diversity theories described above generally only anticipate diverse judges to behave differently from their colleagues in cases directly related to their diversity characteristic, meaning that overly broad data will prove inadequate for theory testing. Second, without this data narrowness, any statistical results are likely to be deceptively deflated or nonexistent. Here, these data are parsed down into two types of categories: sex and pregnancy discrimination and race and color discrimination. The data exclude sex discrimination claims brought on behalf of male plaintiffs and race discrimination claims brought on behalf of white plaintiffs.

Variables

The dependent variable for the study is Pro-plaintiff Outcome. Pro-plaintiff Outcome is coded as 1 if the judge's ruling on the motion provides the EEOC and/or a private plaintiff a clear victory on the motion and 0 if the judge's ruling clearly favors the defendant. These motions are, for example, motions for summary judgment in which a party asks the judge for a pretrial victory because "no genuine dispute as to any material fact" exists and, therefore, he or she “is entitled to judgment as a matter of law” (Federal Rules of Civil Procedure Rule [FRCP] 56) or motions to dismiss in which the defendant argues that the district court lacks the necessary jurisdiction to hear the case (FRCP 12[b]). Any time the judge issues a partial ruling, such as granting a motion for summary judgment in part and denying it in part, the case is excluded from the data.10

The primary independent variables in the study are the race and sex of the judge deciding the motion. Drawn from the Federal Judicial Center (2011), Female Judge is coded as 1 if the judge deciding the motion is a female, 0 otherwise. Black judge is coded as 1 if the judge deciding the motion is an African American racial minority, and 0 if he is white.11

This study also controls for a host of judicial, case, and legal variables that may affect the outcome of substantive motions in EEOC cases. Briefly, the control variables include judge ideology measures for the district judge deciding the motion along with the ideology and ideological variance of his supervising court of appeals (Giles, Hettinger, and Peppers 2001; Epstein et al. 2007). It is also important to control for the type of motion. Earlier stage motions should be more nonmovant favorable than later stage motions. Similarly, as motions practice is designed to favor the nonmoving party, controlling for the identity of the movant is important. So is accounting for whether the nonmoving party opposed the motion in writing, something that, when present, may increase the odds of denial. At the case level, the study also controls for the major issue or issues alleged in the complaint and for whether the underlying case has one or more than one complainant.

Results

To statistically examine whether a federal trial judge's sex and race affect his or her decision making, we turn to a logistic regression analysis of the two sub-datasets of interest—that is, sex and pregnancy employment
discrimination case motions and race and color employment discrimination case motions. To address concerns about independence of observations, the modeling includes robust standard errors clustered on the individual case. The results from the modeling for the two issue area data-sets are provided in Table 1.12

Let us begin with the results for the effect of judge sex on trial court decision making. Model 1 provides these results for sex and pregnancy discrimination case motions. Recall that many diversity theorists anticipate that female judges will be much more likely than male judges to rule in favor of the plaintiff in these sex discrimination cases. As the results indicate, this is exactly what we find. Female Judge is positive and statistically significant.

The substantive effect of this variable is further explored in Figure 1. There, the left side figure plots the predicted probability of a male (open circle) judge and female (solid circle) judge ruling in favor of the EEOC or the private plaintiff in the motion. All other variables in the model are held at their median (for continuous variables) and modal (for dichotomous variables) values. For female trial judges, the average predicted probability of ruling for the sex discrimination plaintiff is .35. For the male, this number is only .20. As the accompanying right side figure indicates, this difference in the average predicted probability is .15 and is statistically different from 0. This result thus provides substantial evidence that female trial judges do indeed behave distinctly from their male colleagues when faced with decision-making opportunities firmly triggering “women’s issues.”

When it comes to race discrimination case motions, however, there is no statistical evidence that male and female judges systematically behave distinctly from one another. The Female Judge variable is positive in the second regression model, but it does not reach statistically significant levels. This point is driven home by the large confidence interval (CI) on the plotted difference in Figure 1’s right side panel that clearly

Table 1 • Logistic Regression Results for Whether a Motion’s Outcome Is Decided in a Pro-Plaintiff Direction

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1: Sex discrimination case motions</th>
<th>Model 2: Race discrimination case motions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female judge</td>
<td>0.795** (0.35)</td>
<td>0.794 (0.73)</td>
</tr>
<tr>
<td>Black judge</td>
<td>1.471** (0.52)</td>
<td>1.655** (0.76)</td>
</tr>
<tr>
<td>Ideology controls</td>
<td>Included</td>
<td>Included</td>
</tr>
<tr>
<td>Motion controls</td>
<td>Included</td>
<td>Included</td>
</tr>
<tr>
<td>Complainant number control</td>
<td>Included</td>
<td>Included</td>
</tr>
<tr>
<td>Issue(s) alleged controls</td>
<td>Included</td>
<td>Included</td>
</tr>
<tr>
<td>Constant</td>
<td>−1.257 (1.67)</td>
<td>−1.806 (2.45)</td>
</tr>
<tr>
<td>Observations</td>
<td>450</td>
<td>186</td>
</tr>
<tr>
<td>Percent reduction in error</td>
<td>15.60%</td>
<td>23.53%</td>
</tr>
</tbody>
</table>

Standard errors are robust, clustered on the individual case, and presented in parentheses. The baseline motion type is Summary Judgment Motion. Full regression results including all control variables are provided in the online appendix.

** p < .05 (two-tailed).
overlaps 0 (CI = [−0.16, 0.53]). This lack of result for Female Judge in the race discrimination dataset is not surprising. Diversity theory expects that female judges will generally only behave differently from their male colleagues when acting as substantive representatives of their class or when using their unique professional experience and knowledge related to being a woman—that is, things that are certainly put into play in sex discrimination cases but likely not in most race discrimination cases.

Copyright ©2019 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.
Turning now to an examination of the effect of judge race on trial court decision making, it makes sense to first focus on Model 2 (race discrimination case motions). Just as with judge sex and sex discrimination cases, there is much theory in support of the expectation that black trial judges will be much more likely to support race discrimination claimants than white trial judges. The positive and statistically significant coefficient for black judge in Model 2 empirically supports this. Figure 2 once again plots the substantive effect. As the bottom portion of the left hand plot indicates, the average predicted probability of a black judge ruling in favor of the race discrimination plaintiff is .70. This number is only .31 for similarly situated white trial judges. The average difference in this predicted probability is .39 and is statistically different from 0 (see right side of Figure 2), meaning that there is a 126 percent increase in the likelihood of a black trial judge ruling in favor of the EEOC’s race discrimination claim in a dispositive motion over a white trial judge.

Interestingly, there is also a positive and statistically significant effect from a judge’s race on voting behavior in sex discrimination case motions. This is apparent by examining black judges in Model 1 and the top section of Figure 2. As the figure shows, black trial judges have an average predicted probability of ruling in a pro-plaintiff manner on sex discrimination case motions of .514; for white judges, this number is .195. This difference, nearly .32, is again impressively large and both substantively and statistically significant. This large effect for black trial judges deciding sex discrimination case motions may be explained by the above-discussed balance theory—that is, the idea that black judges are sensitive to others facing discrimination.

Finally, for the control variables, in both the sex and race discrimination data, the variable capturing the fact that the defendant is the movant is positive and significant. This result reflects that the nonmoving party holds an advantage in district court motions practice. In the sex discrimination data, the results also indicate that multiple types of motions are significantly more likely to lead to pro-plaintiff outcomes than the baseline type of motions, Summary Judgment Motions. In neither dataset is district judge ideology statistically significant. This nonresult is robust to the measurement of district judges’ preferences, with Party of the Appointing President substituted for District Judge Ideology yielding no difference in effect. Although this may come as a surprise to some political scientists used to seeing attitudinal results in judicial behavior, a lack of effect is not uncommon in many district court studies (e.g., Howard 2002; Keele et al. 2009; Walker 1972; Zorn and Bowie 2010).

Discussion

As noted in the outset of this project, previous attempts to find diversity effects in judging in the trial courts have been largely unsuccessful. By reexamining the research design used for analyzing diversity in judging in trial courts and focusing on data generated from judicial decisions on motions rather than simply case outcomes or published opinions, data that filter out frivolous cases, and data that are narrowly focused on individual issue areas of theoretical interest to diversity theory, this project has been able to bring new clarity to this important issue. Indeed, the results here provide some of the first consistent, systematic evidence that female and racial minority trial judges do indeed behave differently, and rather substantially, from their traditional colleagues. Female trial court judges, when deciding dispositive motions in sex discrimination cases, are about 15 percent more likely to rule in favor of the discrimination claimant than male judges. Similarly, in race discrimination cases, black trial court judges are about 39 percent more likely to decide in favor of the race discrimination plaintiff than white judges. These results thus join other recent work finding strong diversity judging effects in certain settings, whether that be outcomes in federal appellate courts (Boyd, Epstein, and Martin 2010; Kastellec 2013), opinion crafting, oral argument questions, conference behavior, and the hiring of clerks and office staff in appellate courts (e.g., Haire, Moyer, and Treier 2013; Kaheny et al. 2015), or nonmerits trial judging activities like whether to actively encourage settlement (Boyd 2013) or what decisions to make during in-court administrative proceedings such as arraignment and bail requests (Fox and Sickel 2000).

Although the findings here provide important insight into trial court judging, particularly in the diversity context, there are inherent limitations in the project’s generalizability that merit note. One of these stems from the focus of the underlying database only on those cases brought by the EEOC into the federal district courts. As discussed in detail above, the EEOC focus serves many empirical and research design benefits and is similar
in scope to other recent studies examining one type of case—whether that be, for example, cases only brought by the Solicitor General (Black and Owens 2012, 2013; Zorn 2002) or amicus curiae (P. M. Collins 2004, 2007; Collins and Martin 2010) or only those that eventually lead to an appeal (Randazzo 2008). Here, just like many of these other litigants and cases, the EEOC is not an average, typical plaintiff in trial courts but, rather, is a strong repeat player that makes smart decisions regarding which cases to file and which cases not to file. Would we observe the same results by examining employment discrimination lawsuits filed by all types of plaintiffs? Although only future research can answer this question with certainty, the EEOC focus in this project means that the results here are most likely to be generalizable to other plaintiff-brought lawsuits that are filtered and well-crafted, meaning that, for example, an average individual employment discrimination plaintiff bringing a lawsuit declined by the EEOC will probably be less likely to hold the same advantage before diverse judges.

A second data-linked limitation of this project rests with its inability to speak directly to the important issue of intersectionality—that is, the idea that, as once stated by Crenshaw (1992, 1467–68), black women face the “dual vulnerability” of racism and sexism that intersect “to create experiences that are sometimes unique.” Applied to judges, T. Collins and Moyer (2008, 225) argue that “minority women have a distinctive identity” that affects their behavior. For trial court judges, this could, in theory, lead to black female judges being more likely to support female sex discrimination plaintiffs, black race discrimination plaintiffs, or black female plaintiffs raising sex and race discrimination claims simultaneously. Indeed, a disproportionate number of the early sex harassment lawsuits were filed by black women (MacKinnon 1979, 53), and some discrimination lawsuits even today include both race and sex discrimination claims (see Note 9). Unfortunately, the current study’s number of black female judge observations is too small to provide an empirical test of intersectionality. There are just three total black female judges in the data and not a single black female judge observation in which the judge presides over a motion with both a race and sex discrimination claim in the underlying lawsuit. Only future research, aided by a larger sample of data and the ever growing number of female, racial minority judges on the bench, will be able to fully tackle this issue at the district court level.

In addition, the data used here focus on two types of cases—race employment discrimination and sex employment discrimination. As such, this project cannot be confidently applied to other issue areas like criminal sentencing, abortion, family law matters, or beyond. The theory of difference in diversity judging is much less uniform in these areas, largely because it is not as easy to connect a judge’s race or sex with expected preferences over issues not directly linked to sex or race. Although female and black trial judges may behave differently from their colleagues in other issue areas, only future empirical tests can provide confirmation. Similarly, this project’s data, by design, do not represent all parts of a district judge’s job. As discussed above, many of a district judge’s duties are not discretionary and likely do not leave room for her preferred outcomes to be implemented into behavior.

Even after acknowledging these limitations, the implications of the diversity effect found here are vast. Federal trial courts are now responsible for hearing more than 300,000 cases per year—a number that far exceeds those heard by U.S. Courts of Appeals and the U.S. Supreme Court—and composed of nearly 32 percent female judges and more than 13 percent black judges. The fact that some trial court cases are affected so drastically by the identity of the judge has repercussions not just for the outcome of the immediate case and the likelihood of the plaintiff being successful. Rather, such an influence reverberates through the judicial hierarchy and beyond (Baum 1997, 2006). Although the federal district court loser retains the right to appeal, fewer than 20 percent do so (Eisenberg 2004). And, when they do appeal, the appellate courts are generally required to be quite deferential to the lower court decision. In other words, what happens in trial court judging matters.

Finally, in returning to this paper’s introduction, this project is particularly timely given President Obama’s judicial appointment priorities. Recall that the Obama Administration has argued that while diversifying the federal trial courts is important for creating a judiciary that “resembles our nation” (Kang 2013), it does not mean that these judges will consider cases differently from traditional judges. The evidence provided here strongly disputes this point, meaning that President Obama has not only been altering the descriptive identity of the nation’s federal trial courts through his diverse appointments but
also affecting the resulting substantive policy and law—now and well into the future given the lifetime tenure held by these appointees.

Notes

1. By comparison, previous presidents had the following diversity appointment percentages to the federal district courts: 33 percent (Carter), 16 percent (Reagan), 36 percent (G. H. Bush), 52 percent (Clinton), and 34 percent (G. W. Bush). These calculated percentages include all judges who were female, minority, or both and do not double count female and racial minorities.

2. Recent scholarship calls for the testing of the intersection of race and sex in diversity judging (intersectionality). This topic is tackled in further detail below in Note 9 and in the “Discussion” section.

3. By contrast, for appellate courts, research paints an increasingly consistent picture: judge diversity matters, but generally only in issue areas directly connected to the diversity characteristic. For female judges, that means a higher likelihood to favor claimants in women’s issue-type cases like sexual discrimination and sexual harassment (e.g., Boyd, Epstein, and Martin 2010; Crowe 1999; Farhang and Wawro 2004; Haire and Moyer 2014; Peresie 2005; Songer, Davis, and Haire 1994). For nonwhite judges, areas like voting rights, race discrimination, affirmative action, and some criminal matters tend to produce differential judicial behavior (e.g., Cox and Miles 2008; Crowe 1999; Kastellec 2013; Scherer 2004).

4. As Siegelman and Donohue (1990) find, less than one out of four employment discrimination cases yield publicly available opinions. Thos opinions that are available are not representative of broader discrimination litigation.

5. More than 29 percent of the cases in the Kim, Schlanger, and Martin (2013) data have no recorded motions.

6. Including discovery motions in the statistical analysis below does not alter the results reported.

7. Statutory and administrative procedure requires nearly all employment discrimination complaints that may eventually be brought into federal court to first be filed with the Equal Employment Opportunity Commission (EEOC).

8. The Pregnancy Discrimination Act of 1978 prohibits “sex discrimination on the basis of pregnancy.” This project includes workplace pregnancy discrimination allegations with sex discrimination. The results do not change if pregnancy discrimination cases are excluded.

9. Plaintiffs can plead multiple types of discrimination in their district court complaints. Within the Kim, Schlanger, and Martin (2013) data, multiple claim pleading is relatively rare. For example, in the sex discrimination data, less than 10 percent of the case motion observations include another form of discrimination. Although all major discrimination claim types within the Kim, Schlanger, and Martin (2013) data (race, national origin, age, disability, and religion) are paired with sex discrimination claims at least once, race is the most common (about 8.5% of all sex discrimination observations).

10. An alternative coding scheme for these partial decisions is to include them in the data and code them as a loss for the movant. Although the moving party received some relief, the judge’s decision denying the motion in part results in the case continuing despite the movant’s efforts to the contrary. The inclusion of these additional observations does not change the statistical results for the key judge sex and race variables.

11. This variable’s coding excludes from the data those district judges that are neither black nor white. Previous work suggests that black judges are distinct from other minority judges in these cases. Here, these nonblack minority judges amount to a handful of observations, and alternative coding schemes, including grouping these judges with white judges (Kastellec 2013) or with black judges (T. Collins and Moyer 2008), have no effect on the results reported below.

12. Additional modeling confirms the robustness of the results. Following propensity score nearest-neighbor matching (Ho et al. 2006), the estimated differences were statistically significant and in the expected direction for judge race and sex. Only descriptive (noncausal) inference is possible in this study of individual judge behavior (Rubin 2006). Multilevel modeling with motions nested within cases, while not ideal here due to a relatively small number of observations and an average of only two motions per case, also confirms the results reported in the main text.

13. As initial evidence on the effect of black female judges in the data, the statistical results reported below remain unchanged after excluding the small number of these judges from the data.

References


Juries have the power, but not the right, to nullify. As a result, judges generally do not instruct jurors on nullification. If jurors learn about this power, they may exercise it. However, if they indicate their intentions during deliberations, they can be removed from the jury. The author of this article contends that this state of affairs, in which the judge does not tell jurors of their full power and jurors have to be careful not to reveal they are aware of their right to nullify, is harmful to the judge–jury relationship. If judges instructed jurors on their power, jurors would be fully informed and would be able to trust the judge.

I. Introduction

Jurors in the American jury system have the power to engage in jury nullification but they do not learn about it from the judge. Judges do not instruct jurors on nullification. If jurors learn about this power from an outside source, they can exercise it, but they cannot mention it to the judge during their deliberations. If they do, they might