Who Shapes the Law? Gender and Racial Bias in Judicial Citations

JOHN SZMER  University of North Carolina at Charlotte, United States
LAURA P. MOYER  University of Louisville, United States
SUSAN B. HAIRE  University of Georgia, United States
ROBERT K. CHRISTENSEN  Brigham Young University, United States

In this letter, we assess whether the contributions of judges from underrepresented groups are under-valued or overlooked, thereby reducing these judges’ influence on legal policy. Drawing on an original dataset of discretionary citations to over 2,000 published federal appellate decisions, we find that the majority of opinions written by female judges receive less attention from other courts than those by similarly situated men and that this is largely attributable to disparities in citing Black women and Latinas. We also find that additional efforts by Black and Latinx judges to ground their opinions in precedent yield a much lower rate of return in subsequent citations by outside circuits than comparable work by white men and women judges. This suggests that, despite gains in diversification in the federal judiciary, stereotypes about social identities will play a powerful role in determining whose ideas receive recognition.

Citation patterns signify the recognition and authority of others’ ideas (Hinkle 2021) and allow us to understand factors that shape the diffusion of norms and policies. In common law systems, appellate judges are expected to cite binding legal precedent to guide their decisions, but there is no such expectation for nonbinding (“persuasive”) precedent, as citations to these rulings are purely discretionary. As such, we can learn a great deal about the influence of decisions, and those who author them, through studying these discretionary citations. With thousands of nonbinding opinions from which to choose, why are decisions by some judges better able to catch the attention of their peers and afforded more influence in the development of law?

The legal model of judicial decision-making suggests that decisions about whether to cite a nonbinding case are based on factors such as the quality of the reasoning or the legal importance of the cited case. Judges and their law clerks may also seek out published decisions by judges with strong educational or professional credentials. Yet, the research on citation bias in academic publishing points to a third explanation: implicit biases and socializing processes connect gender and racial identities to stereotypes about competence and authoritativeness.

In this letter, we analyze out-of-circuit citations to a sample of 2,000 published decisions from 2009 to 2016 to assess whether there is evidence of gender and racial disparities in discretionary citation practices by federal appellate judges. Because precedent from parallel courts is not binding, judges are free to ignore or cite out-of-circuit precedent completely at their discretion. Our results indicate a bias against citing majority opinions authored by women, especially women of color. We also find that the additional time and care put in by Black and Latinx judges in fully developing their opinions yield a lower rate of return than comparable efforts by white judges. These findings suggest that the conflation of prestige and social categories within the federal judiciary systematically limits the impact of judges from underrepresented groups to influence the law.

COMPETENCE, CUES, AND INFLUENCE

To understand how biases may affect legal influence, it is instructive to look to studies in other contexts to see how racial and gender cues affect evaluations of competence (e.g., Ellemers 2018) and are linked to power structures that devalue the contributions of certain groups (Ridgeway 1991). In one line of inquiry, callback studies identify a persistent bias against job applicants with putatively Black or Latinx names (Quillian et al. 2017), while audit studies in male-dominated STEM fields uncover a preference for male students,
researchers, and employees over identically qualified women (Milkman, Akinola, and Chugh 2012).

These studies show that context matters. Where white men are overrepresented in powerful roles (e.g., law and politics), the association between gender, race, and authority will be strong (Rudman and Kilianski 2000). If gender and race operate as legitimacy cues, then white men may be accorded more prestige simply by virtue of their identity, compared with other groups (Berger et al. 1977). In the judicial context, gender and race affect official evaluations of judges: American Bar Association ratings and state judicial performance evaluations have been found to display bias against women and judges of color (Gill, Lazos, and Waters 2011; Sen 2014). Critics note that law schools undermine the self-confidence of minoritized groups by ignoring their vantage points in core classes (Fischer 1996). This “cultivat[es] generations of lawyers, scholars, legal instructors, and judges who accept and promote the dominant white hegemonic view of procedure as neutral” (Pedro 2021, 162).

Studies of academic citations also show how the contributions of individuals from underrepresented groups receive less attention from their peers, affecting their influence in ways that can be applied to the legal context. An analysis of 1 million dissertations found that while women and underrepresented minorities are more likely to generate scientific innovation in their work, these innovations are less likely to be cited by other scholars compared with work by men, white, and Asian scientists (Hofstra et al. 2020). Similarly, academic task forces have also uncovered citation disparities for research authored by women, Latinx, and Black authors, limiting the reach and influence of these groups’ intellectual contributions (e.g., APSA 2022).

Surprisingly, little work exists on gender and racial bias in judicial citations. Choi and Gulati (2008) examine bias in out-of-circuit citations, but only for one year during a period when the federal judiciary was insufficiently diverse to assess racial bias (see also Landes, Lessig, and Solimine 1998). More recently, a comprehensive study of binding and nonbinding precedent on search-and-seizure cases failed to find evidence of out-group bias in white judges’ citation patterns, but did reveal small in-group effects for racial minority and female judges (Hinkle and Nelson 2018). Shared social bonds, as well as educational and professional backgrounds, have also been linked to mutual citation patterns (Choi and Gulati 2008; Hinkle and Nelson 2018).2

When there is discretion about whose nonbinding opinions to cite, concerns about prestige and reputation (Klein 2002) likely come into play as judges and their law clerks draft the majority opinion with an aim toward persuading colleagues, litigants, and the reviewing court of its legal correctness.3 Law clerks also have an interest in impressing their supervising judge when identifying precedents for citation, regardless of their own views about whom to cite. Stereotypes about competence have a reinforcing quality; that is, observing others’ signals about which judges have more authority and prestige solidifies the connection between status and social categories as law clerks are socialized into their work (Johnson, Dowd, and Ridgeway 2006). Moreover, algorithms used by legal databases such as Westlaw may exacerbate human biases by promoting opinions that have already been cited in search results (Hinkle and Nelson 2018), reinforcing a link between perceived status and citations.

DATA AND METHODS

To assess whether the legal influence of an opinion is affected by the author’s race and gender, we utilize a stratified, random sample of 30 published U.S. Courts of Appeals decisions4 per geographic circuit, per year from 2009 to 2016, and added data on discretionary citations to those decisions. We exclude citations to unsigned opinions, en banc cases, certification decisions, and opinions by district court judges, resulting in 2,228 observations.5 Our approach differs from other studies on citation (e.g., Hinkle and Nelson 2018) in several respects. Notwithstanding the value of studying citation patterns in a single-issue area, we focus on discretionary citations across all issue areas to identify broad trends in policy influence across the federal appellate courts. Additionally, we focus on factors that attract attention to a judge’s work product, rather than reciprocal factors that link a citing judge with a cited judge.

As such, the dependent variable is the number of times that the majority opinion is subsequently cited in out-of-circuit opinions over the five-year period after the initial opinion was decided.6 That is, if the initial case was decided by the 1st Circuit on February 15, 2013, we counted the number of cases decided by the 2nd–11th and D.C. circuits between February 16, 2013, and February 15, 2018, that cited the original 1st Circuit Case. Because this is a count variable with strong evidence of overdispersion, we estimated

---

1 With this limited sample, they find no evidence of gender bias.
2 Related work finds that women and minority judges are less likely to be assigned to write for the majority (e.g., Kaheny, Szmer, and Christensen 2020).
3 Law clerks assist federal judges with legal research and writing (Bowie, Songer, and Szmer 2014). White men have long been overrepresented in clerkships (NALP 2017).
4 We use the “Multi-User Database on the U.S. Courts of Appeals (2009-2016)” available on the Harvard Dataverse at https://doi.org/10.7910/DVN/3QKZB7. The period leverages the historic increases in diversity of the Obama appointees while giving us 5 years to observe subsequent citations, with 2021 as the endpoint of the citation window. Full replication files for this article can be found on the APSR Dataverse (Szmer et al. 2023).
5 See the Supplementary material OA1.
6 While some studies distinguish between positive and negative treatments of precedent, the latter is a rare occurrence in our data, as it is usually more expedient to ignore nonbinding precedent. In a sample of our cases, positive treatments of a decision were highly correlated with total citations (r = 0.94). See Supplementary material OA1.
7 Our results are also robust to shifting the start of the citation window to three months after publication. See Supplementary material OA1.
negative binomial regressions and included year-level unconditional fixed effects and standard errors clustered by circuit.

To test whether an authoring judge’s sex or race affected the degree to which subsequent courts draw on their opinions, we utilize the Federal Judicial Center database to code Female Judge (0,1) and Non-white Judge (0,1). The latter was coded 1 if the judge was listed as African American or Latinx and 0 if the judge was listed as white. Because prior work suggests that Asian Americans face different stereotypes (e.g., the “model minority”) and are not disadvantaged in citations (APSA2022; Hofstra et al.2020), we exclude the small number of majority opinions by these judges, though our findings do not change with their inclusion. (For more discussion, see Supplementary material OA3.)

Because well-reasoned opinions are more likely to influence judges in other courts (Hume 2009), we follow previous work (Moyer et al. 2021) that relies on a novel measure of how well an opinion is grounded in the case law. This indicator, Deep Cites, is a Westlaw-generated measure of the number of precedents discussed in the source opinion using at least two paragraphs of text. Our model also includes controls for many other influences on citations, including indicators of the opinion author’s background and elite status identified in other works (e.g., Holmes 2019) and controls for readability, issue salience, and ideological divisions on the authoring panel. For more information on variables, see Table 1 and Supplementary material OA1.

**RESULTS**

Figure 1 displays a coefficient plot of the key independent variables from Model 1 (see Supplementary material OA2). As expected, majority opinions written by women judges receive significantly fewer subsequent citations than male-authored decisions by outside circuits. Over the subsequent five-year period, female judges, on average, are expected to receive 1.4 external citations to their opinions, compared with 1.63 citations to opinions written by similarly situated men. While seemingly small, when accumulated over a career, the effect is substantial. For the average judge producing 53.5 opinions a year, in 10 years of full-time service, male judges would have 120 more citations than a female judge. This effect exists after controlling for factors that affect the likelihood of subsequent citation, including judge attributes and case characteristics. Conversely, there is no statistically significant difference in the total number of citations for opinions authored by non-white judges when compared to those by white judges.

To unpack our results further, we look to research that evaluates Justice Sotomayor’s observation that “[i]f you are a person of color, you have to work harder than everybody else to succeed” (de Vogue 2021). For
instance, Biernat and Kobrynowicz (1997) find that women and Black individuals are evaluated as less competent and consequently have to work harder to meet a higher standard than white and male individuals. Hofstra et al. (2020) similarly find that when women and underrepresented minorities produce more innovative outputs, they receive less credit than their white, Asian, and male peers. In the judicial context, recent work finds that both non-white and female appellate judges do more to justify their rulings when assigned to write the majority opinion (Moyer et al. 2021). So, are all judges similarly situated in their ability to translate effort dedicated to strengthening opinion quality into legal influence?

To explore the possibility that gender and racial identity will dampen the influence of opinion quality for underrepresented groups, we estimate a second model with two multiplicative terms: Female Judge x Deep Cites and Non-white Judge x Deep Cites. The coefficient plot in Figure 2 (top panel) shows the degree to which a majority opinion thoroughly engages precedents and translates to increased attention by outside circuits. Moreover, the coefficients for the multiplicative terms indicate racial, but not gender, bias. This reveals an asymmetry in influence, whereby white judge-authored opinions are more influential when they contain more lengthy precedential discussions, while similarly well-developed opinions written by Black and Latinx judges are recognized less often by their judicial peers from other courts. We then dropped the judge-gender multiplicative term (Figure 2, bottom panel) and reestimated models focused on racial bias with expected counts presented in Figure 3.

Figure 3 plots the expected number of subsequent citations by the number of Deep Cites and shows the dramatic difference between white and non-white judges. As the number of Deep Cites increases in opinions written by white judges, the expected number of out-of-circuit citations rises exponentially. White judges clearly experience a large benefit from the additional effort explaining their ruling, an advantage that is not shared by their Black and Latinx colleagues.

To further explore these dynamics, we test whether the race and gender of the opinion author affects the influence of an opinion outside of the circuit. Figures 4–6 show the results of models that include intersectional variables: Non-white Female, Non-white Male, and White Female, with White Male as the reference group. In Figures 4 and 5, Black female and Latina judges receive significantly fewer subsequent citations than white men. The coefficient for white women is also negative (but narrowly misses significance at the 0.05 level). Although white women and men receive more discretionary citations when their opinions contain more discussion of precedent, non-white women and men judges do not see the same benefit (see Figure 6).

DISCUSSION

Overall, our findings paint a concerning picture of the correlates of judicial influence linked to gender, race, and their intersection. First, we find, even after controlling for other factors, women receive fewer discretionary citations for their majority opinions than their male peers. Further analysis (see Supplementary material OA5) suggests that this gender gap in citations diminishes over time, particularly for white women. While we could not identify the cause of the shrinking
gap due to our limited time frame, we hope future research will examine possible factors, such as increasingly gender-diverse appellate courts.

The discrepancy is most substantial for opinions authored by Latinas and Black women, who receive 28 percent fewer discretionary citations than white males. This difference does not shift over the time period (see Supplementary material OA5). Over 10 years of full-time judicial service, this translates to a sizeable disadvantage for women of color: about 239 fewer citations than for white men.

Second, we find that when white men and women judges write opinions containing more extended discussions of precedent, they are rewarded with more external citations. In contrast, when Black and Latinx judges engage more fully with legal precedent in their
opinions, this has a much smaller impact. To return to Justice Sotomayor’s comment, “work(ing) harder” is not enough to counteract systematic bias.

While these findings align with other research on how status beliefs contribute to inequality in influence (Johnson, Dowd, and Ridgeway 2006), we recognize that our study does not explicitly test for in-group favoritism as Hinkle and Nelson (2018) do. Looking at an earlier time period, they do not find evidence of in-group bias by white male judges; however, women and non-white judges were somewhat more likely to cite in-group opinions. Our approach complements
their study by examining overall citation patterns during a more diverse era in the federal judiciary and by incorporating an intersectional approach. We see that doctrinal contributions by women of color judges are consistently undervalued, whereas opinions authored by white women become more influential over time. Additional research is needed to explore how these dynamics shift over time and whether citation trends are driven by factors such as explicit bias or individual differences in self-promotion. Moreover, the time period of our study limits our ability to assess the impact of Asian American judicial appointees, who experience a different set of performance stereotypes that should be explored. Although we do not find an effect for judges affiliated with the Federalist Society (see Supplementary material OA4), it is possible that professional organizations fuel the linkage between social categories and prestige in a way that affects influence.

Despite these limitations, our findings have important implications beyond judicial decision-making, including for policy diffusion. When a political actor has complete discretion about which policy to emulate, our work suggests that stereotypes about social categories, competence, and status will play a role in determining whose ideas receive recognition. Future research could explore whether this extends to the legislative context by examining the impact of bills introduced by legislators from underrepresented groups. Our findings also suggest that it is not enough to diversify the judiciary; socialization processes will also need to be addressed to produce equitable outcomes and allow the law to benefit from the full spectrum of perspectives on the bench.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0003055423000801.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: https://doi.org/10.7910/DVN/NWO52N.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge that the data for this project were collected with support from the National Science Foundation (NSF SES #1655159, 1654614, 1654559, and 1654697). Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors thank the editors and the anonymous referees for their helpful feedback; all remaining errors are our own. We also thank Ethan Boldt, Vanisha Kudumuri, Alyson Hendricks-Benton, Mariliz Kastberg-Leonard, Jake Truscott, Madison Eckels, Teneia Segi, and Samuel LaVergne for their research assistance. Finally, the authors specially thank Ben Radford for generating the readability measure and Barry Edwards for his years of hard work on the project.

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

Note: N=2,228. See Supplementary material OA2 for full results.
ETHICAL STANDARDS

The authors affirm this research did not involve human subjects.

REFERENCES


