

## *Chapter 6*

# **Biases in Judicial Decision-Making**

Naci Mocan

### **INTRODUCTION**

It is a well-documented fact that judicial outcomes are correlated with defendants' personal attributes. For example, Blacks are more than four times as likely to be incarcerated as Whites (Wagner, 2012). In 2015, the U.S. population was 13 percent Black (U.S. Department of Commerce, 2018) while the U.S. state prison population was 38 percent Black (Carson, 2015), and Black death row inmates constituted more than 40 percent of all death row inmates (Statistica Research Department, 2019). Such disparities in judicial outcomes between groups of individuals, however, do not imply disparate treatment. A number of differences between the groups can cause these disparities. For instance, the arrest rate of Blacks for murder and nonnegligent manslaughter (arrests per Black population) is more than six times higher in comparison to the arrest rate of Whites for the same crime (FBI, 2018; U.S. Department of Commerce, 2018). Even if there was no racial inequality in any intermediate step of the judicial process between arrest and sentencing, a higher Black arrest rate for murder would lead to disproportionate representation of Blacks on death row. In other words, given such differences in arrest rates, no matter how impartial and fair judges are, outcomes would still be racially disparate.<sup>1</sup>

Income could also be responsible for some of the racial differences in judicial outcomes. Median household income for Black households was \$40,500 in 2017, while it was \$68,145 for White households (U.S. Department of Commerce, Census Bureau, 2018). If Black defendants cannot afford high-quality defense attorneys, poor defense during trial will lead to higher conviction probabilities and longer sentences for minorities (Anderson & Heaton, 2012). Along the same lines, minorities are more likely to rely upon court-appointed attorneys, and it has been shown that these attorneys may put

forth lower effort to defend their clients, which could lead to worse outcomes (Agan, Freedman, & Owens, 2018). Thus, the level of income, which partly determines the quality of legal representation, may stand as a contributor to racial inequality in case outcomes. It is therefore important to determine the specific factors, some of which can be outside of the judicial system, that can cause race and gender disparities in judicial outcomes, and to control for the impact of these factors.

The criminal justice process is lengthy, and it is multilayered. There are a number of key actors—prosecutors, criminal defense attorneys, among others—of which judges are just one. The ultimate decisions made by judges embody the decisions made by these other actors before the case has arrived in front of the judge. This realization poses a challenge to the investigation of judicial bias because it is important to filter out various confounding factors to pinpoint the sources of any potential bias, if it exists, and to design effective public policy to eliminate the bias. Thus, an overarching examination of judicial bias would investigate the actions/decisions of each actor in the judicial system separately, while accounting for the fact that one actor's decision may impact the decision of another actor, and that these decisions may be interdependent. For example, judges' decisions may reveal racial bias even when judges themselves are racially unbiased, if it is the case that prosecutors show bias. Alternatively, racially unbiased prosecutors may adjust their behavior in how they handle cases based on the race of the defendant if they recognize that judges are racially biased. Thus, analyses of judicial decision-making need to be statistically sophisticated to account for observable and unobservable factors that determine the judicial decisions under investigation. They also need to take into account the institutional details of the judicial system, and they should consider and adjust for the implications of the decisions made at earlier stages of the process. Inability to tackle these issues in a statistically satisfactory way would diminish the ability to make causal interpretations regarding the impact of the defendants' race or gender on the outcome of the case.

In succeeding sections of this chapter, I will focus on the decisions made by judges, and I will summarize the results of recent research that has paid attention, to the extent possible, to the issues discussed above. I focus on judges because they are key players in the judicial system and because there is more research on judges in comparison to other actors in the judicial system, such as prosecutors, defense attorneys, or jurors.<sup>2</sup>

In the next section, I describe briefly how to think about statistical analysis of judicial decision-making, highlighting potential complications. I will then summarize the state of knowledge about the impact of defendant's race and gender, the characteristics of judges, the importance of motions, and the impact of race-matching on judicial decisions.

## HOW TO ANALYZE THE DECISIONS OF JUDGES

To identify the “true” and direct impact of the defendant’s race, ethnicity, or gender on the outcome of a case, one needs to account for every aspect of a case that may be correlated with the race, ethnicity, or gender of the defendant and which may also impact the outcome of the case. Thus, researchers aim to find two groups of cases that are identical in all possible dimensions (e.g., the aggravating and mitigating circumstances, the criminal history of the defendant, the attributes of the prosecutor, the characteristics of the defense attorney, and the income and education level of the defendant), but that differ only in the race of the defendant. If the resolution of these cases is different (e.g., the defendants in the first group are convicted more frequently in comparison to the defendants in the other group), then the disparity can be attributed to the racial differences between the defendants. To make this idea clearer, consider the regression Equation (1) below, which can be run on a sample of adjudicated cases, such as cases decided in federal courts over a given period of time. The specification shown in Equation (1) and all other discussions throughout the paper assume that individual defendants can be matched with the adjudicating judge.

$$Y_{djc} = \alpha_1 + \beta_1 \text{Minority Defendant}_d + X_d\Phi + v_j + \lambda_c + \varepsilon_{djc} \quad (1)$$

where  $Y_{djc}$  stands for the outcome of a case with defendant  $d$ , adjudicated by judge  $j$  in courthouse  $c$ .  $Y$  could be an indicator to represent whether the defendant is convicted by judge  $j$ ; it could represent the sentence length meted out by the judge, or it could be some other outcome. The vector  $X$  contains the variables that capture various relevant attributes of the defendant and the case, ranging from criminal history of the defendant to the details of the alleged crime. *Minority Defendant* is a dichotomous indicator which takes the value of one if the defendant is a racial minority and zero if the defendant is White. If  $Y$  represents an undesirable outcome for the defendant, such as conviction, and if the estimated coefficient  $\beta_1$  turns out to be positive, this indicates racial discrimination because it implies that the average minority defendant has a higher probability of conviction in comparison to White defendants, holding constant personal attributes and case characteristics.

Judges may differ from each other in their propensity to convict and in their harshness in sentencing. Because each judge decides many cases, such idiosyncratic unobservable differences between judges can be accounted for by including a dichotomous indicator for each judge  $v_j$  (judge fixed effects) in Equation (1). Similarly, there could be unobserved characteristics and cultures of courthouses that may impact the behavior of each judge in that courthouse (and the outcome of each case in that court) in a particular way;

but these characteristics and cultures could vary between courthouses. That is, all judges in a particular courthouse of a given district as a group may be more or less strict in their judicial decisions in comparison to judges in a different courthouse. Such courthouse-specific differences that may influence case outcomes in the same way in a given courthouse but that may differ between courthouses can be controlled for by including a set of dichotomous courthouse indicators ( $\lambda_c$ ).

The error term  $\varepsilon_{jic}$  captures all other factors that impact the outcome  $Y$ . Any determinant of  $Y$  that is left out and not included in Equation (1) is embedded in the error term  $\varepsilon$ . The key assumption in the analysis of judicial decisions with a regression model, as in Equation (1), is that any left-out, unobserved factor, captured by the error term, is not correlated with the variable *Minority Defendant*. For example, if defendant income is not included in regression Equation (1) (if income is not part of the vector  $X$ ), this implies that income is embedded in the error term ( $\varepsilon$ ). To the extent that minority status and income are correlated, the error term  $\varepsilon$  (which now captures income) and the variable *Minority Defendant* will be correlated. Thus, any variation in incomes between defendants will translate into variation in the error terms and therefore in the case outcomes between the defendants, and such income variation will also impact the key explanatory variable *Minority Defendant*. In this situation it is impossible to obtain an unbiased estimate of the parameter  $\beta_1$  to determine the true impact of the defendant's minority status on the outcome of interest. That is, it would not be possible to disentangle the true "causal" impact of minority status from the impact of income on the outcome.

One complication in these analyses is the potential of systematic "sorting" of defendants to judges. For example, assume that there are two types of judges in the same courthouse: Type-L and Type-H. Type-L judges are lenient—they convict defendants at lower rates and assign shorter prison terms. Type-H judges are harsh by nature—they convict defendants at higher rates and assign longer sentences. Now assume that most minority defendants are channeled to judges of Type-H, and that most White defendants are put in front of Type-L judges. Even if judges were not racially biased, the end result of this sorting would be that minority defendants would end up with higher conviction probabilities and longer sentences in comparison to White defendants. One obvious method to overcome the confounding that is generated by this type of sorting is to employ data from jurisdictions that use random assignment of cases to judges. Most of the researches cited below utilize data obtained from judicial units where case files are randomly assigned to judges.<sup>3</sup>

There are still other, even more subtle, complications in the analysis of judicial decisions that may confound the estimated impact of race, ethnicity, and gender on judicial outcomes. For example, assume that prosecutors

are more aggressive inside the courtroom during trials involving minority defendants, and they are milder and more accommodating against White defendants. If this is the case, the conviction and sentencing decisions of a judge may resemble favoritism toward White defendants, but this would be misleading. The disparate outcomes would be reflection of prosecutors' actions in the courtroom. This particular potential confounding can be controlled for by including in the model (Equation 1) indicator variables for individual prosecutors who were involved in each case (prosecutor fixed effects), although in most publicly available data sets such information about prosecutors is not recorded.

There exist different methods and statistical procedures that can be utilized to attempt to control for observable as well as unobservable attributes of the case and the defendant, and to account for the issue of selection, when relevant. In doing so, the goal is to create a "counterfactual" scenario to a judicial decision in an effort to mimic a research design where the race of the defendant is altered while all else that matters remains the same. In this chapter, I will not get into the statistical properties and relative merits of various procedures and research designs that have been employed in the analyses of judicial decision-making. Suffice it to say that there is no study that relies on perfect counterfactuals to identify the causal impact of the attributes of key players (e.g., the race of defendant, the gender of the judge) in judicial outcomes. Nevertheless, the studies cited in this article employ credible statistical procedures that arguably make causal inferences regarding the impact of race and gender on judicial decisions. Thus, instead of research methodologies, I will focus on what we have learned so far from the body of existing credible research regarding bias in judicial decision-making. Instead of summarizing an exhaustive list of studies, I will present a small but representative group of studies that shed light on various sub-domains of the terrain of "judicial discrimination." I will focus on research that analyzed the judicial system in the United States, although I will discuss some important papers that utilized data from other countries. In taking this approach, I have organized the findings into the three groups.

### **1: Defendants' Race and Gender and Judges' Attributes Impact Judicial Outcomes**

There is substantial research demonstrating that defendants' race influences case outcomes. In the notation of Equation (1), this research, using a variety of data sets from different settings, has identified that the coefficient  $\beta_1$  is positive and statistically different from zero when the dependent variable  $Y$  signifies a negative outcome from the defendant's viewpoint. For example, it has been shown that race and gender of the death row inmates matter

regarding which defendants gets their sentence commuted to a life term, and which ones get executed (Argys & Mocan, 2004). Arnold, Dobbie, and Yang (2018) found that bail judges are biased against Black defendants. Rehavi and Starr (2014) reported that Blacks received longer sentences than Whites, all else being equal. Mustard (2001) analyzed federal offenders and found that Black and male defendants received longer sentences. Starr (2015) reported substantial gender disparities in sentence lengths in favor of females in federal criminal cases. Schanzenbach (2005) analyzed case outcomes of offenders sentenced under federal sentencing guidelines, finding that all else equal female defendants received shorter sentences, while the opposite was true for Blacks and Hispanics.

These are important findings as they demonstrate that justice is not blind, and that the race and gender of the defendant has arguably a causal impact on the resolution of cases. This implies that two otherwise identical defendants would end up with different judgments if one was White and the other a racial minority (or if one was male and the other female). However, disparate treatment of defendants because of their race and gender does not portray the whole picture, which is quite a bit more complicated and nuanced.

Specifically, studies that analyze judicial decisions as a function of case characteristics and defendant attributes implicitly assume that judges are homogenous in their treatment of race. This is, of course, a strong assumption, and it is shown not to be true. For example, Abrams, Bertrand, and Mullainathan (2012) analyzed case files from the Circuit Court of Cook County in Illinois. Exploiting random assignment of cases to judges, they detected significant variation between judges in how they sentence minority defendants. The authors found that there was substantial inter-judge variation in incarceration rates, but not in sentence lengths. For example when moving from a judge at the 10th percentile to a judge at the 90th percentile in the racial gap distribution, the gap in incarceration rates between Black and White defendants increases by 18 percentage points (where the mean incarceration rates of Black and White defendants are 38 percent and 51 percent, respectively). This indicates that some judges systematically convict Black defendants at higher rates, and some other judges do the opposite. Thus, substantially different probabilities of incarceration emerge for a Black defendant, depending on the judge to whom the case is assigned.

Judges differ on a number of dimensions, including age, experience, race, gender, and political affiliation, among others. Although legal formalism suggests that the facts of a case, the fundamental principles of judicial decision-making, and the letter of law should determine judicial outcomes, it has long been argued by legal realists that attributes of judges, ranging from life experiences to ideological beliefs, can have an impact on judicial decisions. Leading legal realist Judge Jerome Frank, who served on the U.S. Court of

Appeals for the Second Circuit, has even argued that a judge's decision may be impacted by mundane things, including what he/she ate for breakfast (see Frank, 1973).

A line of research has shown that differences in personal attributes of judges can translate into how judges make decisions and how they treat different groups of defendants. The general framework can be depicted by Equation (2):

$$Y_{djc} = \alpha_2 + \beta_2 \text{ Minority Defendant}_d + \gamma_2 \text{ Rp Judge}_j + \tau_2 \text{ Minority Defendant}_d \times \text{ Rp Judge}_j + \mathbf{X}_d \Phi + \lambda_c + \varepsilon_{djc} \quad (2)$$

Equation (2) is similar to Equation (1), but there is an additional variable—Rp Judge—which represents an attribute of the judge. In this example, Rp Judge is an indicator variable that takes the value of one if the judge is a Republican-appointed federal judge and zero otherwise. Equation (2) also contains an interaction term between Minority Defendant and Rp Judge. This formulation allows for the identification of more detailed impacts: the incremental impact of Republican-appointed judges on outcomes of nonminority defendants is  $\gamma_2$ . The incremental impact of Republican-appointed judges on outcomes of minority defendants is  $(\gamma_2 + \tau_2)$ . Thus,  $\tau_2$  represents the impact of Republican-appointed judges on outcomes of minority defendants in comparison to nonminority defendants.<sup>4</sup>

Using this framework, Cohen and Yang (2019) find that political affiliation of judges has an impact on their decisions. The authors find that Republican-appointed federal judges are harsher against Black defendants and are more lenient toward female defendants in comparison to judges appointed by Democrats. Similarly, analyzing data from the U.S. Courts of Appeals, Sunstein and colleagues (2006) showed that judges who were Democratic appointees voted—in political elections—in a more liberal fashion in comparison to judges who were appointed by Republicans. Glynn and Sen (2015) found that appellate court judges who have daughters consistently voted in a feminist direction. Knepper (2018) found that female plaintiffs who filed sex discrimination cases with the Equal Employment Opportunity Commission were more likely to win their cases in federal district courts if a female judge was assigned to the case.

These findings underscore that race and gender of the defendant, and race, gender, and political affiliation of the judge, impact case outcomes. Such characteristics of defendants and judges should not matter if justice is to be blind and if cases are to be decided on their merits alone. One important point that needs to be highlighted is that these estimated effects reveal unfairness in the treatment of two (groups of) defendants who have otherwise similar

cases, but they do not reveal whether one group is treated harshly or the other group is treated favorably. For example, the result that Black defendants receive longer sentences than Whites may imply that Whites were assigned sentences of “optimal” length but Blacks received sentences that were longer than optimal. Alternatively, it could imply that Blacks were assigned sentences that were correct, but that Whites were treated leniently and received shorter than optimal-length sentences. This point notwithstanding, the results discussed in this section indicate unfair treatment of one group based on factors that should not be determinants of case outcomes.

## **2: Judicial Decisions Are Impacted by Factors Unrelated to Case Merits**

The previous section showed that demographic characteristics of both judges and defendants play a role in judicial decisions. In this section, I present evidence suggesting that additional factors unrelated to the essentials of a case (i.e., unrelated to case merits) also play a role. One such factor is the emotions of judge. It has been shown that emotions, such as anger and sadness, impact people’s judgment and decisions (Bodenhausen, Sheppard, & Kramer, 1994, Keltner, Ellsworth, & Edwards, 1993). Because judicial decisions are supposed to be made with careful deliberation and because they are consequential, it would be reasonable to assume that judges’ decisions would not be influenced by exogenous events that might trigger negative emotions. Eren and Mocan (2018), however, showed that if the sentencing date in juvenile court cases in Louisiana happened to be during the week following a surprise loss of the Louisiana State University (LSU) football team, defendants received longer sentences—and this impact was driven by judges who received their undergraduate degrees from LSU. Importantly, the authors showed that, disregarding the impact of football games, there was no evidence of discrimination (in the sense of Equation (1)): the race of the defendant had no impact on judicial decisions after controlling for relevant case and defendant attributes. However, after a surprising LSU loss, racial bias emerged: the impact on sentence length following an upset loss was borne by Black defendants, suggesting an implicit bias against Blacks that emerged only when the judge had suffered an emotional shock.

Danziger, Levav, and Avnaim-Pesso (2011) reported another type of cognitive bias in judicial decisions. They analyzed data of parole decisions in Israel and found that the probability of a favorable parole decision declined significantly as judges adjudicated the cases sequentially, but that the propensity to grant parole increased after the judge took a break for food and tea. In other words, the more cases judges saw in a row, the worse the outcome for the defendant; and, if the judge took a break for food, the next set of



defendants he or she saw received favorable treatment. This suggests that the time of the day when the judge hears a case has an impact on the outcome, as the irritability of judges varies over the course of the day (and in somewhat predictable fashion). Chen and Philippe (2019) found that, in France, if the date on which the judge made his/her decision happened to be the birthday of the defendant, the sentence was reduced by 3 percent. In the United States, when the decision date and the defendant's birthday were the same, defendants received federal sentences that were 33 percent shorter.

These results are significant because they document that judicial decisions can be arbitrary and capricious, depending on the mood of the judge, or depending on whether the judge feels generous toward a defendant for a reason that is irrelevant to the case. The added layer that emerges in the paper by Eren and Mocan (2018) is the existence of implicit bias, which surfaces when the judge is exposed to an emotional trauma. The potential good news here is that making this information available to judges could alter their behavior. Specifically, if judges recognize that their propensity to punish increases following a negative emotional shock or when they get tired and hungry, they may take this information into account in their subsequent adjudications and may adjust their instinct-based decisions. Similarly, if judges are made aware that "rewarding" a defendant based on his/her birthday or based on some other custom or celebration (e.g., a trial coinciding with a national or religious holiday) generates arbitrary disparities in sentencing between defendants, they may correct this behavior.

### 3: Judges Show Significant In-Group Bias

As discussed above, there exist evidence of racial and gender bias in judicial decisions. Case outcomes are influenced by defendants' race and gender. There is also substantial evidence indicating that race and gender and other attributes of judges matter in these decisions. Thus, an important question is whether judges treat defendants "who are like them" differently in comparison to defendants who are not like them. That is, do judges treat members of their "group" differently, and does race-matching between judges and defendants matter for case resolution? To address this question, consider Equation (3):

$$Y_{djc} = \alpha_3 + \beta_3 \text{Minority Defendant}_d + \gamma_3 \text{Minority Judge}_j + \delta_3 \text{Minority Defendant}_d * \text{Minority Judge}_j + \mathbf{X}_d \Theta + \mu_c + \omega_{djc} \quad (3)$$

where the notation is the same as before, and *Minority Judge* takes the value of one if the judge is a racial minority and zero otherwise.

In Equation (3),  $\gamma_3$  captures the difference in outcomes between White defendant-minority judge pairs and White defendant-White judge pairs, and  $\gamma_3 + \delta_3$  is the differential impact on outcomes on minority defendants by minority versus White judges. Thus,  $\delta_3$  in Equation (3) signifies the differential decisions of minority judges versus White judges in their treatment of minority defendants over White defendants. Put plainly, if  $\delta_3$  is different from zero, this is evidence for in-group bias. For example, if  $Y$  stands for sentence length and if  $\delta_3$  is negative, this indicates that judges assign lower sentences to those defendants who are in their "in-group"; for example, those who are of the same race. Thus, the formulation in Equation (3) allows for an investigation of whether the race of judges makes a difference in how they treat defendants who share the same race.

Despite its importance, there is very limited research on in-group bias in judicial decisions. Potential reasons for this include the difficulty of finding suitable data sets. To do such analyses, one needs to have access to data with sensitive information about cases, including the race/ethnicity of both the defendant and the judge. Furthermore, there needs to be enough variation in the race of both defendants and judges for the analysis sample to be able to statistically identify an effect, if it exists. This means that such analysis cannot be performed in situations where either minority defendants or minority judges constitute a small fraction of the cases in the analysis sample. In other words, one needs a detailed data set, a large data set, and a data set that is reasonably well-balanced across these different attributes.

A trio of studies considered the effect of in-group race/ethnicity on outcomes in Israeli courts. Shayo and Zussman (2011) analyzed data obtained from claims courts in Israel, where cases are randomly assigned to Arab or Israeli judges. The authors showed that a claim was more likely to be accepted if the judge was of the same race/ethnicity (Arab or Israeli) as the plaintiff, indicating positive in-group bias. They also found that this effect was associated with the intensity of terrorist events around the location of the court preceding the date of the judicial decision. Gazal-Ayal and Sulitzeanu-Kenan (2010) also reported positive in-group bias in Israeli Arab and Jewish judges' decisions on criminal cases. Grossman et al. (2016) analyzed panel decisions of appellate court judges, instead of solo bench decisions of individual judges as in Shayo and Zussman (2011), and found that appeal outcomes of Jewish defendants were independent of the ethnic composition of the panel of judges, but that Arab defendants received more lenient punishment when the panel included at least one Arab judge.

Depew, Eren, and Mocan (2017) analyzed the decisions of juvenile court judges in Louisiana and found negative in-group bias. If juvenile defendants were assigned to a judge of his/her own race, they were more likely to be placed on custody and received longer sentences. As explained by the

authors, this finding could be driven by a number of alternative mechanisms. For example, it could be the result of "tough love." Judges may believe that future criminal activity of a juvenile, who is of the same race as the judge, can be prevented by being tough on that juvenile. That is, judges may be imposing stiffer punishment now to reduce recidivism later. Alternatively, judges may be concerned about generating an impression of being biased against defendants who are of another race. Thus, they may mete out the "correct" amount of punishment to defendants who are in their in-group, but they may treat defendants who are of another race more leniently.<sup>5</sup>

It is important to note that in Equation (3),  $(\alpha_3 + \beta_3)$  represents the impact of White judges on minority defendants and  $\alpha_3$  stands for the impact of White judges on White defendants. Thus,  $\beta_3$  is the magnitude of the differential treatment of minority defendants vis-à-vis White defendants by White judges. Similarly,  $(\beta_3 + \delta_3)$  stands for the differential treatment of minority defendants by minority judges. If we let  $\pi$  represent the proportion of White judges in the analysis sample, then  $(1 - \pi)$  stands for the proportion of minority judges. This means that the *overall* differential treatment of minority defendants by all judges (White and minority, which is  $\beta_1$  of Equation 1) can be calculated as  $(\pi \beta_3) + (1 - \pi)(\beta_3 + \delta_3)$  which is equal to  $[\beta_3 + \delta_3 \pi - \delta_3] = [\beta_3 + (1 - \pi) \delta_3]$ . Thus, there is a direct link between the overall racial bias that can be estimated using a standard workhorse model (Equation 1) and the in-group bias model (Equation 3).

This is significant because it demonstrates that the effects of overt racial bias ( $\beta_1$  in Equation 1) can be recovered from the coefficients of the in-group bias regression (Equation 3). To make this clear, assume that  $Y$  stands for sentence length, measured in months, and assume that  $\delta_3$  is estimated as -6 from Equation (3). This means that there is in-group bias in sentencing and that defendants receive sentences that are six months shorter if they are matched by a judge of their own race. This important finding can be broken down into its components.

Assume that estimating Equation (3) also revealed that  $\beta_3$  is 2. This means that *White judges* assign two months *longer* sentences to minority defendants vis-à-vis White defendants. Put differently, White judges exhibit positive in-group bias by assigning White defendants two months *shorter* sentences vis-à-vis minority defendants. Note that, in this situation,  $(\beta_3 + \delta_3)$ , which signifies how *minority defendants are treated* vis-à-vis White defendants by *minority judges*, is equal to -4. In turn, this means that minority defendants receive sentences that are four months shorter when they are matched with a minority judge in comparison to the situation where they would be matched with a White judge. In summary, both White and minority judges favor their in-group by assigning shorter sentences. Now assume that in 2/3 of the judges in this analysis are White and the other 1/3 are minority; that is,  $\pi = 2/3$  and

$(1-\pi) = 1/3$ . Thus,  $\beta_1$ , which can be calculated as  $[\beta_3 + (1-\pi)\delta_3]$ , is equal to  $[2 + (1/3)(-6)] = 0$ , and we know that 0 indicates no racial bias in sentencing.

The intuition is straightforward. A researcher who has access to a complete data set with all relevant variables that are needed to investigate racial bias can run Equation (1), and he/she may estimate the coefficient of *Minority Defendant* as being equal to zero, indicating no racial bias in sentencing. This inference, although correct at the aggregate, could mask more nuanced and layered discriminatory behavior. Specifically, this “no evidence for discrimination” would be an artifact of how White judges discriminate against minorities *and* minority judges discriminate against Whites, *and* the magnitudes of these effects and the proportions of White and minority judges in the data are such that they could cancel out at the aggregate.

To put it differently, in a judicial system in which judges favor “their own type” and where there are sufficient number of cases of each type handled by both types of judges, no racial bias effect will be detected in the data (using an analysis in the spirit of Equation 1) if in-group bias effect is similar for both types of judges. This finding would fail to reject the hypothesis of no racial discrimination, and therefore would incorrectly report “no evidence of racial discrimination.”

## CONCLUSION

Although “Equal Justice under Law” is etched on the façade of the U.S. Supreme Courthouse, and justice is expected to be blind, it has been shown that judicial decisions are correlated with such personal attributes as defendants’ race and gender. Because judicial decisions are made by judges in most cases (e.g., in bench trials in federal and districts courts or by a panel of judges in appellate courts) the focus has been on the verdicts handed down by judges. While lay observers attribute the race and gender disparities in judges’ decisions to discrimination, these differences do not necessarily indicate disparate treatment. This is because a number of defendants’ attributes may be correlated with race and gender, and these attributes may directly or indirectly impact case outcomes. Examples of these attributes include observable factors such as socioeconomic circumstances, as well as difficult-to-observe characteristics such as criminal proclivity. Thus, detailed characteristics of the defendants and those of their case files need to be accounted for in order to tease out their impact on judicial decisions. Only then is it possible to statistically estimate the true impact of race and gender on judicial outcomes.

There are other important issues that complicate the investigation of judicial bias. First, there are many actors in the criminal justice system (CJS) other than judges, including prosecutors and defense attorneys, and the actions of these actors can impact the decisions of judges. Put differently, even if judges

are completely unbiased with respect to race and gender, their decisions may resemble bias as an artifact of these other actors' decisions and actions that took place before the case arrived at the docket of the judge. Furthermore, the decisions of these actors are interrelated. For example, if a particular judge is known to be racially biased against a particular group of defendants, defense attorneys may be more likely to plea bargain with prosecutors to avoid the case being adjudicated by the judge. In this scenario, the composition of the cases that are brought in front of this judge, as well as their case characteristics, would be different in comparison to counterfactual scenarios in which the judge was not racially unbiased. This type systematic "sorting" is produced by the defendant (or his/her defense attorney). Systematic sorting can also be generated by other non-judge actors. For example, prosecutors have substantial discretion regarding whether to press charges and what type of charges to press against a defendant. If this decision is racially biased, the case files that end up in front of judges will have gone through sorting based on the race or gender of the defendants. This would confound the analysis of judges' decisions, if sorting is not adjusted for.

Researchers over the last decade have increasingly recognized these issues, and by implementing appropriate statistical procedures on comprehensive data sets, they tried to address these complications. The research studies summarized in this chapter are based on credible empirical designs, and they address many of these statistical challenges.

Research shows that defendants' race and gender impact case outcomes, holding constant all factors that should matter for the adjudication of the case. This means, for example, that two otherwise identical defendants, charged for the same crime with the same mitigating and aggregating circumstances, face different probabilities of conviction and receive different sentences based on their race and gender.

This is a troublesome result, and it is understandable that objections can be raised. For instance, some might argue that it is impossible for researchers to control for every single factor that can matter in a judge's decision; there could be additional features of the case that can be observed by the judge but unobserved by researchers, and these omitted variables could be driving the findings. This objection, while reasonable, does not have the power it used to have thirty years ago when empirical analyses were much less sophisticated and big and detailed data sets were not available. This "omitted variable bias" argument has less credibility now because researchers apply increasingly sophisticated statistical techniques to large data sets which include rich information on the case: most, if not all, of the information that is available to the judge.<sup>6</sup>

Research has also shown that attributes of judges (i.e., race, gender, and political affiliation) have an impact on how cases are adjudicated. This is also a troublesome result, although perhaps more "understandable." After all, it

has been acknowledged that judges bring to the bench their own background and life experiences. That judges would see the world differently on account of their different backgrounds/religions/political ideologies is understandable, but evaluation of the merits of cases through such individualized lenses violates the principle of fairness for the defendant. The issue is even more salient when one recognizes that the U.S. judicial system has a structure where some judges are elected to the bench. If elected judges are concerned about representing the preferences of voters, and if re-election is a concern, it is easy to see how judicial decisions can stray from the letter of the law.<sup>7</sup>

That judges' own personal attributes (e.g., being a female judge or being a minority judge) impact case outcomes has been promoted by some analysts as a tool to counteract the discriminatory behavior of the courts. Specifically, it has been argued that increasing racial and gender diversity on the courts could be a mechanism by which different perspectives could be introduced into the judicial decision-making process (Ifill, 2000).

As discussed in this chapter, there is credible evidence that if judges and defendants share the same racial or ethnic identity, this will have an impact on case outcomes. This also means that this "in-group bias" may mask the overall racial bias. In other words, the analysis of judicial decisions may reveal no bias, but this "nil result" could be an artifact of judges of different racial groups treating the members of their in-group favorably and these subgroup biases canceling each other in the aggregate. Finally, research shows that judicial decisions are impacted by random events that may trigger emotional shocks or irritability for judges, which in turn produce arbitrary and capricious implementation of punishment.

What are the solutions to these problems? The researches summarized above imply two clear methods to minimize the extent of the biases detected in judicial decision-making. First, the results of research on judicial decision-making can be made available to judges, the decisions of whom were analyzed in those studies. Recognizing these biases can be a vehicle through which these biases can be reduced. This is especially true for implicit biases that are shown to emerge under stressful events faced by judges and decisions made under conditions that create irritability and discomfort for judges.

Second, given the research results showing that attributes of judges as well as in-group bias impact their decisions, an obvious policy implication is to avoid solo bench decisions. Instead, decisions can be made by a panel of judges, where the panel consists of at least two judges and the racial and gender composition of the panel is heterogeneous. This, of course, is an expensive proposition, and it would at least double the required number of judges at lower courts. Nevertheless, it could be implemented selectively in those jurisdictions that reveal the most significant racial bias or in-group bias.

Some scholars argue that biases in the judicial system are so deep-rooted that little, if anything, can be done to eliminate them. These scholars further

argue that biases can be eradicated and fairness in the application of the law and equality between defendants can be achieved by utilizing machine learning techniques and by the use of artificial intelligence-based (AI) algorithms to make judicial decisions. Some jurisdictions have started to utilize these methods, and future research will reveal the effectiveness and costs and benefits of removing personal judgment from judicial decision-making.

## NOTES

1. Differential arrest rates between racial groups are a function of differential criminal participation and the arrest propensity of police conditional on criminal activity.

2. That there is more research on judges is partly a reflection of availability of data in enough detail to allow for meaningful analysis.

3. The requirement of the judicial authority that case files be randomly assigned to judges does not ensure that random assignment is implemented in practice. There are, however, simple statistical tests to verify random assignment. For example, if case files are randomly assigned to judges, the attributes of the cases and those of judges should be uncorrelated. For an example, see Eren and Mocan (2019).

4. Because the status of being a Republican or Democrat-appointed judge does not change for any judge, Equation (2) does not include judge fixed effects.

5. For other potential explanations, see Depew, Eren, and Mocan (2017).

6. While there is clear evidence that defendants' race has an impact on case outcomes, it is unclear whether the deep root of this is racial animus or statistical discrimination. This is a difficult question to answer, and how we answer it may have policy implications. That said, from the perspective of defendants who face differential treatment on account of race, the root cause does not matter.

7. Consider the following example. Assume three things: one, that the judge is elected from a mostly White district; two, that recidivism rates (i.e., rates of reoffending after leaving prison) are higher among convicted Black criminals; and, three, that voters are concerned about crime. In this situation, an elected judge who "represents voter preferences" is expected to assign longer prison sentences to Black defendants both to incapacitate them (in prison) and to deter them from recidivism (with stiffer punishment). This, of course, is a violation of fair and equal treatment, but it could be the optimal decision for a judge wanting to be re-elected. In this situation, it is the voters who de facto determine how the law should be interpreted.

## REFERENCES

- Abrams, D. S., Bertrand, M., & Mullainathan, S. (2012). Do Judges Vary in Their Treatment of Race? *The Journal of Legal Studies*, 41(2), 347–383.
- Agan, A., Freedman, M., & Owens, E. (2018). Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense. *NBER Working Paper*, No. 24579.

- Alesina, A., & Ferrara, E. L. (2014). A Test of Racial Bias in Capital Sentencing. *The American Economic Review*, 104(11), 3397–3433.
- Anderson, J. M., & Heaton, P. (2012). How Much Difference Does the Lawyer Make? The Effect of Defense Counsel on Murder Case Outcomes. *Yale Law Journal*, 122(154), 154–217.
- Anwar, S., Bayer, P., & Hjalmarsson, R. (2012). The Impact of Jury Race in Criminal Trials. *The Quarterly Journal of Economics*, 127(2), 1017–1055.
- Anwar, S., Bayer, P., & Hjalmarsson, R. (2014). The Role of Age in Jury Selection and Trial Outcomes. *The Journal of Law and Economics*, 57(4), 1001–1030.
- Argys, L. M., & Mocan, N. (2004). Who Shall Live and Who Shall Die? An Analysis of Prisoners on Death Row in the United States. *The Journal of Legal Studies*, 33(2), 255–282.
- Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial Bias in Bail Decisions. *The Quarterly Journal of Economics*, 133(4), 1885–1932.
- Bodenhausen, G. V., Sheppard, L. A., & Kramer, G. P. (1994). Negative Affect and Social Judgment: The Differential Impact of Anger and Sadness. *European Journal of Social Psychology*, 24(1), 45–62.
- Boyd, C. L., Epstein, L., & Martin, A. D. (2010). Untangling the Causal Effects of Sex on Judging. *American Journal of Political Science*, 54, 389–411.
- Carson, E. A. (2015). Prisoners in 2014. *U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics*. Retrieved from [www.bjs.gov](http://www.bjs.gov).
- Chen, D. L., & Philippe, A. (2019). Clash of Norms: Judicial Leniency on Defendant Birthdays. *Working Paper*. Retrieved from SSRN: <http://dx.doi.org/10.2139/ssrn.3203624>.
- Cohen, A., & Yang, C. S. (2019). Judicial Politics and Sentencing Decisions. *American Economic Journal: Economic Policy*, 11(1), 160–191.
- Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous Factors in Judicial Decisions. *Proceedings of the National Academy of Sciences*, 108(17), 6889–6892.
- Depew, B., Eren, O., & Mocan, N. (2017). Judges, juveniles, and in-group bias. *The Journal of Law and Economics*, 60(2), 209–239.
- Eren, O., & Mocan, N. (2018). Emotional Judges and Unlucky Juveniles. *American Economic Journal: Applied Economics*, 10(3), 171–205.
- Federal Bureau of Investigations. (2018). Crime in the United States, 2017. Federal Bureau of Investigations, Uniform Crime Statistics. Retrieved from <https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017/topic-pages/tables/table-43>.
- Flanagan, F. X. (2018). Race, Gender, and Juries: Evidence from North Carolina. *The Journal of Law and Economics*, 61(2), 189–214.
- Frank, J. (1973). *Courts on Trial: Myth and Reality in American Justice*. Princeton, NJ: Princeton University Press.
- Gazal-Ayal, O., & Sulitzeanu-Kenan, R. (2010). Let My People Go: Ethnic In-Group Bias in Judicial Decisions- Evidence from a Randomized Natural Experiment. *Journal of Empirical Legal Studies*, 7(3), 403–428.
- Glaeser, E. L., & Sacerdote, B. (2003). Sentencing in Homicide Cases and the Role of Vengeance. *The Journal of Legal Studies*, 32(2), 363–382.



- Glynn, A., & Sen, M. (2015). Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Women's Issues? *American Journal of Political Science*, 59(1), 37–54.
- Grossman, G., Gazal-Ayal, O., Pimentel, S., & Weinstein, J. (2016). Descriptive Representation and Judicial Outcomes in Multiethnic Societies. *American Journal of Political Science*, 60(1), 44–69.
- Hèyes, A., & Saberian, S. (2019). Temperature and Decisions: Evidence from 207,000 Court Cases. *American Economic Journal: Applied Economics*, 11(2), 238–265.
- Ifill, S. (2000) Racial Diversity on the Bench: Beyond Role Models and Public Confidence. *Washington and Lee Law Review*, 96, 863–876.
- Kastellec, J. P. (2013). Racial Diversity and Judicial Influence on Appellate Courts. *American Journal of Political Science*, 57, 167–183.
- Keltner, D., Ellsworth, P. C., & Edwards, K. (1993). Beyond simple pessimism: Effects of sadness and anger on social perception. *Journal of Personality and Social Psychology*, 64(5), 740–752.
- Knepper, M. (2018). When the Shadow Is the Substance: Judge Gender and the Outcomes of Workplace Sex Discrimination Cases. *Journal of Labor Economics*, 36(3), 623–664.
- Mustard, D. B. (2001). Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. *Federal Courts*. *Journal of Law and Economics*, 44(1), 285–314.
- Philippe, A., & Ouss, A. (2018). No Hatred or Malice, Fear or Affection: Media and Sentencing. *Journal of Political Economy*, 126(5), 2134–2178.
- Rehavi, M. M., & Starr, S. B. (2014). Racial Disparity in Federal Criminal Sentences. *Journal of Political Economy*, 122(6), 1320–1354.
- Schanzenbach, M. (2005). Racial and Sex Disparities in Prison Sentences: The Effect of District-Level Judicial Demographics. *Journal of Legal Studies*, 34, 57–92.
- Starr, S. B. (2015). Estimating Gender Disparities in Federal Criminal Cases. *American Law and Economics Review*, 17(1), 127–159.
- Statistica Research Department. (2019). U.S. Capital Punishment: Racial Distribution of Death Row Inmates 2019. *Statistica*. Retrieved from <https://www.statista.com>.
- Sunstein, C. R., Schkade, D., Ellman, L. M., & Sawicki, A. (2006). *Are Judge Political?* Washington, DC: Brookings Institution Press.
- United States Census Bureau. (2018). Population Estimates. Retrieved from <https://www.census.gov/quickfacts/fact/table/US/PST045218>.
- United States Department of Commerce, Census Bureau. (2015). Quick Facts: United States. Retrieved from [www.census.gov](http://www.census.gov).
- U.S. Department of Commerce, Census Bureau. (2018). Income and Poverty in the United States: 2017. Current Population Reports. Retrieved from <https://www.census.gov/content/dam/Census/library/publications/2018/demo/p60-263.pdf>.
- Wagner, P. (2012, August 28). Incarceration is Not an Equal Opportunity Punishment. *Prison Policy Initiative*. Retrieved from [www.prisonpolicy.org](http://www.prisonpolicy.org).
- Yang, C. S. (2015). Free at Last? Judicial Discretion and Racial Disparities in Federal Sentencing. *The Journal of Legal Studies*, 44(1), 75–111.