RACIAL BIAS IN BAIL DECISIONS*

David Arnold
Will Dobbie
Crystal S. Yang

May 2018

Abstract

This paper develops a new test for identifying racial bias in the context of bail decisions—a high-stakes setting with large disparities between white and black defendants. We motivate our analysis using Becker's model of racial bias, which predicts that rates of pretrial misconduct will be identical for marginal white and marginal black defendants if bail judges are racially unbiased. In contrast, marginal white defendants will have higher rates of misconduct than marginal black defendants if bail judges are racially biased, whether that bias is driven by racial animus, inaccurate racial stereotypes, or any other form of bias. To test the model, we use the release tendencies of quasi-randomly assigned bail judges to identify the relevant race-specific misconduct rates. Estimates from Miami and Philadelphia show that bail judges are racially biased against black defendants, with substantially more racial bias among both inexperienced and part-time judges. We find suggestive evidence that this racial bias is driven by bail judges relying on inaccurate stereotypes that exaggerate the relative danger of releasing black defendants. *JEL Codes:* C26, J15.

*We gratefully acknowledge the coeditors Lawrence Katz and Andrei Shleifer, and five anonymous referees for many valuable insights and suggestions. We also thank Josh Angrist, David Autor, Pedro Bordalo, Leah Platt Boustan, David Deming, Hanming Fang, Hank Farber, Roland Fryer, Jonah Gelbach, Nicola Gennaioli, Edward Glaeser, Paul Goldsmith-Pinkham, Christine Jolls, Louis Kaplow, Michal Kolesár, Amanda Kowalski, Ilyana Kuziemko, Magne Mogstad, Nicola Persico, Steven Shavell, David Silver, Alex Torgovitsky, and numerous seminar participants for helpful comments and suggestions. Molly Bunke, Kevin DeLuca, Nicole Gandre, James Reeves, and Amy Wickett provided excellent research assistance.

Racial disparities exist at every stage of the U.S. criminal justice system. Compared to observably similar whites, blacks are more likely to be searched for contraband (Antonovics and Knight 2009), more likely to experience police force (Fryer 2016), more likely to be charged with a serious offense (Rehavi and Starr 2014), more likely to be convicted (Anwar, Bayer, and Hjalmarrson 2012), and more likely to be incarcerated (Abrams, Bertrand, and Mullainathan 2012). Racial disparities are particularly prominent in the setting of bail: in our data, black defendants are 3.6 percentage points more likely to be assigned monetary bail than white defendants and, conditional on being assigned monetary bail, receive bail amounts that are $9,923 greater.[1] One view is that these racial disparities are driven by statistical discrimination, or the use of observable group traits such as race to form accurate beliefs about the unobservable characteristics of defendants (e.g., Phelps 1972; Arrow 1973). A second view is that statistical discrimination alone cannot explain these disparities, leaving a role for various forms of racial bias, such as racial animus (e.g., Becker 1957) or inaccurate racial stereotypes (e.g., Bordalo et al. 2016). However, distinguishing between these two contrasting explanations remains an empirical challenge.

To test whether racial bias is empirically relevant, Becker (1957, 1993) proposed an "outcome test" that compares the success or failure of decisions across groups at the margin. In our setting, the outcome test is based on the idea that rates of pretrial misconduct will be identical for marginal white and marginal black defendants if bail judges are racially unbiased and the disparities in bail setting are solely due to accurate statistical discrimination. In contrast, marginal white defendants will have higher rates of pretrial misconduct than marginal black defendants if these bail judges are racially biased against blacks, whether that racial bias is driven by racial animus, inaccurate racial stereotypes, or any other form of racial bias. The outcome test has been difficult to implement in practice, however, as comparisons based on average defendant outcomes are biased when whites and blacks have different risk distributions—the well-known infra-marginality problem (e.g., Ayres 2002).

In recent years, two seminal papers have developed outcome tests of racial bias that partially circumvent this infra-marginality problem. In the first paper, Knowles, Persico, and Todd (2001) show that if motorists respond to the race-specific probability of being searched, then all motorists of a given race will carry contraband with equal probability. As a result, the marginal and average success rates of police searches will be identical and OLS estimates are not biased by infra-marginality concerns. Knowles et al. (2001) find no difference in the average success rate of police searches for white and black drivers, leading them to conclude that there is no racial bias in police searches. In a second important paper, Anwar and Fang (2006) develop a test of relative racial bias based on the idea that the ranking of search and success rates by white and black police officers should be unaffected by the race of the motorist even when there are infra-marginality problems. Consistent with Knowles et al. (2001), Anwar and Fang (2006) find no evidence of relative racial bias in police

[1] Authors' calculation for Miami-Dade and Philadelphia using the data described in Section II. Racial disparities in bail setting are also observed in other jurisdictions. For example, black felony defendants in state courts are nine percentage points more likely to be detained pretrial compared to otherwise similar white defendants (McIntyre and Baradaran 2013).

searches, but note that their approach cannot be used to detect absolute racial bias.[2] However, the prior literature has been critiqued for its reliance on restrictive assumptions about the unobserved risk of blacks and whites (e.g., Brock et al. 2012).

In this paper, we propose a new outcome test for identifying racial bias in the context of bail decisions. Bail is an ideal setting to test for racial bias for a number of reasons. First, the legal objective of bail judges is narrow, straightforward, and measurable: to set bail conditions that allow most defendants to be released while minimizing the risk of pretrial misconduct. In contrast, the objectives of judges at other stages of the criminal justice process, such as sentencing, are complicated by multiple hard-to-measure objectives, such as the balance between retribution and mercy. Second, mostly untrained bail judges must make on-the-spot judgments with limited information and little to no interaction with defendants. These institutional features make bail decisions particularly prone to the kind of inaccurate stereotypes or categorical heuristics that exacerbate racial bias (e.g., Fryer and Jackson 2008; Bordalo et al. 2016). Finally, bail decisions are extremely consequential for both white and black defendants, with prior work suggesting that detained defendants suffer about $30,000 in lost earnings and government benefits alone (Dobbie, Goldin, and Yang 2018).[3]

In the first section of the paper, we formally develop two complementary estimators that use variation in the release tendencies of quasi-randomly assigned bail judges to identify the differences in pretrial misconduct rates at the margin of release required for the Becker outcome test. Our first estimator uses the standard instrumental variables (IV) framework to identify differences in the local average treatment effects (LATEs) for white and black defendants near the margin of release. Though IV estimators are often criticized for the local nature of the estimates, we exploit the fact that the Becker test relies on (the differences between) exactly these kinds of local treatment effects to test for racial bias. In our context, our IV estimator measures the weighted average of racial bias across all bail judges with relatively few auxiliary assumptions, but at the potential cost that we cannot estimate judge-specific treatment effects and the weighting scheme underlying the IV estimator is not always policy relevant. In contrast, our second estimator uses the marginal treatment effects (MTE) framework developed by Heckman and Vytlacil (1999, 2005) to estimate judge-specific treatment effects for white and black defendants at the margin of release. Our MTE estimator therefore allows us to put equal weight on each judge in our sample, but with the estimation of the judge-specific estimates coming at the cost of additional auxiliary assumptions.

The second part of the paper tests for racial bias in bail setting using administrative court data from Miami and Philadelphia. We find evidence of significant racial bias against black defendants using both our IV and MTE estimators, ruling out statistical discrimination as the sole explanation for the racial disparities in bail. We find that marginally released white defendants are 22.2 to 23.1

[2]We replicate the Knowles et al. (2001) and Anwar and Fang (2006) tests in our data, finding no evidence of racial bias in either case. The differences between our test and the Knowles et al. (2001) and Anwar and Fang (2006) tests are that (1) we identify treatment effects for marginal defendants rather than the average defendant, and (2) we identify absolute rather than relative bias. See Section III.C for additional details on why the Knowles et al. (2001) and Anwar and Fang (2006) tests yield different results than our test.

[3]See Dobbie et al. (2018), Gupta, Hansman, and Frenchman (2016), Leslie and Pope (2017), and Stevenson (2016) for evidence on the nonfinancial consequences of bail decisions.

percentage points more likely to be rearrested prior to disposition than marginally released black defendants using our IV and MTE estimators, respectively. Our estimates of racial bias are nearly identical if we account for other observable crime and defendant differences by race, suggesting that our results cannot be explained by black–white differences in certain types of crimes (e.g., the proportion of felonies versus misdemeanors) or black–white differences in defendant characteristics (e.g., the proportion with prior offenses versus no prior offenses). In sharp contrast to these results, naïve OLS estimates indicate, if anything, racial bias against white defendants, highlighting the importance of accounting for both infra-marginality and omitted variables when estimating racial bias in the criminal justice system.

In the final part of the paper, we explore which form of racial bias is driving our findings. The first possibility is that, as originally modeled by Becker (1957, 1993), racial animus leads judges to discriminate against black defendants at the margin of release. This type of taste-based racial bias may be a particular concern in our setting due to the relatively low number of minority bail judges, the rapid-fire determination of bail decisions, and the lack of face-to-face contact between defendants and judges. A second possibility is that bail judges rely on incorrect inferences of risk based on defendant race due to anti-black stereotypes, leading to the relative overdetention of black defendants at the margin. These inaccurate anti-black stereotypes can arise if black defendants are overrepresented in the right tail of the risk distribution, even when the difference in the riskiness of the average black defendant and the average white defendant is very small (Bordalo et al. 2016). As with racial animus, these racially biased prediction errors in risk may be exacerbated by the fact that bail judges must make quick judgments on the basis of limited information, with virtually no training and, in many jurisdictions, little experience working in the bail system.

We find three sets of facts suggesting that our results are driven by bail judges relying on inaccurate stereotypes that exaggerate the relative danger of releasing black defendants versus white defendants at the margin. First, we find that both white and black bail judges exhibit racial bias against black defendants, a result that is inconsistent with most models of racial animus. Second, we find that our data are strikingly consistent with the theory of stereotyping developed by Bordalo et al. (2016). For example, we find that black defendants are sufficiently overrepresented in the right tail of the predicted risk distribution, particularly for violent crimes, to rationalize observed racial disparities in release rates under a stereotyping model. We also find that there is no racial bias against Hispanics, who, unlike blacks, are not significantly overrepresented in the right tail of the predicted risk distribution. Finally, we find substantially more racial bias when prediction errors of any kind are more likely to occur. For example, we find substantially less racial bias among both the full-time and more experienced part-time judges who are least likely to rely on simple race-based heuristics, and substantially more racial bias among the least experienced part-time judges who are most likely to rely on these heuristics.

Our findings are broadly consistent with parallel work by Kleinberg et al. (2018), who use machine learning techniques to show that bail judges make significant prediction errors for defendants of all races. Using a machine learning algorithm to predict risk using a variety of inputs such as prior and current criminal charges, but excluding defendant race, they find that the algorithm could reduce crime and jail populations while simultaneously reducing racial disparities. Their results also suggest that variables that are unobserved in the data, such as a judge's mood or a defendant's demeanor at the bail hearing, are the source of prediction errors, not private information that leads to more accurate risk predictions. Our results complement Kleinberg et al. (2018) by documenting one specific source of these prediction errors—racial bias among bail judges.

Our results also contribute to an important literature testing for racial bias in the criminal justice system. As discussed above, Knowles et al. (2001) and Anwar and Fang (2006) are seminal works in this area. Subsequent work has used outcome tests to examine racial bias in police search decisions (Antonovics and Knight 2009), capital sentencing (Alesina and La Ferrara 2014), and parole board release decisions (Mechoulan and Sahuguet 2015; Anwar and Fang 2015). Racial bias in bail setting has been studied using the prices charged by bail bond dealers (Ayres and Waldfogel 1994) and a parametric framework to account for unobserved heterogeneity across defendants (Bushway and Gelbach 2011). Our paper is also related to work using LATEs provided by IV estimators to obtain effects at the margin of the instrument (e.g., Card 1999; Gruber, Levine, and Staiger 1999) and work using MTEs to extrapolate to other estimands of interest (e.g., Heckman and Vyltacil 2005; Heckman, Urzua, and Vytlacil 2006; Cornelissen et al. 2016; Brinch, Mogstad, and Wiswall 2017).

The remainder of the paper is structured as follows. Section I provides an overview of the bail system, describes the theoretical model underlying our analysis, and develops our empirical test for racial bias. Section II describes our data and empirical methodology. Section III presents the main results. Section IV explores potential mechanisms, and Section V concludes. The Online Appendix provides all additional results, theoretical proofs, and detailed information on our setting.

I. An Empirical Test of Racial Bias

In this section, we motivate and develop our empirical test for racial bias in bail setting. Our theoretical framework closely follows the previous literature on the outcome test in the criminal justice system (e.g., Becker 1957, 1993; Knowles et al. 2001; Anwar and Fang 2006). Consistent with the prior literature, we show that we can test for racial bias by comparing treatment effects for the marginal black and marginal white defendants. We then develop two complementary estimators to identify these race-specific treatment effects using the quasi-random assignment of cases to judges. Online Appendix B provides additional details and proofs.

I.A. Overview of the Bail System

In the United States, bail judges are granted considerable discretion to determine which defendants should be released before trial. Bail judges are meant to balance two competing objectives when deciding whether to detain or release a defendant before trial. First, bail judges are directed to release all but the most dangerous defendants before trial to avoid undue punishment for defendants who have not yet been convicted of a crime. Second, bail judges are instructed to minimize the risk of pretrial misconduct by setting the appropriate conditions for release. In our setting, pretrial misconduct includes both the risk of new criminal activity and the risk of failure to appear for a required court appearance. Importantly, bail judges are not supposed to assess guilt or punishment at the bail hearing.

The conditions of release are set at a bail hearing typically held within 24 to 48 hours of a defendant's arrest. In most jurisdictions, bail hearings last only a few minutes and are held through a videoconference to the detention center such that judges can observe each defendant's demeanor. During the bail hearing, the assigned bail judge considers factors such as the nature of the alleged offense, the weight of the evidence against the defendant, the nature and probability of danger that the defendant's release poses to the community, the likelihood of flight based on factors such as the defendant's employment status and living situation, and any record of prior flight or bail violations, among other factors (Foote et al. 1954). Because bail judges are granted considerable discretion in setting the appropriate bail conditions, there are substantial differences across judges in the same jurisdiction (e.g., Dobbie et al. 2018; Gupta et al. 2016; Leslie and Pope 2017; Stevenson 2016).

The assigned bail judge has a number of potential options when setting a defendant's bail conditions. For example, the bail judge can release low-risk defendants on a promise to return for all court appearances, known as release on recognizance (ROR). For defendants who pose a higher risk of flight or new crime, the bail judge can allow release but impose nonmonetary conditions such as electronic monitoring or periodic reporting to pretrial services. The judge can also require defendants to post a monetary amount to secure release, typically 10 percent of the total bail amount. If the defendant fails to appear at the required court appearances or commits a new crime while out on bail, either he or the bail surety forfeits the 10 percent payment and is liable for the remaining 90 percent of the total bail amount. In practice, the median bail amount is $6,000 in our sample, and only 57 percent of defendants meet the required monetary conditions to secure release. Bail may also be denied altogether for defendants who commit the most serious crimes such as first- or second-degree murder.

One important difference between jurisdictions is the degree to which bail judges specialize in conducting bail hearings. In our setting, the bail judges we study in Philadelphia are full-time specialists who are tasked with setting bail seven days a week throughout the entire year. In contrast, the bail judges we study in Miami are part-time nonspecialists who assist the bail court by serving weekend shifts once or twice per year. These weekend bail judges spend their weekdays as trial judges. We explore the potential importance of these institutional features in Section IV.

I.B. Model of Judge Behavior

This section develops a stylized theoretical framework that allows us to define an outcome-based test of racial bias in bail setting. We begin with a model of taste-based racial bias or racial animus that closely follows Becker (1957, 1993). We then present an alternative model of racially biased prediction errors, which generates similar empirical predictions as the taste-based model.

1. Taste-Based Discrimination. Let i denote a defendant and $\mathbf{V}_i$ denote all case and defendant characteristics considered by the bail judge, excluding defendant race r_i. The expected cost of release for defendant i conditional on observable characteristics $\mathbf{V}_i$ and race r_i is equal to the expected probability of pretrial misconduct $\mathbb{E}[\alpha_i|\mathbf{V}_i, r_i]$, which includes the likelihood of both new crime and failure to appear, times the cost of misconduct C, which includes the social cost of any new crime or failures to appear. For simplicity, we normalize $C = 1$, so that the expected cost of release conditional on observable characteristics is equal to $\mathbb{E}[\alpha_i|\mathbf{V}_i, r_i]$. Moving forward, we also simplify our notation by letting the expected cost of release conditional on observables be denoted by $\mathbb{E}[\alpha_i|r_i]$.

The perceived benefit of release for defendant i assigned to judge j is denoted by $t_r^j(\mathbf{V}_i)$, which is a function of observable case and defendant characteristics $\mathbf{V}_i$. The perceived benefit of release $t_r^j(\mathbf{V}_i)$ includes social cost savings from reduced jail time, private gains to defendants from an improved bargaining position with the prosecutor or increased labor force participation, and personal benefits to judge j from any direct utility or disutility from being known as either a lenient or tough judge, respectively. Importantly, we allow the perceived benefit of release $t_r^j(\mathbf{V}_i)$ to vary by race $r \in W, B$ to allow for judge preferences to differ for white and black defendants.

Definition 1. Following Becker (1957, 1993), we define judge j as racially biased against black defendants if $t_W^j(\mathbf{V}_i) > t_B^j(\mathbf{V}_i)$. Thus, for racially biased judges, there is a higher perceived benefit of releasing white defendants than releasing observably identical black defendants.

For simplicity, we assume that bail judges are risk neutral and maximize the perceived net benefit of pretrial release. We also assume that the bail judge's sole task is to decide whether to release or detain a defendant given that this decision margin is the most important and consequential (Kleinberg et al. 2018; Dobbie et al. 2018). In simplifying each judge's task to this single decision, we abstract away from the fact that bail judges may set different levels of monetary bail that take into account a defendant's ability to pay. We discuss possible extensions to the model that account for these features below.

Under these assumptions, the model implies that bail judge j will release defendant i if and only if the expected cost of pretrial release is less than the perceived benefit of release:

$$\mathbb{E}[\alpha_i|r_i = r] \leq t_r^j(\mathbf{V}_i) \tag{1}$$

Given this decision rule, the marginal defendant for judge j and race r is the defendant i for whom the expected cost of release is exactly equal to the perceived benefit of release, i.e. $\mathbb{E}[\alpha_i^j|r_i = r] = t_r^j(\mathbf{V}_i)$. We simplify our notation moving forward by letting this expected cost of release for the marginal defendant for judge j and race r be denoted by α_r^j.

Based on the above framework and Definition 1, the model yields the familiar outcome-based test for racial bias from Becker (1957, 1993):

Proposition 1. If judge j is racially biased against black defendants, then $\alpha_W^j > \alpha_B^j$. Thus, for racially biased judges, the expected cost of release for the marginal white defendant is higher than the expected cost of release for the marginal black defendant.

Proposition 1 predicts that marginal white and marginal black defendants should have the same probability of pretrial misconduct if judge j is racially unbiased, but marginal white defendants should have a higher probability of misconduct if judge j is racially biased against black defendants.

2. Racially Biased Prediction Errors in Risk. In the taste-based model of discrimination outlined above, we assume that judges agree on the (true) expected cost of release, $\mathbb{E}[\alpha_i|r_i]$, but not the perceived benefit of release, $t_r^j(\mathbf{V}_i)$. An alternative approach is to assume that judges disagree on their (potentially inaccurate) predictions of the expected cost of release, as would be the case if judges systematically overestimate the probability of pretrial misconduct for black defendants relative to white defendants. We show that a model motivated by these kinds of racially biased prediction errors in risk can generate the same predictions as a model of taste-based discrimination.

Let i again denote defendants and $\mathbf{V}_i$ denote all case and defendant characteristics considered by the bail judge, excluding defendant race r_i. The perceived benefit of releasing defendant i assigned to judge j is now defined as $t(\mathbf{V}_i)$, which does not vary by judge.

The perceived cost of release for defendant i conditional on observable characteristics $\mathbf{V}_i$ is equal to the perceived probability of pretrial misconduct, $\mathbb{E}^j[\alpha_i|\mathbf{V}_i, r_i]$, which is now allowed to vary across judges. We can write the perceived cost of release as:

$$\mathbb{E}^j[\alpha_i|\mathbf{V}_i, r_i] = \mathbb{E}[\alpha_i|\mathbf{V}_i, r_i] + \tau_r^j(\mathbf{V}_i) \tag{2}$$

where $\tau_r^j(\mathbf{V}_i)$ is a prediction error that is allowed to vary by judge j and defendant race r_i. To simplify our notation, we let the true expected probability of pretrial misconduct conditional on all variables observed by the judge be denoted by $\mathbb{E}[\alpha_i|r_i]$.

Definition 2. We define judge j as making racially biased prediction errors in risk against black defendants if $\tau_B^j(\mathbf{V}_i) > \tau_W^j(\mathbf{V}_i)$. Thus, judges making racially biased prediction errors systematically overestimate the true cost of release for black defendants relative to white defendants.

Following the taste-based model, bail judge j will release defendant i if and only if the benefit of pretrial release is greater than the perceived cost of release:

$$\mathbb{E}^j[\alpha_i|\mathbf{V}_i, r_i = r] = \mathbb{E}[\alpha_i|r_i = r] + \tau_r^j(\mathbf{V}_i) \leq t(\mathbf{V}_i) \tag{3}$$

Given the above setup, it is straightforward to show that the prediction error model can be reduced to the taste-based model of discrimination outlined above if we relabel $t(\mathbf{V}_i) - \tau_r^j(\mathbf{V}_i) = t_r^j(\mathbf{V}_i)$. As a result, we can generate identical empirical predictions using the prediction error and taste-based models.

Following this logic, our model of racially biased prediction errors in risk yields a similar outcome-based test for racial bias:

Proposition 2. If judge j systematically overestimates the true expected cost of release of black defendants relative to white defendants, then $\alpha_W^j > \alpha_B^j$. Thus, for judges who make racially biased prediction errors in risk, the true expected cost of release for the marginal white defendant is higher than the true expected cost of release for the marginal black defendant.

Parallel to Proposition 1, Proposition 2 predicts that marginal white and marginal black defendants should have the same probability of pretrial misconduct if judge j does not systematically make prediction errors in risk that vary with race, but marginal white defendants should have a higher probability of misconduct if judge j systematically overestimates the true expected cost of release of black defendants relative to white defendants.

Regardless of the underlying behavioral model that drives the differences in judge behavior, the empirical predictions generated by these outcome-based tests are identical: if there is racial bias against black defendants, then marginal white defendants will have a higher probability of misconduct than marginal black defendants. In contrast, marginal white defendants will not have a higher probability of misconduct than marginal black defendants if observed racial disparities in bail setting are solely due to statistical discrimination.[4] Of course, finding higher misconduct rates for marginal white versus marginal black defendants does have a different interpretation depending on the underlying behavioral model. We will return to this issue in Section IV when we discuss more speculative evidence that allows us to differentiate between these two forms of racial bias.

I.C. Empirical Test of Racial Bias in Bail Setting

The goal of our analysis is to empirically test for racial bias in bail setting using the rate of pretrial misconduct for white defendants and black defendants at the margin of release. Following the theory model, let the weighted average of pretrial misconduct rates for defendants of race r at the margin for judge j, α_r^j, for some weighting scheme, w^j, across all bail judges, $j = 1...J$, be given by:

$$\alpha_r^{*,w} = \sum_{j=1}^{J} w^j \alpha_r^j = \sum_{j=1}^{J} w^j t_r^j \tag{4}$$

where w^j are non-negative weights which sum to one that will be discussed in further detail below. By definition, $\alpha_r^j = t_r^j$, where t_r^j represents judge j's threshold for release for defendants of race r. In

[4] In contrast to the two models we consider in this section, models of (accurate) statistical discrimination suggest that blacks may be treated worse than observably identical whites if either (1) blacks are, on average, riskier given an identical signal of risk (e.g., Phelps 1972; Arrow 1973) or (2) blacks have less precise signals of risk (e.g., Aigner and Cain 1977). In both types of (accurate) statistical discrimination models, however, judges use race to form accurate predictions of risk, both on average and at the margin of release. As a result, neither form of (accurate) statistical discrimination will lead to marginal white defendants having a higher probability of misconduct than marginal black defendants.

our context, pretrial misconduct rates can be identified by the treatment effect of pretrial release on misconduct, as defendants detained before trial cannot, by definition, commit pretrial misconduct. Thus, $\alpha_r^{*,w}$ represents a weighted average of the treatment effects for defendants of race r at the margin of release across all judges.

Following this notation, the average level of racial bias among bail judges, $D^{*,w}$, for the weighting scheme w^j is given by:

$$
\begin{aligned}
D^{*,w} &= \sum_{j=1}^{J} w^j \left(t_W^j - t_B^j \right) \\
&= \sum_{j=1}^{J} w^j t_W^j - \sum_{j=1}^{J} w^j t_B^j \\
&= \alpha_W^{*,w} - \alpha_B^{*,w}
\end{aligned}
\tag{5}
$$

From Equation (4), we can express $D^{*,w}$ as a weighted average across all judges of the difference in treatment effects for white and black defendants at the margin of release.

Standard OLS estimates will typically not recover unbiased estimates of the weighted average of racial bias, $D^{*,w}$, for two reasons. First, characteristics observable to the judge but not the econometrician may be correlated with pretrial release, resulting in omitted variable bias when estimating the treatment effects for black and white defendants. The second, and more important, reason OLS estimates will not recover unbiased estimates of racial bias is that the average treatment effect identified by OLS will not equal the treatment effect at the margin required by the outcome test unless one is willing to assume either identical risk distributions for black and white defendants or constant treatment effects across the entire distribution of both black and white defendants (e.g., Ayres 2002). Thus, even if the econometrician observes the full set of observables known to the bail judge, OLS estimates are still not sufficient to test for racial bias without restrictive assumptions.[5]

We therefore develop two complementary estimators for racial bias that use variation in the release tendencies of quasi-randomly assigned bail judges to identify differences in pretrial misconduct rates at the margin of release. Our first estimator uses the standard IV framework to identify the difference in LATEs for white and black defendants near the margin of release. Our IV estimator allows us to estimate a weighted average of racial bias across bail judges with relatively few auxiliary assumptions, but with the caveats that we cannot estimate judge-specific treatment effects and the weighting scheme underlying the IV estimator may not be policy relevant. In contrast, our second estimator uses the MTE framework developed by Heckman and Vytlacil (1999, 2005) to estimate judge-specific treatment effects for white and black defendants at the margin of release, allowing us to choose our own weighting scheme when calculating racial bias in our data. In practice, we choose to impose equal weights on each judge—a parameter with a clear economic interpretation—meaning

[5] In Online Appendix C, we use a series of simple graphical examples to illustrate how a standard OLS estimator suffers from infra-marginality bias whenever there are differences in the risk distributions of black and white defendants. We then use a simple two-judge example to illustrate how a judge IV estimator can alleviate the infra-marginality bias.

that our MTE estimates can be interpreted as the average level of bias across judges.

1. Setup. We first briefly review the baseline assumptions that underlie both our IV and MTE estimators. Online Appendix B provides empirical tests of each assumption.

Let Z_i be a scalar measure of the assigned judge's propensity for pretrial release for defendant-case i that takes on values ordered $\{z_0, ..., z_J\}$, where $J+1$ is the number of total judges in the bail system. For example, a value of $z_j = 0.5$ indicates that judge j releases 50 percent of all defendants. In practice, we construct Z_i using a standard leave-out procedure that captures the pretrial release tendencies of judges. We calculate Z_i separately for white and black defendants to relax the standard monotonicity assumption that the judge ordering produced by the scalar Z_i is the same for both white and black defendants, implicitly allowing judges to exhibit different levels of racial bias.

Following Imbens and Angrist (1994), a race-specific estimator using Z_i as an instrumental variable for pretrial release is valid and well-defined under the following three assumptions:

Assumption 1. (Existence) $Cov(Released_i, Z_i) \neq 0$
Assumption 2. (Exclusion) $Cov(Z_i, \mathbf{v}_i) = 0$
Assumption 3. (Monotonicity) $Released_i(z_j) - Released_i(z_{j-1}) \geq 0$

where $\mathbf{v}_i = \mathbf{U}_i + \varepsilon_i$ consists of characteristics unobserved by the econometrician but observed by the judge, $\mathbf{U}_i$, and idiosyncratic variation unobserved by both the econometrician and judge, ε_i. Assumption 1 ensures that there is a first-stage relationship between our instrument Z_i and the probability of pretrial release $Released_i$. Assumption 2 ensures that our instrument Z_i is orthogonal to characteristics unobserved by the econometrician, $\mathbf{v}_i$. In other words, Assumption 2 assumes that the assigned judge only affects pretrial misconduct through the channel of pretrial release. Assumption 3 implies that for a given case, any defendant released by a strict judge would also be released by a more lenient judge, and any defendant detained by a lenient judge would also be detained by a more strict judge.

2. IV Estimator for Racial Bias Given Assumptions 1–3, we now formally define our IV estimator for racial bias, provide conditions for consistency, and discuss the interpretation of the IV weights.

a. Defining our IV estimator. Let the true IV-weighted level of racial bias, $D^{*,IV}$ be defined as:

$$D^{*,IV} = \sum_{j=1}^{J} w^j \left(t_W^j - t_B^j\right) = \sum_{j=1}^{J} \lambda^j \left(t_W^j - t_B^j\right) \tag{6}$$

where $w^j = \lambda^j$, the standard IV weights defined in Imbens and Angrist (1994).

Let our IV estimator that uses judge leniency as an instrumental variable for pretrial release be defined as:

$$D^{IV} = \alpha_W^{IV} - \alpha_B^{IV} \tag{7}$$
$$= \sum_{j=1}^{J} \lambda_W^j \alpha_W^{j,j-1} - \sum_{j=1}^{J} \lambda_B^j \alpha_B^{j,j-1}$$

where λ_r^j are again the standard IV weights and each pairwise treatment effect $\alpha_r^{j,j-1}$ captures the treatment effects of compliers within each j, $j-1$ pair. As we discuss in Online Appendix B, compliers for judge j and $j-1$ are individuals such that $\alpha_r^{j,j-1} \in (t_r^{j-1}, t_r^j]$.

b. Consistency of our IV estimator: Our IV estimator D^{IV} provides a consistent estimate of $D^{*,IV}$ under two conditions: (1) Z_i is continuous and (2) λ_r^j is constant by race. See Online Appendix B for proofs of consistency. The first condition is that our judge leniency measure Z_i is continuously distributed over some interval $[\underline{z}, \bar{z}]$. Intuitively, each defendant becomes marginal to a judge as the distance between any two judge leniency measures converges to zero, i.e. the instrument becomes more continuous. Under this first condition, each race-specific IV estimate, α_r^{IV}, approaches a weighted average of treatment effects for defendants at the margin of release. In Online Appendix B, we discuss the potential infra-marginality bias that may result if our instrument is discrete, as is the case in our data. In practice, we find that the maximum infra-marginality bias of our IV estimator D^{IV} from $D^{*,IV}$ is 1.1 percentage points in our setting. The second condition for consistency is that the weights on the pairwise LATEs must be equal across race. This equal weights assumption ensures that the race-specific IV estimates from Equation (7), α_W^{IV} and α_B^{IV}, provide the same weighted averages of $\alpha_W^{j,j-1}$ and $\alpha_B^{j,j-1}$. In Online Appendix B, we empirically test whether the IV weights λ_r^j are constant by race in our data, finding that the distributions of black and white IV weights are visually indistinguishable from each other and that the IV weights for each judge-by-year cell are highly correlated across race.

c. Interpretation of the IV weights: As discussed above, our IV estimator yields a weighted average of racial bias across bail judges, where the weights λ^j are the standard IV weights defined in Imbens and Angrist (1994). To better understand the economic interpretation of an IV-weighted estimate of racial bias, Online Appendix B investigates the relationship between our IV weights and judge-by-year characteristics. We find that our IV weights are positively correlated with both the number of cases in a judge-by-year cell and judge-by-year specific estimates of racial bias, implying that the IV-weighted estimate of racial bias may be larger than an equal-weighted estimate of racial bias. We return to this issue below when discussing the difference between our IV and MTE estimates.

3. MTE Estimator for Racial Bias Finally, we formally define our MTE estimator of racial bias and provide conditions for consistency. Without loss of generality, we focus on an estimate of racial bias that places equal weight on each bail judge.

a. Defining our MTE estimator: Let the true equal-weighted MTE estimate of racial bias, $D^{*,MTE}$ be defined as:

$$D^{*,MTE} = \sum_{j=1}^{J} w^j \left(t_W^j - t_B^j \right) = \sum_{j=1}^{J} \frac{1}{J} \left(t_W^j - t_B^j \right) \tag{8}$$

where $w^j = \frac{1}{J}$, such that $D^{*,MTE}$ can be interpreted as the average level of racial bias across judges.

Let our equal-weighted MTE estimator of racial bias, D^{MTE}, be defined as:

$$D^{MTE} = \sum_{j=1}^{J} \frac{1}{J} \left(MTE_W(p_W^j) - MTE_B(p_B^j) \right) \tag{9}$$

where p_r^j is the probability that judge j releases a defendant of race r calculated using only the variation in pretrial release due to our judge leniency measure Z_i (i.e. judge j's race-specific propensity score). $MTE_r(p_r^j)$ is the estimated MTE at the propensity score for judge j calculated separately for each defendant race r. In Online Appendix B, we show that $MTE_r(p_r^j) = \alpha_r^j$ when we map each judge j's release decision under our theory model to the MTE framework developed by Heckman and Vytlacil (2005).

b. Consistency of our MTE estimator: Our MTE estimator D^{MTE} provides a consistent estimate of $D^{*,MTE}$ if the race-specific MTEs are identified over the entire support of the propensity score calculated using variation in Z_i. If Z_i is continuous, the local instrumental variables (LIV) estimand provides a consistent estimate of the MTE over the support of the propensity score with no additional assumptions (Heckman and Vytlacil 2005; Cornelissen et al. 2016). With a discrete instrument, however, our MTE estimator is only consistent under additional functional form restrictions that allow us to interpolate the MTEs between the values of the propensity score we observe in the data. Following Heckman and Vytlacil (2005) and Doyle (2007), we use a local polynomial function and information from the observed values of the propensity score to estimate the MTE curve over the full support of the propensity score. In Online Appendix B, we provide support for our functional form assumption by showing that we can recover each nonparametric LATE using the appropriately weighted MTE up to sampling error (Cornelissen et al. 2016).

I.D. Discussion and Extensions

In this section, we discuss the interpretation of our test of racial bias under different assumptions and extensions.

1. Racial Differences in Arrest Probability. Our test for racial bias assumes that any measurement error in the outcome is uncorrelated with race. This assumption would be violated if, for example, judges minimize new crime, not just new arrests, and police are more likely to rearrest black defendants conditional on having committed a new crime (Fryer 2016; Goncalves and Mello 2018). In this scenario, we will overestimate the probability of pretrial misconduct for black versus white defendants at the margin and, as a result, underestimate the true amount of racial bias in bail setting. It is therefore possible that our estimates reflect a lower bound on the true amount of racial bias among bail judges to the extent that judges minimize new crime.[6]

2. Omitted Objectives for Release. We also assume that judges do not consider other objectives or outcomes, or what Kleinberg et al. (2018) refer to as "omitted payoff bias." We will have this kind of omitted payoff bias if, for example, bail judges consider how pretrial detention impacts a defendant's employment status and this outcome is correlated with race.

We explore the empirical relevance of omitted payoff bias in several ways. First, as will be discussed below, we find that our estimates are nearly identical if we measure pretrial misconduct using only rearrests versus using rearrests or failures to appear. These results are also consistent with Kleinberg et al. (2018), who find similar evidence of prediction errors using rearrests or failures to appear. Second, as will be discussed below, we also find similar estimates when we measure pretrial misconduct using crime-specific rearrest rates to address the concern that judges may be most concerned about reducing violent crimes. Third, we note that Dobbie et al. (2018) find that white defendants at the margin of release are no more likely to be employed in the formal labor market up to four years after the bail hearing compared to black defendants at the margin of release. This goes against the idea that judges may be trading off minimizing pretrial misconduct with maximizing employment. Finally, as will be discussed below, we find that racial bias against black defendants is larger for part-time and inexperienced judges compared to full-time and experienced judges. There are few conceivable stories where omitted payoffs differ by judge experience. Taken together, we therefore believe that any omitted payoff bias is likely to be small in practice.

3. Racial Differences in Ability to Pay Monetary Bail. In our model, we abstract away from the fact that bail judges may set different levels of monetary bail that, by law, should take into account a defendant's ability to pay. Extending our model to incorporate these institutional details means that racial bias could also be driven by judges systematically over-predicting the relative ability of black defendants to pay monetary bail at the margin.

We explore the empirical relevance of racial differences in ability to pay monetary bail in two ways. First, we test whether the assignment of non-monetary bail (i.e., either ROR or non-monetary conditions) versus monetary bail has a larger impact on the probability of release for marginal black defendants,[7] which could occur if judges systematically over-predict black defendants' ability to pay monetary bail at the margin. To test this idea, Panel A of Online Appendix Table A1 presents

[6]A related concern is that bail judges may be influenced by other court actors (e.g. prosecutors) when making decisions, such that racial bias stems from judges not overriding racially biased bail recommendations. However, we find substantial variation in pretrial release tendencies across judges, inconsistent with the idea that judges "rubber-stamp" bail recommendations. We also find that racial bias decreases with judge experience, inconsistent with other court actors driving the racial bias unless experience affects the probability of overriding biased recommendations.

[7]Dobbie et al. (2018) show that the assignment of ROR and non-monetary conditions have a statistically identical impact on defendant outcomes, including pretrial misconduct. We therefore combine ROR and non-monetary conditions into a single category in our analysis.

two-stage least squares estimates of the impact of non-monetary versus monetary bail on pretrial release using a leave-out measure based on non-monetary bail decisions as an instrumental variable. We find that the assignment of non-monetary bail versus monetary bail has a nearly identical impact on the pretrial release rates for marginal black defendants and marginal white defendants. Second, we directly estimate racial bias in the setting of non-monetary versus monetary bail to incorporate any additional bias stemming from this margin. We estimate these effects using a two-stage least squares regression of pretrial misconduct on non-monetary bail, again using a leave-out measure based on non-monetary bail decisions as an instrumental variable. Panel B of Online Appendix Table A1 presents these estimates. We find similar estimates of racial bias when focusing on the non-monetary versus monetary bail decision when we scale the estimated treatment effects by the "first stage" effect of non-monetary bail on pretrial release from Panel A.

4. Judge Preferences for Non-Race Characteristics. Another extension to our model concerns two distinct views about what constitutes racial bias. The first is that racial bias includes not only any bias due to phenotype, but also bias due to seemingly non-race factors that are correlated with, if not driven by, race. For example, judges could be biased against defendants charged with drug offenses because blacks are more likely to be charged with these types of crimes. Our preferred estimates are consistent with this broader view of racial bias, measuring the disparate treatment of black and white defendants at the margin for <u>all</u> reasons unrelated to true risk of pretrial misconduct, including reasons related to seemingly non-race characteristics such as crime type.

A second view is that racial bias is disparate treatment due to phenotype alone, not other correlated factors such as crime type. In Online Appendix B, we show that it is possible to test for this narrower form of racial bias using a re-weighting procedure that weights the distribution of observables of blacks to match observables of whites in the spirit of DiNardo, Fortin, and Lemieux (1996) and Angrist and Fernández-Val (2013). This narrower test for racial bias relies on the assumption that judge preferences vary only by observable characteristics $\mathbf{X}_i$, i.e. $t_r^j(\mathbf{V}_i) = t_r^j(\mathbf{X}_i)$. We find nearly identical estimates of racial bias using this re-weighting procedure, suggesting that judge preferences over non-race characteristics are a relatively unimportant contributor to our findings. We discuss these results in robustness checks below.

II. Data and Instrument Construction

This section summarizes the most relevant information regarding our administrative court data from Philadelphia and Miami-Dade, describes the construction of our judge leniency measure, and provides support for the baseline assumptions required for our IV and MTE estimators of racial bias. Further details on the cleaning and coding of variables are contained in Online Appendix D.

II.A. Data Sources and Descriptive Statistics

Philadelphia court records are available for all defendants arrested and charged between 2010–2014 and Miami-Dade court records are available for all defendants arrested and charged between

2006–2014. For both jurisdictions, the court data contain information on defendant's name, gender, race, date of birth, and zip code of residence. Because our ethnicity identifier does not distinguish between non-Hispanic white and Hispanic white, we match the surnames in our dataset to census genealogical records of surnames. If the probability a given surname is Hispanic is greater than 70 percent, we label this individual as Hispanic. In our main analysis, we include all defendants and compare outcomes for marginal black and white (Hispanic and non-Hispanic) defendants. In robustness checks, we present results comparing marginal black and non-Hispanic white defendants.[8]

The court data also include information on the original arrest charge, the filing charge, and the final disposition charge. We also have information on the severity of each charge based on state-specific offense grades, the outcome for each charge, and the punishment for each guilty disposition. Finally, the case-level data include information on attorney type, arrest date, and the date of and judge presiding over each court appearance from arraignment to sentencing. Importantly, the case-level data also include information on bail type, bail amount when monetary bail is set, and whether bail was met. Because the data contain defendant identifiers, we can measure whether a defendant was subsequently arrested for a new crime before case disposition. In Philadelphia, we also observe whether a defendant failed to appear for a required court appearance.

We make three restrictions to the court data to isolate cases that are quasi-randomly assigned to judges. First, we drop a small set of cases with missing bail judge information or missing race information. Second, we drop the 30 percent of defendants in Miami-Dade who never have a bail hearing because they post bail immediately following arrest; below we show that the characteristics of defendants who have a bail hearing are uncorrelated with our judge leniency measure. Third, we drop all weekday cases in Miami-Dade because, as explained in Online Appendix E, bail judges in Miami-Dade are assigned on a quasi-random basis only on the weekends. The final sample contains 162,836 cases from 93,914 unique defendants in Philadelphia and 93,417 cases from 65,944 unique defendants in Miami-Dade.

Table I reports summary statistics for our estimation sample separately by race and pretrial release status. On average, black defendants are 3.6 percentage points more likely to be assigned monetary bail compared to white defendants and receive bail amounts that are $7,281 greater than white defendants (including zeros). Conversely, black defendants are 2.0 percentage points and 1.6 percentage points less likely to be released on their own recognizance or to be assigned non-monetary conditions compared to white defendants, respectively. As a result, black defendants are 2.4 percentage points more likely to be detained pretrial compared to white defendants.

Compared to white defendants, released black defendants are also 1.9 percentage points more likely to be rearrested for a new crime before case disposition, our preferred measure of pretrial misconduct. Released black defendants are also 0.9 percentage points, 0.7 percentage points, and 3.0 percentage points more likely to be rearrested for a drug, property, and violent crime, respectively. In Philadelphia, released black defendants are 1.4 percentage points more likely to fail to appear

[8]Online Appendix Table A2 presents results for marginal Hispanic white defendants compared to non-Hispanic white defendants. Perhaps in some part because of measurement error in our coding of Hispanic ethnicity, we find no evidence of racial bias against Hispanics.

in court compared to white defendants. Defining pretrial misconduct as either failure to appear or rearrest in Philadelphia, and only rearrest in Miami, released black defendants are 4.1 percentage points more likely to commit any form of pretrial misconduct compared to white defendants. We also find that approximately four percent of detained defendants are rearrested for a new crime prior to case disposition—an outcome that should be impossible. We show that our results are unaffected by dropping these cases in robustness checks.[9]

II.B. Construction of the Instrumental Variable

We estimate the causal impact of pretrial release for the marginal defendant using a measure of the tendency of a quasi-randomly assigned bail judge to release a defendant as an instrument for release. In both Philadelphia and Miami-Dade, there are multiple bail judges serving at each point in time, allowing us to utilize variation in bail setting across judges. Both jurisdictions also assign cases to bail judges in a quasi-random fashion in order to balance caseloads: Philadelphia utilizes a rotation system where three judges work together in five-day shifts, with one judge working an eight-hour morning shift (7:30AM–3:30PM), another judge working the eight-hour afternoon shift (3:30PM–11:30PM), and the final judge working the eight-hour evening shift (11:30PM–7:30AM). Similarly, bail judges in Miami-Dade rotate through the weekend felony and misdemeanor bail hearings. See Online Appendix E for additional details.

Following Dobbie et al. (2018), we construct our instrument using a residualized, leave-out judge leniency measure that accounts for the case assignment processes in Philadelphia and Miami-Dade. To construct this residualized judge leniency measure, we first regress pretrial release decisions on an exhaustive set of court-by-time fixed effects, the level at which defendants are quasi-randomly assigned to judges. In Miami, these court-by-time fixed effects include court-by-bail year-by-bail day of week fixed effects and court-by-bail month-by-bail day of week fixed effects. In Philadelphia, we add bail-day of week-by-bail shift fixed effects. We then use the residuals from this regression to calculate the leave-out mean judge release rate for each defendant. We calculate our instrument across all case types, but allow the instrument to vary across years and defendant race.[10]

Figure I presents the distribution of our residualized judge leniency measure for pretrial release at the judge-by-year level for all defendants, white defendants, and black defendants. Our sample includes seven total bail judges in Philadelphia and 170 total bail judges in Miami-Dade. In Philadelphia, the average number of cases per judge is 23,262 during the sample period of 2010–2014, with the typical judge-by-year cell including 5,253 cases. In Miami-Dade, the average number of cases per judge is 550 during the sample period of 2006–2014, with the typical judge-by-year cell including 179 cases. Controlling for the exhaustive set of court-by-time fixed effects, the judge

[9] To understand how these miscodings impact the interpretation of results, we follow Dahl et al. (2014) in calculating rearrest rates for marginally detained defendants. These estimates imply that the rearrest rate for marginally released defendants is approximately 2.0 to 3.0 percentage points higher than our estimated treatment effects.

[10] Our leave-out procedure is essentially a reduced-form version of jackknife IV, with the leave-out leniency measure for judge j being algebraically equivalent to judge j's fixed effect from a leave-out regression of residualized pretrial release on the full set of judge fixed effects and court-by-time fixed effects. In unreported results, jackknife IV and LIML estimates using the full set of judge fixed effects as instruments yield similar results.

release measure ranges from -0.283 to 0.253 with a standard deviation of 0.040. In other words, moving from the least to most lenient judge increases the probability of pretrial release by 53.6 percentage points, a 76.8 percent change from the mean release rate of 69.8 percentage points.

II.C. Instrument Validity

1. Existence of First Stage. To examine the first-stage relationship between judge leniency (Z_{itj}) and whether a defendant is released pretrial ($Released_{itj}$), we estimate the following equation for defendant-case i, assigned to judge j at time t using a linear probability model, estimated separately for white and black defendants:

$$Released_{itj} = \gamma_W Z_{itj} + \pi_W \mathbf{X}_{it} + v_{itj} \tag{10}$$

$$Released_{itj} = \gamma_B Z_{itj} + \pi_B \mathbf{X}_{it} + v_{itj} \tag{11}$$

where the vector $\mathbf{X}_{it}$ includes court-by-time fixed effects. The error term v_{itj} is composed of characteristics unobserved by the econometrician but observed by the judge, as well as idiosyncratic variation unobserved to both the judge and econometrician. As described previously, Z_{itj} are leave-out (jackknife) measures of judge leniency that are allowed to vary across years and defendant race. Robust standard errors are two-way clustered at the individual and judge-by-shift level.

Figure I provides graphical representations of the first stage relationship, for all defendants and separately by race, between our residualized measure of judge leniency and the residualized probability of pretrial release that accounts for our exhaustive set of court-by-time fixed effects. The graphs are a flexible analog to Equations (10) and (11), where we plot a local linear regression of residualized pretrial release against judge leniency. The individual rate of residualized pretrial release is monotonically increasing in our leniency measure for both races.

Table II presents formal first stage results from Equations (10) and (11) for all defendants, white defendants, and black defendants. Columns (1), (3), and (5) begin by reporting results with only court-by-time fixed effects. Columns (2), (4), and (6) add our baseline crime and defendant controls: race, gender, age, whether the defendant had a prior offense in the past year, whether the defendant had a prior history of pretrial crime in the past year, whether the defendant had a prior history of failure to appear in the past year, the number of charged offenses, indicators for crime type (drug, DUI, property, violent, or other), crime severity (felony or misdemeanor), and indicators for any missing characteristics.

We find that our residualized judge instrument is highly predictive of whether a defendant is released pretrial. Our results show that a defendant assigned to a bail judge that is 10 percentage points more likely to release a defendant pretrial is 3.89 percentage points more likely to be released pretrial. Judge leniency is also highly predictive of pretrial release for both white and black defendants, with the first-stage coefficient being 0.360 and 0.415, respectively.[11]

[11] Consistent with prior work using judge leniency as an instrumental variable (e.g., Bhuller et al. 2016), the probability of being released pretrial does not increase one-for-one with our measure of judge leniency, likely because

2. Exclusion Restriction. Table III verifies that assignment of cases to bail judges is random after we condition on our court-by-time fixed effects. Columns (1), (3), and (5) of Table III use a linear probability model to test whether case and defendant characteristics are predictive of pretrial release. These estimates capture both differences in the bail conditions set by bail judges and differences in defendants' ability to meet the bail conditions. We control for court-by-time fixed effects and two-way cluster standard errors at the individual and judge-by-shift level. For example, we find that black male defendants are 10.4 percentage points less likely to be released pretrial compared to similar female defendants, while white male defendants are 8.6 percentage points less likely to be released pretrial compared to similar female defendants. White defendants with at least one prior offense in the past year are 16.8 percentage points less likely to be released compared to similar defendants with no prior offenses, while black defendants with at least one prior offense in the past year are 13.4 percentage points less likely to be released compared to similar defendants with no prior offenses. Columns (2), (4), and (6) assess whether these same case and defendant characteristics are predictive of our judge leniency measure using an identical specification. We find that judges with differing leniencies are assigned cases with very similar defendants.

Even with random assignment, the exclusion restriction could be violated if bail judge assignment impacts the probability of pretrial misconduct through channels other than pretrial release. The assumption that judges only systematically affect defendant outcomes through pretrial release is fundamentally untestable, and our estimates should be interpreted with this potential caveat in mind. However, we argue that the exclusion restriction assumption is reasonable in our setting. Bail judges exclusively handle one decision, limiting the potential channels through which they could affect defendants. In addition, we are specifically interested in short-term outcomes (pretrial misconduct) which occur prior to disposition, further limiting the role of alternative channels that could affect longer-term outcomes. Finally, Dobbie et al. (2018) find that there are no independent effects of the money bail amount or non-monetary bail conditions on defendant outcomes, and that bail judge assignment is uncorrelated with the assignment of public defenders and trial judges.

3. Monotonicity. The final condition needed for our IV and MTE estimators is that the impact of judge assignment on the probability of pretrial release is monotonic across defendants of the same race. In our setting, the monotonicity assumption requires that individuals released by a strict judge would also be released by a more lenient judge, and that individuals detained by a lenient judge would also be detained by a stricter judge. The monotonicity assumption is required in order to identify and interpret our IV estimator as a well-defined LATE and to estimate marginal treatment effects using the standard local instrument variables (LIV) approach. See Angrist et al. (1996) and Heckman and Vytlacil (2005) for additional details. Importantly, we allow our judge leniency measure to vary by defendant race to allow for the possibility that the degree of racial bias varies

of attenuation bias due to sampling variation in the construction of our instrument. Consistent with this explanation, we find first stage coefficients ranging from 0.6 to 0.7 in Monte Carlo simulations when judge tendencies are fixed over the course of the year, and 0.2 to 0.4 when judge tendencies are allowed to change within each year. It is important to note that attenuation bias due to sampling variation in our leniency measure does not bias our estimates since it affects both the first stage and reduced form proportionally.

across judges. In practice, we observe that judge behavior is only imperfectly monotonic with respect to race (see Online Appendix Figure A1), with a regression of the ranking of each judge's leniency measure for whites on the ranking of each judge's leniency measure for blacks yielding a coefficient equal to 0.827 (se=0.010). The non-monotonic behavior we observe with respect to race is driven by approximately 17.9 percent of judges who hear about 8.2 percent of all cases. Consistent with the monotonicity assumption within race, we find a strong first-stage relationship across various case and defendant types (see Online Appendix Table A3).[12]

III. Results

In this section, we present our main results applying our empirical test for racial bias. We then show the robustness of our results to alternative specifications, before comparing the results from our empirical test with the alternative outcome-based tests developed by Knowles et al. (2001) and Anwar and Fang (2006).

III.A. Empirical Test for Racial Bias

1. IV Estimates. We begin by presenting IV estimates of racial bias that rely on relatively few auxiliary assumptions, but with the caveat that the weighting scheme underlying the estimator may not always be policy relevant. We estimate these IV results using the following two-stage least squares specifications for defendant-case i assigned to judge j at time t, estimated separately for white and black defendants:

$$Y_{itj} = \alpha_W^{IV} Released_{itj} + \beta_W \mathbf{X}_{it} + \mathbf{v}_{itj} \tag{12}$$

$$Y_{itj} = \alpha_B^{IV} Released_{itj} + \beta_B \mathbf{X}_{it} + \mathbf{v}_{itj} \tag{13}$$

where Y_{itj} is the probability of pretrial misconduct, as measured by the probability of rearrest prior to case disposition. The vector $\mathbf{X}_{it}$ includes court-by-time fixed effects and baseline crime and defendant controls: race, gender, age, whether the defendant had a prior offense in the past year, whether the defendant had a prior history of pretrial crime in the past year, whether the defendant had a prior history of failure to appear in the past year, the number of charged offenses, indicators for crime type (drug, DUI, property, violent, or other), crime severity (felony or misdemeanor), and indicators for any missing characteristics. As described previously, the error term $\mathbf{v}_{itj} = \mathbf{U}_{itj} + \varepsilon_{itj}$ consists of characteristics unobserved by the econometrician but observed by the judge, $\mathbf{U}_{itj}$, and idiosyncratic variation unobserved by both the econometrician and judge, ε_{itj}. We instrument for

[12] One specific concern is that lenient judges may be better at using unobservable information to predict the risk of pretrial misconduct, as this would result in some high-risk defendants being released by only strict judges. Following Kleinberg et al. (2018), we test for this possibility in Online Appendix Figure A2 by examining pretrial misconduct rates among observably identical defendants released by either lenient or strict judges. We find that predicted risk largely tracks true risk in all judge leniency quintiles, suggesting that lenient judges are neither more nor less skilled in predicting defendant risk. These results are broadly consistent with Kleinberg et al. (2018), who find that judges more or less agree on how to rank-order defendants based on their observable characteristics.

pretrial release, $Released_{itj}$, with our measure of judge leniency, Z_{itj}, that is allowed to vary across years and defendant race. Robust standard errors are two-way clustered at the individual and judge-by-shift level.

Estimates from Equations (12) and (13) are presented in columns (1)–(2) of Table IV. Column (3) reports our IV estimate of racial bias D^{IV}. Panel A of Table IV presents results for the probability of rearrest for any crime prior to case disposition, while Panel B presents results for rearrest rates for drug, property, and violent offenses separately. In total, 17.8 percent of defendants are rearrested for a new crime prior to disposition, with 7.9 percent of defendants rearrested for a crime that includes a drug offense, 6.7 percent of defendants rearrested for a crime that includes a property offense, and 6.1 percent of defendants rearrested for a crime that includes a violent offense.[13]

We find convincing evidence of racial bias against black defendants using our IV estimator. We find that marginally released white defendants are 23.6 percentage points more likely to be rearrested for any crime compared to marginally detained white defendants (column (1)). In contrast, the effect of pretrial release on rearrest rates for marginally released black defendants is a statistically insignificant 1.4 percentage points (column (2)). Loosely, these estimates imply that there is a 23.6 percent rate of rearrest for marginally released white defendants and a 1.4 percent rate of rearrest for marginally released black defendants, as detained defendants cannot be rearrested before trial.[14] Taken together, these IV estimates imply that marginally released white defendants are 22.2 percentage points more likely to be rearrested prior to disposition than marginally released black defendants (column (3)), consistent with racial bias against blacks ($p = .027$). Importantly, we can reject the null hypothesis of no racial bias even assuming the maximum infra-marginality bias in our IV estimator of 1.1 percentage points (see Online Appendix B).

In Panel B, we find suggestive evidence of racial bias against black defendants across all crime types, although the point estimates are too imprecise to make definitive conclusions. For example, we find that marginally released whites are about 8.0 percentage points more likely to be rearrested for a violent crime prior to disposition than marginally released blacks ($p = .173$). Marginally released white defendants are also 4.7 percentage points more likely to be rearrested for a drug crime prior to case disposition than marginally released black defendants ($p = .430$), and 16.3 percentage points more likely to be rearrested for a property crime ($p = .025$). These results suggest that judges are likely racially biased against black defendants even if they are most concerned about minimizing

[13]For completeness, Figure I provides a graphical representation of our reduced form results separately by race. Following the first stage results, we plot the reduced form relationship between our judge leniency measure and the residualized rate of rearrest prior to case disposition, estimated on the full sample using local linear regression. Consistent with the first stage estimates in Table II and IV estimates in Table IV, the reduced form relationship between judge leniency and rearrest rates is much flatter for black defendants compared to white defendants.

[14]Online Appendix Table A4 presents OLS results that measure the average level of risk among released white and black defendants conditional on observables. Taken together with our results in Table IV, these OLS estimates imply that the marginally released white defendant is riskier than the average released white defendant, while the marginally released black defendant is *less* risky than the average released black defendant. These results suggest that judges make substantial errors in predicting rearrest rates for black defendants, with all judges releasing relatively risky black defendants while disagreeing over relatively less risky black defendants. These findings are consistent with Kleinberg et al. (2018), who find that bail judges release many observably high-risk defendants while detaining many observably low-risk defendants.

specific types of new crime, such as violent crimes.

2. MTE Estimates. Our second set of estimates comes from our MTE estimator that allows us to put equal weight on each judge in our sample, but at the cost of additional auxiliary assumptions. We estimate these MTE results using a two-step procedure. First, we estimate the entire distribution of MTEs using the derivative of residualized rearrest before case disposition, $\ddot{Y}_{itj}$, with respect to variation in the propensity score provided by our instrument, p_r^j, separately for white and black defendants:

$$MTE_W(p_W^j) = \frac{\partial}{\partial p_W^j}\mathbb{E}(\ddot{Y}_{itj}|p_W^j, W) \tag{14}$$

$$MTE_B(p_B^j) = \frac{\partial}{\partial p_B^j}\mathbb{E}(\ddot{Y}_{itj}|p_B^j, B) \tag{15}$$

where p_r^j is the propensity score for release for judge j and defendant race r and $\ddot{Y}_{itj}$ is rearrest residualized using the full set of court-by-time fixed effects and baseline crime and defendant controls, $\mathbf{X}_{it}$. Following Heckman, Urzua, and Vytlacil (2006) and Doyle (2007), we also residualize Z_{itj} and $Released_{itj}$ using $\mathbf{X}_{it}$. We then regress the residualized release variable on the residualized judge leniency measure to calculate p_r^j, a race-specific propensity score. Next, we compute the numerical derivative of a local quadratic estimator relating $\ddot{Y}_{itj}$ to p_r^j to estimate race-specific MTEs. See Figure II for estimates of the full distribution of MTEs by defendant race.

Second, we use the race-specific MTEs to calculate the level of racial bias for each judge j. We calculate the average level of bias across all bail judges using a simple average of these judge-specific estimates:

$$\sum_{j=1}^{J}\frac{1}{J}\left(MTE_W(p_W^j) - MTE_B(p_B^j)\right) \tag{16}$$

We calculate standard errors by bootstrapping this two-step procedure at the judge-by-shift level. See Online Appendix B for additional details.

Estimates from Equations (14) and (15) are presented in columns (4)–(5) of Table IV, with column (6) reporting our MTE equal-weighted estimate of racial bias D^{MTE} from Equation (16). Consistent with our IV estimates, we find that marginally released white defendants are 24.9 percentage points more likely to be rearrested for any crime compared to marginally detained white defendants (column (4)), while the effect of pretrial release on rearrest rates for marginally released black defendants is a statistically insignificant 1.7 percentage points (column (5)). Our MTE estimates therefore imply that marginally released white defendants are 23.1 percentage points more likely to be rearrested prior to disposition than marginally released black defendants (column (6)), consistent with racial bias against black defendants ($p = .048$).

In addition, Figure II shows that the MTEs for white defendants lie strictly above the MTEs for black defendants, implying that marginally released white defendants are riskier than marginally released black defendants at all points in the judge leniency distribution. In other words, the results from Figure II show that there is racial bias against black defendants at every part of the judge leniency distribution. These results, along with the fact that both IV and MTE approaches yield qualitatively similar estimates of racial bias, suggest that both the choice of IV weights and the additional parametric assumptions required to estimate the race-specific MTEs do not greatly affect our estimates of racial bias.

III.B. Robustness

Online Appendix Tables A5–A6 explore whether our main findings are subject to omitted payoff bias. We find that our estimates are qualitatively similar when we use a measure of pretrial misconduct defined as failure to appear in Philadelphia, the only city where we observe this information (columns (1)–(2) of Online Appendix Table A5), or when we define pretrial misconduct as either failure to appear or rearrest in Philadelphia and only rearrest in Miami (columns (5)–(6) of Online Appendix Table A5). We also find that marginally released white defendants generate larger social costs than marginally released black defendants when we estimate results separately for a subset of more serious crimes and weight each individual estimate by the corresponding social cost (Online Appendix Table A6).

Online Appendix Table A7 explores the sensitivity of our main results to a number of different specifications. Columns (1) and (6) drop a small number of defendants who the data indicate were rearrested prior to disposition despite never being released. Column (2) presents reweighted estimates with the weights chosen to match the distribution of observable characteristics by race (see Section I.D and Online Appendix B for details). Columns (3) and (7) present results comparing outcomes for marginal non-Hispanic white defendants and black defendants. Columns (4) and (8) present results clustering more conservatively at the individual and judge level. Column (5) assesses whether monetary bail amounts have an independent effect on the probability of pretrial misconduct—a potential violation of the exclusion restriction—by controlling for monetary bail amount as an additional regressor in both our first- and second-stage regressions.[15] Under these alternative specifications, we continue to find evidence of racial bias against black defendants.

III.C. Comparison to Other Outcome Tests

Online Appendix Tables A4 and A9 replicate the outcome tests from Knowles et al. (2001) and Anwar and Fang (2006). The Knowles et al. (2001) test relies on the prediction that, under the null hypothesis of no racial bias, the average pretrial misconduct rate given by standard OLS estimates will not vary by defendant race. In contrast to our IV and MTE tests, however, standard OLS estimates suggest racial bias against <u>white</u> defendants. The Anwar and Fang (2006) test instead relies on the prediction that, under the null hypothesis of no relative racial bias, the treatment of

[15] In these specifications, the coefficient on monetary bail amount is -0.002 (p = .500) for white defendants and -0.001 (p = .184) for black defendants, suggesting that monetary bail amount has no significant independent effect on pretrial misconduct, consistent with findings reported in Dobbie et al. (2018).

black and white defendants will not depend on judge race. However, this test also fails to find racial bias in our setting because both white and black judges are racially biased against black defendants. We also find that the IV and MTE estimates of racial bias are similar among white and black judges, although the confidence intervals for these estimates are large. Taken together, these results highlight the importance of accounting for both infra-marginality and omitted variables, as well as the importance of developing empirical tests that can detect absolute racial bias in the criminal justice system. See Arnold, Dobbie, and Yang (2017) for additional details on these results.

IV. Potential Mechanisms

In this section, we attempt to differentiate between two alternative forms of racial bias that could explain our findings: (1) racial animus (e.g., Becker 1957, 1993) and (2) racially biased prediction errors in risk (e.g., Bordalo et al. 2016).

IV.A. Racial Animus

The first potential explanation for our results is that judges either knowingly or unknowingly discriminate against black defendants at the margin of release as originally modeled by Becker (1957, 1993). Bail judges could, for example, harbor explicit animus against black defendants that leads them to value the freedom of black defendants less than the freedom of observably similar white defendants. Bail judges could also harbor implicit biases against black defendants, leading to the relative over-detention of blacks despite the lack of any explicit animus. Racial animus may be a particular concern in bail setting due to the relatively low number of minority bail judges, the rapid-fire determination of bail decisions, and the lack of face-to-face contact between defendants and judges. Prior work has shown that it is exactly these types of settings where racial prejudice is most likely to translate into the disparate treatment of minorities (e.g., Greenwald et al. 2009).

One suggestive piece of evidence against this hypothesis is provided by the Anwar and Fang (2006) test of relative racial bias discussed above, which indicates that bail judges are monolithic in their treatment of white and black defendants. Consistent with these results, we also find that IV and MTE estimates of racial bias are similar among white and black judges. These estimates suggest that either racial animus is not driving our results or that black and white bail judges harbor equal levels of racial animus towards black defendants.

IV.B. Racially Biased Prediction Errors in Risk

A second explanation for our results is that bail judges are making racially biased prediction errors in risk, potentially due to inaccurate anti-black stereotypes. Bordalo et al. (2016) show, for example, that representativeness heuristics—probability judgments based on the most distinctive differences between groups—can exaggerate perceived differences between groups. In our setting, these kinds of race-based heuristics or anti-black stereotypes could lead bail judges to exaggerate the relative danger of releasing black defendants versus white defendants at the margin of release.

These race-based prediction errors could also be exacerbated by the fact that bail judges must make quick judgments on the basis of limited information and with virtually no training.

1. Representativeness of Black and White Defendants. We first explore whether our data are consistent with the formation of anti-black stereotypes that could lead to racially biased prediction errors. Extending Bordalo et al. (2016) to our setting, these anti-black stereotypes should only be present if blacks are overrepresented among the right tail of the predicted risk distribution relative to whites (both Hispanic and non-Hispanic). To test this idea, Figure III presents the distribution of the predicted risk of rearrest prior to case disposition calculated using the full set of crime and defendant characteristics, as well as the likelihood ratios, $\frac{\mathbb{E}(x|Black)}{\mathbb{E}(x|White)}$, throughout the risk distribution.[16] Results for each individual characteristic in our predicted risk measure are also presented in Online Appendix Table A10. Consistent with the potential formation of anti-black stereotypes, we find that black defendants are significantly underrepresented in the left tail of the predicted risk distribution and overrepresented in the right tail of the predicted risk distribution. For example, black defendants are 1.2 times less likely than whites to be represented among the bottom 25 percent of the predicted risk distribution, but 1.1 times more likely to be represented among the top 25 percent and 1.2 times more likely to be represented among the top five percent of the predicted risk distribution.

In Online Appendix F, we show that these black–white differences in the predicted risk distribution are large enough to rationalize the black–white differences in pretrial release rates under the Bordalo et al. (2016) stereotypes model. First, as a benchmark for the stereotypes model, we compute the fraction of black defendants that would be released if judges applied the same release threshold for whites to blacks. We rank-order both black and white defendants using our predicted risk measure, finding that 70.8 percent of black defendants would be released pretrial if judges use the white release threshold for both black and white defendants. By comparison, only 68.8 percent of black defendants are actually released pretrial. Thus, to rationalize the black–white difference in release rates, the stereotypes model will require that judges believe that black defendants are riskier than they actually are.

In the stereotypes model, judges form beliefs about the distribution of risk through a representativeness-based discounting model, where the weight attached to a given risk type t is increasing in the representativeness of t. Formally, let $\pi_{t,r}$ be the probability that a defendant of race r is in risk category t. The stereotyped beliefs for black defendants, $\pi^{st}_{t,B}$, is given by:

$$\pi^{st}_{t,B} = \pi_{t,B} \frac{\left(\frac{\pi_{t,B}}{\pi_{t,W}}\right)^{\theta}}{\sum_{s \in T} \pi_{s,B} \left(\frac{\pi_{s,B}}{\pi_{s,W}}\right)^{\theta}} \quad (17)$$

[16] Our measures of representativeness and predicted risk may be biased if judges base their decisions on variables that are not observed by the econometrician (e.g., demeanor at the bail hearing). Following Kleinberg et al. (2018), we can test for the importance of unobservables in bail decisions by splitting our sample into a training set to generate the risk predictions and a test set to test those predictions. We find that our measure of predicted risk from the training set is a strong predictor of true risk in the test set, indicating that our measure of predicted risk is not systematically biased by unobservables (see Online Appendix Figure A3).

where θ captures the extent to which representativeness distorts beliefs and the representativeness ratio, $\frac{\pi_{t,B}}{\pi_{t,W}}$, is equal to the probability a defendant is black given risk category t divided by the probability a defendant is white given risk category t.

Using the definition of $\pi^{st}_{t,B}$ from Equation (17), we can calculate the full stereotyped risk distribution for black defendants under different values of θ. For each value of θ, we can then calculate the implied release rate for black defendants under the above assumption that judges use the white release threshold for both black and white defendants. By iterating over different values of θ, we can then find the level of θ that equates the implied and true release rates for black defendants. Using this approach, we find that $\theta = 1.9$ can rationalize the true average release rate for blacks. To understand how far these beliefs are from the true distribution of risk, we plot the stereotyped distribution for blacks with $\theta = 1.9$ alongside the true distribution of risk for blacks in Online Appendix Figure A4. The mean predicted risk is 0.235 under the true distribution of risk for blacks, compared to 0.288 under the stereotyped distribution for blacks with $\theta = 1.9$.[17] These results indicate that a relatively modest shift in the true risk distribution for black defendants is sufficient to explain the large racial disparities we observe in our setting. See Online Appendix F for additional details on the stereotypes model and these calculations.

Further evidence in support of anti-black stereotypes comes from a comparison of the crime-specific distributions of risk. Black defendants are most overrepresented in the right tail of the predicted risk distribution for new violent crimes (see Online Appendix Figure A5), where we also tend to find strong evidence of racial bias.

A final piece of evidence in support of stereotyping comes from a comparison of the Hispanic and black distributions of risk relative to the non-Hispanic white distribution. Recall that we find no evidence of racial bias against Hispanic defendants (see Online Appendix Table A2). Consistent with the stereotyping model, we also find that the risk distributions of Hispanic and white defendants overlap considerably. In contrast, the risk distribution for blacks is shifted to the right relative to both the Hispanic and white distributions (see Online Appendix Figure A6). Thus, all of our results are broadly consistent with bail judges making race-based prediction errors due to anti-black stereotypes and representativeness-based thinking, which in turn leads to the overdetention of black defendants at the margin of release.

2. Racial Bias and Prediction Errors in Risk. We can also test for race-based prediction errors by examining situations where prediction errors of any kind are more likely to occur. One such test for race-based prediction errors uses a comparison of experienced and inexperienced judges. When a defendant violates the conditions of release, such as by committing a new crime, he or she is taken into custody and brought to court for a hearing during which a bail judge decides whether to revoke bail. As a result, judges may be less likely to rely on inaccurate racial stereotypes as they acquire greater on-the-job experience, at least in settings with limited information and contact. Consistent with this idea, we find that more experienced bail judges are more likely to release defendants, but

[17] Our estimate of θ is quantitatively similar to the magnitude of stereotypes in explaining investor overreaction to stock market news and the formation of credit cycles (Bordalo et al. 2018; Bordalo et al. 2017).

not make misclassification errors (see Online Appendix Figure A7). In contrast, while it appears plausible that race-based prediction errors will decrease with experience, there is no reason to believe that racial animus will change with experience.

To test this idea, columns (1)–(4) of Table V presents our estimates of racial bias, D^{IV} and D^{MTE}, separately by court. Although we caution that there are likely many differences in the criminal justice systems of the two cities in our sample, one distinction is the degree to which bail judges specialize in conducting bail hearings. In Philadelphia, bail judges are full-time judges who specialize in setting bail 24 hours a day, seven days a week, hearing an average of 5,253 cases each year. Conversely, the Miami bail judges in our sample are part-time generalists who work as trial court judges on weekdays and assist the bail court on weekend, hearing an average of only 179 bail cases each year. Consistent with racially biased prediction errors being more common among inexperienced judges, we find that racial bias is higher in Miami than Philadelphia ($p = .325$ for IV, $p = .442$ for MTE). In Miami, we find that marginally released white defendants are 25.1 percentage points more likely to be rearrested using our IV estimator ($p = .027$) and 24.9 percentage points more likely to be rearrested using our MTE estimator ($p = .040$), compared to marginally released black defendants. In Philadelphia, we find no statistically significant evidence of racial bias under either our IV or MTE estimates, suggesting the possible importance of experience in alleviating any prediction errors.[18]

Columns (5)–(8) of Table V provide additional evidence on this issue by exploiting the substantial variation in the experience profiles of the Miami bail judges in our sample. Splitting by the median number of years hearing bail cases, the average experienced Miami judge has 9.5 years of experience working in the bail system, while the average inexperienced Miami judge has only 2.5 years of experience. Consistent with our across-court findings, we find suggestive evidence that inexperienced judges are more racially biased than experienced judges ($p = .193$ for IV, $p = .095$ for MTE). Among inexperienced judges, we find that marginally released white defendants are 48.7 percentage points more likely to be rearrested using our IV estimator ($p = .040$) and 51.0 percentage points more likely to be rearrested using our MTE estimator ($p = .029$), compared to marginally released black defendants. Among experienced judges, we find no statistically significant evidence of racial bias under either our IV or MTE estimates.

Taken together, our results suggest that bail judges make racially biased prediction errors in risk. In contrast, we find limited evidence in support of the hypothesis that bail judges harbor racial animus towards black defendants. These results are broadly consistent with recent work by Kleinberg et al. (2018) showing that bail judges make significant prediction errors in risk for all defendants, perhaps due to overweighting the most salient case and defendant characteristics such as race and the nature of the charged offense. Our results also provide additional support for the

[18]Our IV estimate of racial bias in Philadelphia should be interpreted with some caution given that we only observe seven judges for this city in our data. The maximum infra-marginality bias of our IV estimator in Philadelphia is 16.4 percentage points, compared to only 1.6 percentage points in Miami-Dade. We note, however, that there is no infra-marginality bias of our MTE estimator for either city if we have correctly specified the shape of the MTE function.

stereotyping model developed by Bordalo et al. (2016), which suggests that probability judgments based on the most distinctive differences between groups—such as the significant overrepresentation of blacks relative to whites in the right tail of the risk distribution—can lead to anti-black stereotypes and, as a result, racial bias against black defendants.

V. Conclusion

In this paper, we test for racial bias in bail setting using the quasi-random assignment of bail judges to identify pretrial misconduct rates for marginal white and marginal black defendants. We find evidence that there is substantial bias against black defendants, ruling out statistical discrimination as the sole explanation for the racial disparities in bail. Our estimates are nearly identical if we account for observable crime and defendant differences by race, indicating that our results cannot be explained by black–white differences in the probability of being arrested for certain types of crimes (e.g., the proportion of felonies versus misdemeanors) or black–white differences in defendant characteristics (e.g., the proportion of defendants with prior offenses versus no prior offenses).

We find several pieces of evidence consistent with our results being driven by racially biased prediction errors in risk, as opposed to racial animus among bail judges. First, we find that both white and black bail judges are racially biased against black defendants, a finding that is inconsistent with most models of racial animus. Second, we find that black defendants are sufficiently overrepresented in the right tail of the predicted risk distribution to rationalize observed racial disparities in release rates under a theory of stereotyping. Finally, racial bias is significantly higher among both part-time and inexperienced judges, and descriptive evidence suggests that experienced judges can better predict misconduct risk for all defendants. Taken together, these results are most consistent with a model of bail judges relying on inaccurate stereotypes that exaggerate the relative danger of releasing black defendants versus white defendants at the margin.

The findings from this paper have a number of important implications. If racially biased prediction errors among inexperienced judges are an important driver of black–white disparities in pretrial detention, providing judges with increased opportunities for training or on-the-job feedback could play an important role in decreasing racial disparities in the criminal justice system. Consistent with recent work by Kleinberg et al. (2018), our findings also suggest that providing judges with data-based risk assessments may also help decrease unwarranted racial disparities.

The empirical test developed in this paper can also be used to test for bias in other settings. Our test for bias is appropriate whenever there is the quasi-random assignment of decision makers and the objective of these decision makers is both known and well-measured. Our test can therefore be used to explore bias in settings as varied as parole board decisions, Disability Insurance applications, bankruptcy filings, and hospital care decisions.

Princeton University
Princeton University and NBER
Harvard Law School and NBER

Supplementary Material

An Online Appendix for this article can be found at The Quarterly Journal of Economics online. Data and code replicating tables and figures in this article can be found in (12), in the Harvard Dataverse, doi: 10.7910/DVN/REUOXC.

References

[1] Abrams, David S., Marianne Bertrand, and Sendhil Mullainathan, “Do Judges Vary in Their Treatment of Race?” *Journal of Legal Studies*, 41 (2012), 347-383.

[2] Aigner, Dennis J., and Glen G. Cain, “Statistical Theories of Discrimination in Labor Markets,” *ILR Review*, 30 (1977), 175-187.

[3] Alesina, Alberto, and Eliana La Ferrara, “A Test of Racial Bias in Capital Sentencing,” *American Economic Review*, 104 (2014), 3397-3433.

[4] Angrist, Joshua, and Iván Fernández-Val, “ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework,” in *Advances in Economics and Econometrics: Volume 3, Econometrics: Tenth World Congress*, Cambridge University Press, 51 (2013), 401-434.

[5] Angrist, Joshua, Kathryn Graddy, and Guido W. Imbens, “The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish,” *Review of Economic Studies*, 67 (2000), 499-527.

[6] Angrist, Joshua, Guido W. Imbens, and Donald Rubin, “Identification of Causal Effects Using Instrumental Variables,” *Journal of the American Statistical Association*, 91 (1996), 444-455.

[7] Antonovics, Kate, and Brian Knight, “A New Look at Racial Profiling: Evidence from the Boston Police Department,” *Review of Economics and Statistics*, 91 (2009), 163-177.

[8] Anwar, Shamena, and Hanming Fang, “An Alternative Test of Racial Bias in Motor Vehicle Searches: Theory and Evidence,” *American Economic Review*, 96 (2006), 127-151.

[9] ———, “Testing for Racial Prejudice in the Parole Board Release Process: Theory and Evidence,” *Journal of Legal Studies*, 44 (2015), 1-37.

[10] Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson, “The Impact of Jury Race in Criminal Trials,” *Quarterly Journal of Economics*, 127 (2012), 1017-1055.

[11] Arnold, David, Will Dobbie, and Crystal S. Yang, “Racial Bias in Bail Decisions,” NBER Working Paper No. 23421, 2017.

[12] Arnold, David, Will Dobbie, and Crystal S. Yang, “Replication Data for: ‘Racial Bias in Bail Decisions’,” Harvard Dataverse (2018), doi: 10.7910/DVN/REUOXC.

[13] Arrow, Kenneth, “The Theory of Discrimination,” in *Discrimination in Labor Markets*, 3 (1973), 3-33.

[14] Ayres, Ian, “Outcome Tests of Racial Disparities in Police Practices,” *Justice Research and Policy*, 4 (2002), 131-142.

[15] Ayres, Ian, and Joel Waldfogel, “A Market Test for Race Discrimination in Bail Setting,” *Stanford Law Review*, 46 (1994), 987-1047.

[16] Becker, Gary S., *The Economics of Discrimination* (Chicago: University of Chicago Press, 1957).

[17] ———, "Nobel Lecture: The Economic Way of Looking at Behavior," *Journal of Political Economy*, 101 (1993), 385-409.

[18] Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad, "Incarceration, Recidivism and Employment," NBER Working Paper No. 22648, 2016.

[19] Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Stereotypes," *Quarterly Journal of Economics*, 131 (2016), 1753-1794.

[20] Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, "Diagnostic Expectations and Stock Returns," NBER Working Paper No. 23863, 2017.

[21] Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Diagnostic Expectations and Credit Cycles," *Journal of Finance*, 73 (2018), 199-227.

[22] Brinch, Christian N., Magne Mogstad, and Matthew Wiswall, "Beyond LATE with a Discrete Instrument," *Journal of Political Economy*, 125 (2017), 985-1039.

[23] Brock, William A., Jane Cooley, Steven N. Durlauf, and Salvador Navarro, "On the Observational Implications of Taste-Based Discrimination in Racial Profiling," *Journal of Econometrics*, 166 (2012), 66-78.

[24] Bushway, Shawn D., and Jonah B. Gelbach, "Testing for Racial Discrimination in Bail Setting Using Nonparametric Estimation of a Parametric Model," Unpublished Working Paper, 2011.

[25] Card, David, "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, 3 (1999), 1801-1863.

[26] Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg, "From LATE to MTE: Alternative Methods for the Evaluation of Policy Interventions," *Labour Economics*, 41 (2016), 47-60.

[27] Dahl, Gordon B., Andreas Ravndal Kostøl, and Magne Mogstad, "Family Welfare Cultures," *The Quarterly Journal of Economics*, 129 (2014), 1711-1752.

[28] DiNardo, John, Nicole Fortin, and Thomas Lemieux, "Labor Market Institutions and The Distribution of Wages, 1973-1993: A Semi-Parametric Approach," *Econometrica*, 64 (1996), 1001-1044.

[29] Dobbie, Will, Jacob Goldin, and Crystal Yang, "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges," *American Economic Review*, 108 (2018), 201-240.

[30] Doyle, Joseph, "Child Protection and Child Outcomes: Measuring the Effects of Foster Care," *American Economic Review*, 97 (2007), 1583-1610.

[31] Evdokimov, Kirill S., and Michal Kolesár, "Inference in Instrumental Variables Analysis with Heterogeneous Treatment Effects," Unpublished Working Paper, 2018.

[32] Foote, Caleb, James Markle and Edward Woolley, "Compelling Appearance in Court: Administration of Bail in Philadelphia," *University of Pennsylvania Law Review*, 102 (1954), 1031-1079.

[33] Frölich, Markus, "Nonparametric IV Estimation of Local Average Treatment Effects with Covariates," *Journal of Econometrics*, 139 (2007), 35-75.

[34] Fryer, Roland G., "An Empirical Analysis of Racial Differences in Police Use of Force," NBER Working Paper No. 22399, 2016.

[35] Fryer, Roland G., and Matthew Jackson, "A Categorical Model of Cognition and Biased Decision-Making," *The B.E. Journal of Theoretical Economics*, 8 (2008), 1-42.

[36] Goldkamp, John S., and Michael R. Gottfredson, "Development of Bail/Pretrial Release Guidelines in Maricopa County Superior Court, Dade County Circuit Court and Boston Municipal Court. Washington, D.C," *National Institute of Justice*, 1988.

[37] Goncalves, Felipe, and Steven Mello, "A Few Bad Apples? Racial Bias in Policing." Unpublished Working Paper, 2018.

[38] Greenwald, Anthony G., T. Andrew Poehlman, Eric L. Uhlmann, and Mahzarin R. Banaji, "Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity," *Journal of Personality and Social Psychology*, 97 (2009), 17-41.

[39] Gruber, Jonathan, Phillip Levine, and Douglas Staiger, "Abortion Legalization and Child Living Circumstances: Who is the "Marginal Child"?" *Quarterly Journal of Economics*, 114 (1999), 263-291.

[40] Gupta, Arpit, Christopher Hansman, and Ethan Frenchman, "The Heavy Costs of High Bail: Evidence from Judge Randomization," *Journal of Legal Studies*, 45 (2016), 471-505.

[41] Heckman, James J., Sergio Urzua, and Edward Vytlacil, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *The Review of Economics and Statistics*, 88 (2006), 389-432.

[42] Heckman, James J., and Edward Vytlacil, "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects. Proceedings of the National Academy of Sciences," 96 (1999), 4730-4734.

[43] ———, "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73 (2005), 669-738.

[44] Imbens, Guido W., and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62 (1994), 467-475.

[45] Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan, "Human Decisions and Machine Predictions," *Quarterly Journal of Economics*, 133 (2018),

237-293.

[46] Knowles, John, Nicola Persico, and Petra Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 109 (2001), 203-229.

[47] Kowalski, Amanda E., "Doing More When You're Running LATE: Applying Marginal Treatment Effect Methods to Examine Treatment Effect Heterogeneity in Experiments," NBER Working Paper No. 22363, 2016.

[48] Leslie, Emily, and Nolan G. Pope, "The Unintended Impact of Pretrial Detention on Case Outcomes: Evidence from NYC Arraignments," *Journal of Law and Economics*, 60 (2017), 529-557.

[49] McIntyre, Frank, and Shima Baradaran, "Race, Prediction, and Pretrial Detention," *Journal of Empirical Legal Studies*, 10 (2013), 741-770.

[50] Mechoulan, Stéphane, and Nicolas Sahuguet, "Assessing Racial Disparities in Parole Release," *Journal of Legal Studies*, 44 (2015), 39-74.

[51] Phelps, Edmund S., "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62 (1972), 659-661.

[52] Rehavi, M. Marit, and Sonja B. Starr, "Racial Disparity in Federal Criminal Sentences," *Journal of Political Economy*, 122 (2014), 1320-1354.

[53] Shubik-Richards, Claire, and Don Stemen, "Philadelphia's Crowded, Costly Jails: The Search for Safe Solutions," *Technical Report, Pew Charitable Trusts Philadelphia Research Initiative*, 2010.

[54] Stevenson, Megan, "Distortion of Justice: How the Inability to Pay Bail Affects Case Outcomes," Unpublished Working Paper, 2016.

TABLE I

DESCRIPTIVE STATISTICS

	All Defendants		White		Black	
	Released	Detained	Released	Detained	Released	Detained
Panel A: Bail Type	(1)	(2)	(3)	(4)	(5)	(6)
Release on Recognizance	0.258	0.000	0.269	0.000	0.249	0.000
Nonmonetary Bail w/ Conditions	0.195	0.030	0.203	0.033	0.189	0.028
Monetary Bail	0.547	0.970	0.527	0.967	0.562	0.972
Bail Amount ($1,000s)	13.235	35.286	11.957	24.782	14.180	42.227
Panel B: Defendant Characteristics						
Male	0.811	0.893	0.796	0.890	0.822	0.895
Age at Bail Decision	33.911	35.092	34.070	36.296	33.794	34.296
Prior Offense in Past Year	0.287	0.466	0.272	0.464	0.299	0.466
Arrested on Bail in Past Year	0.185	0.262	0.181	0.256	0.188	0.266
Failed to Appear in Court in Past Year	0.071	0.057	0.070	0.054	0.071	0.059
Panel C: Charge Characteristics						
Number of Offenses	2.722	3.162	2.544	2.587	2.854	3.541
Felony Offense	0.482	0.538	0.450	0.473	0.506	0.581
Misdemeanor Only	0.518	0.462	0.550	0.527	0.494	0.419
Any Drug Offense	0.390	0.260	0.373	0.244	0.403	0.271
Any DUI Offense	0.084	0.007	0.091	0.007	0.079	0.007
Any Violent Offense	0.310	0.331	0.288	0.241	0.326	0.390
Any Property Offense	0.238	0.387	0.237	0.406	0.239	0.376
Panel D: Outcomes						
Rearrest Prior to Disposition	0.237	0.042	0.226	0.037	0.245	0.045
Rearrest Drug Crime	0.111	0.006	0.106	0.005	0.115	0.006
Rearrest Property Crime	0.086	0.022	0.082	0.022	0.089	0.022
Rearrest Violent Crime	0.078	0.021	0.061	0.013	0.091	0.026
Failure to Appear in Court (Phl only)	0.258	0.006	0.250	0.006	0.264	0.007
Failure to Appear in Court or Rearrest	0.348	0.044	0.325	0.039	0.366	0.048
Observations	178,765	77,488	76,015	30,831	102,750	46,657

Notes. This table reports descriptive statistics for the sample of defendants from Philadelphia and Miami-Dade counties. The sample consists of bail hearings that were quasi-randomly assigned from Philadelphia between 2010–2014 and from Miami-Dade between 2006–2014, as described in the text. Information on race, gender, age, and criminal outcomes is derived from court records. Released is defined as being released at any point before trial. Detained is defined as never being released before trial. Bail amount (in $1,000s) includes zeros. Failure to appear in court is defined only in Philadelphia. See Online Appendix D for additional details on the sample and variable construction.

Table II

First-Stage Results

	All Defendants		White		Black	
	(1)	(2)	(3)	(4)	(5)	(6)
Pretrial Release	0.405***	0.389***	0.373***	0.360***	0.434***	0.415***
	(0.027)	(0.025)	(0.036)	(0.032)	(0.036)	(0.033)
	[0.698]	[0.698]	[0.711]	[0.711]	[0.688]	[0.688]
Court x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	No	Yes	No	Yes	No	Yes
Observations	256,253	256,253	106,846	106,846	149,407	149,407

Notes. This table reports the first-stage relationship between pretrial release and judge leniency. The regressions are estimated on the sample as described in the notes to Table 1. Judge leniency is estimated using data from other cases assigned to a bail judge in the same year, constructed separately by defendant race, following the procedure described in Section II.B. All regressions include court-by-time fixed effects. Baseline controls include race, gender, age, whether the defendant had a prior offense in the past year, whether the defendant had a prior history of pretrial crime in the past year, whether the defendant had a prior history of failure to appear in the past year, the number of charged offenses, indicators for crime type (drug, DUI, property, violent, and other), crime severity (felony and misdemeanor), and indicators for any missing controls. The sample mean of the dependent variable is reported in brackets. Robust standard errors two-way clustered at the individual and judge-by-shift level are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

TABLE III

TEST OF RANDOMIZATION

	All		White		Black	
	Pretrial Release	Judge Leniency	Pretrial Release	Judge Leniency	Pretrial Release	Judge Leniency
	(1)	(2)	(3)	(4)	(5)	(6)
Male	−0.09424***	−0.00005	−0.08593***	0.00004	−0.10379***	−0.00014
	(0.00235)	(0.00024)	(0.00325)	(0.00038)	(0.00323)	(0.00031)
Age at Bail Decision	−0.01725***	−0.00009	−0.02250***	−0.00015	−0.01512***	−0.00005
	(0.00086)	(0.00009)	(0.00127)	(0.00016)	(0.00104)	(0.00010)
Prior Offense in Past Year	−0.14922***	−0.00017	−0.16817***	0.00030	−0.13411***	−0.00044
	(0.00287)	(0.00028)	(0.00445)	(0.00046)	(0.00362)	(0.00036)
Arrested on Bail in Past Year	0.01066***	0.00004	0.01967***	−0.00166***	0.00495	0.00116***
	(0.00355)	(0.00034)	(0.00552)	(0.00057)	(0.00439)	(0.00042)
Failed to Appear in Court in Past Year	0.03318***	0.00012	0.03253***	0.00104**	0.03245***	−0.00047
	(0.00413)	(0.00025)	(0.00631)	(0.00043)	(0.00529)	(0.00031)
Number of Offenses	−0.02090***	−0.00001	−0.01829***	−0.00002	−0.02131***	0.00000
	(0.00053)	(0.00003)	(0.00085)	(0.00006)	(0.00063)	(0.00004)
Felony Offense	−0.17618***	−0.00003	−0.18817***	−0.00014	−0.16948***	0.00004
	(0.00257)	(0.00012)	(0.00397)	(0.00020)	(0.00323)	(0.00014)
Any Drug Offense	0.03514***	−0.00038	0.02558***	−0.00002	0.04069***	−0.00063*
	(0.00258)	(0.00026)	(0.00357)	(0.00039)	(0.00332)	(0.00032)
Any Property Offense	−0.04272***	−0.00013	−0.05560***	0.00009	−0.03188***	−0.00029
	(0.00285)	(0.00026)	(0.00388)	(0.00041)	(0.00354)	(0.00033)
Any Violent Offense	0.01640***	0.00028	0.07515***	0.00033	−0.02443***	0.00023
	(0.00389)	(0.00025)	(0.00497)	(0.00045)	(0.00429)	(0.00029)
Joint F-Test	[0.00000]	[0.60067]	[0.00000]	[0.21951]	[0.00000]	[0.08289]
Observations	256,253	256,253	106,846	106,846	149,407	149,407

Notes. This table reports reduced form results testing the random assignment of cases to bail judges. The regressions are estimated on the sample as described in the notes to Table 1. Judge leniency is estimated using data from other cases assigned to a bail judge in the same year, constructed separately by defendant race, following the procedure described in Section II.B. Columns (1), (3), and (5) report estimates from an OLS regression of pretrial release on the variables listed and court-by-time fixed effects. Columns (2), (4), and (6) report estimates from an OLS regression of judge leniency on the variables listed and court-by-time fixed effects. The p-value reported at the bottom of the columns is for an F-test of the joint significance of the variables listed in the rows. Robust standard errors two-way clustered at the individual and the judge-by-shift level are reported in parentheses. ***=significant at 1 percent level, **=significant at 5 percent level, *=significant at 10 percent level.

Table IV
Pretrial Release and Criminal Outcomes

	IV Results			MTE Results		
	White	Black	D^{IV}	White	Black	D^{MTE}
Panel A: Rearrest for All Crimes	(1)	(2)	(3)	(4)	(5)	(6)
Rearrest Prior to Disposition	0.236***	0.014	0.222**	0.249***	0.017	0.231**
	(0.073)	(0.070)	(0.101)	(0.084)	(0.080)	(0.117)
	[0.172]	[0.182]	–	[0.172]	[0.182]	–
Panel B: Rearrest by Crime Type						
Rearrest for Drug Crime	0.067	0.019	0.047	0.074	−0.024	0.097
	(0.043)	(0.043)	(0.060)	(0.048)	(0.054)	(0.074)
	[0.077]	[0.081]	–	[0.077]	[0.081]	–
Rearrest for Property Crime	0.158***	−0.005	0.163**	0.149**	0.043	0.106
	(0.057)	(0.047)	(0.073)	(0.066)	(0.053)	(0.084)
	[0.065]	[0.068]	–	[0.065]	[0.068]	–
Rearrest for Violent Crime	0.079**	−0.000	0.080	0.082*	−0.001	0.083
	(0.039)	(0.042)	(0.058)	(0.044)	(0.050)	(0.068)
	[0.047]	[0.071]	–	[0.047]	[0.071]	–
Observations	106,846	149,407	–	106,846	149,407	–

Notes. This table reports estimates of racial bias in pretrial release based on rearrest prior to case disposition. Columns (1)–(2) report two-stage least squares results of the impact of pretrial release on the probability of pretrial misconduct separately by race, while column (3) reports the difference between the white and black two-stage least squares coefficients, or D^{IV} as described in the text. Columns (1)–(3) use IV weights for each specification and report robust standard errors two-way clustered at the individual and judge-by-shift level in parentheses. Columns (4)–(5) report the average marginal treatment effect of the impact of pretrial release on the probability of pretrial misconduct separately by race, while column (6) reports the difference between the white and black MTE coefficients, or D^{MTE} as described in the text. Columns (4)–(6) use equal weights for each judge and report bootstrapped standard errors clustered at the judge-by-shift level in parentheses. All specifications control for court-by-time fixed effects and defendant race, gender, age, whether the defendant had a prior offense in the past year, whether the defendant had a prior history of pretrial crime in the past year, whether the defendant had a prior history of failure to appear in the past year, the number of charged offenses, indicators for crime type (drug, DUI, property, violent, or other), crime severity (felony or misdemeanor), and indicators for any missing characteristics. The sample means of the dependent variables are reported in brackets. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Table V

Racial Bias in Pretrial Release by Judge Experience

	Judge Specialization				Judge Experience			
	Miami D^{IV}	Miami D^{MTE}	Phl D^{IV}	Phl D^{MTE}	Miami Low Exp D^{IV}	Miami Low Exp D^{MTE}	Miami High Exp D^{IV}	Miami High Exp D^{MTE}
Panel A: Rearrest for All Crimes	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rearrest Prior to Disposition	0.251**	0.249**	0.040	0.078	0.487**	0.510**	0.144	0.036
	(0.114)	(0.121)	(0.184)	(0.195)	(0.237)	(0.233)	(0.178)	(0.164)
	[0.149]	[0.149]	[0.194]	[0.194]	[0.148]	[0.148]	[0.152]	[0.152]
Panel B: Rearrest by Crime Type								
Rearrest for Drug Crime	0.053	0.103	0.008	0.015	0.141	0.185	−0.013	0.006
	(0.066)	(0.077)	(0.138)	(0.150)	(0.119)	(0.138)	(0.101)	(0.110)
	[0.057]	[0.057]	[0.092]	[0.092]	[0.057]	[0.057]	[0.057]	[0.057]
Rearrest for Property Crime	0.196**	0.127	−0.031	−0.014	0.296**	0.293*	0.146	0.035
	(0.084)	(0.096)	(0.110)	(0.199)	(0.140)	(0.163)	(0.111)	(0.124)
	[0.078]	[0.078]	[0.060]	[0.060]	[0.078]	[0.078]	[0.079]	[0.079]
Rearrest for Violent Crime	0.082	0.079	0.065	0.066	0.204	0.218*	0.032	−0.036
	(0.065)	(0.075)	(0.115)	(0.119)	(0.134)	(0.119)	(0.100)	(0.099)
	[0.050]	[0.050]	[0.067]	[0.067]	[0.048]	[0.048]	[0.051]	[0.051]
Observations	93,417	93,417	162,836	162,836	47,692	47,692	45,725	45,725

Notes. This table reports estimates of racial bias for different subgroups of judges. D^{IV} is the difference between the white and black two-stage least squares coefficient estimates of pretrial release on pretrial misconduct using IV weights. D^{MTE} is the difference between the average white and black MTE estimates of pretrial release on pretrial misconduct using equal weights by judge. Columns (1)–(2) report estimates for nonspecialist bail judges in Miami-Dade. Columns (3)–(4) report estimates for specialist bail judges in Philadelphia. Columns (5)–(8) report estimates for nonspecialist bail judges in Miami with below- and above-median years of experience. The sample is described in the notes to Table 1. The dependent variable is listed in each row. All specifications control for court-by-time fixed effects and defendant race, gender, age, whether the defendant had a prior offense in the past year, whether the defendant had a prior history of pretrial crime in the past year, whether the defendant had a prior history of failure to appear in the past year, the number of charged offenses, indicators for crime type (drug, DUI, property, violent, or other), crime severity (felony or misdemeanor), and indicators for any missing characteristics. The sample means of the dependent variables are reported in brackets. For IV specifications, robust standard errors two-way clustered at the individual and judge-by-shift level are reported in parentheses. For MTE specifications, bootstrapped standard errors clustered at the judge-by-shift level are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

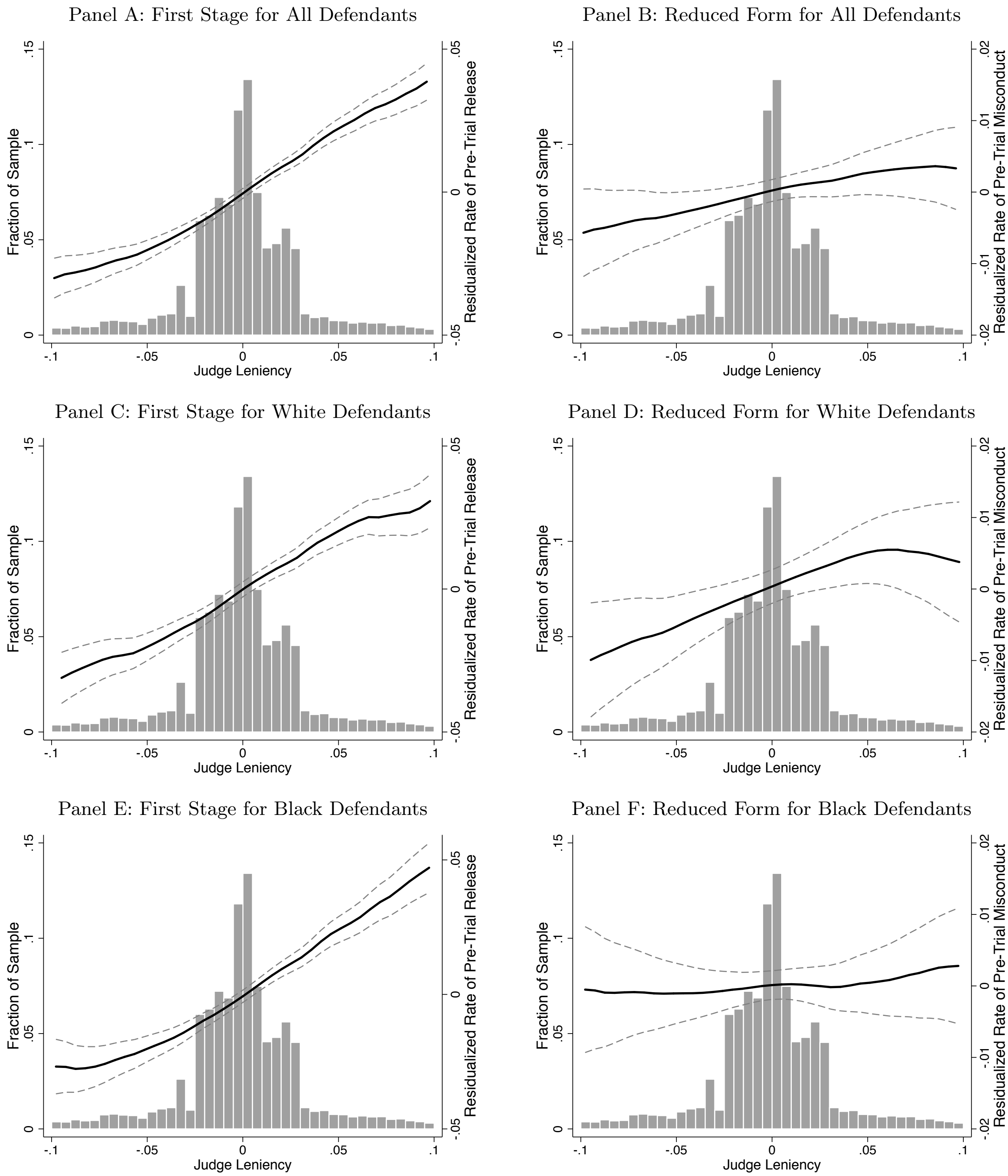

Figure I

First-Stage and Reduced Form Results

These figures report the first-stage and reduced form relationships between defendant outcomes and judge leniency. The regressions are estimated on the sample as described in the notes to Table 1. Judge leniency is estimated using data from other cases assigned to a bail judge in the same year, constructed separately by defendant race, following the procedure described in Section II.B. In the first-stage regressions, the solid line is a local linear regression of residualized pretrial release on judge leniency. In the reduced form regressions, the solid line is a local linear regression of residualized pretrial misconduct on judge leniency. Pretrial release and pretrial misconduct are residualized using court-by-time fixed effects in the full sample. Standard errors are two-way clustered at the individual and judge-by-shift level.

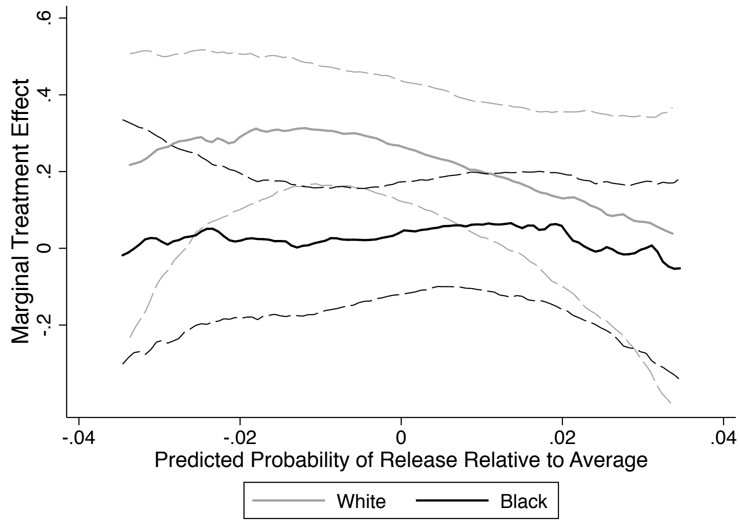

FIGURE II

Marginal Treatment Effects

This figure reports the marginal treatment effects (MTEs) of pretrial release on pretrial rearrest separately by race. To estimate each MTE, we first estimate the predicted probability of release using only judge leniency. We then estimate the relationship between the predicted probability of release and rearrest prior to disposition using a local quadratic estimator (bandwidth = 0.030). Finally, we use the numerical derivative of the local quadratic estimator to calculate the MTE at each point in the distribution. Standard errors are computed using 500 bootstrap replications clustered at the judge-by-shift level. See the text for additional details.

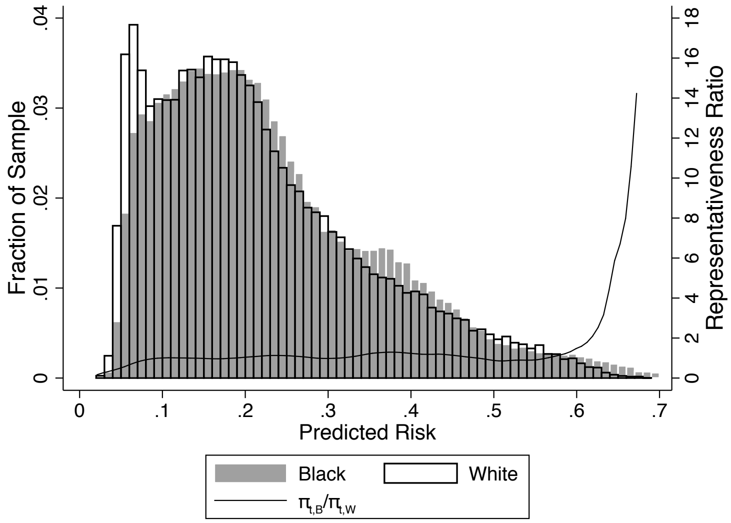

Figure III

Predicted Risk Distribution by Defendant Race

This figure reports the predicted distribution of pretrial misconduct risk separately by race. Pretrial misconduct risk is estimated using the machine learning algorithm described in Online Appendix F. The solid line represents the representativeness ratio for black versus white defendants as described in the text, or the estimated misconduct risk for blacks divided by the estimated misconduct risk for whites. See the text for additional details.