Are juries racially discriminatory? Evidence from the race-blind charging of grand jury defendants with and without racially distinctive names

Mark Hoekstra[1], Suhyeon Oh[2], Meradee Tangvatcharapong[3]

[1]Department of Economics, Baylor University; NBER; IZA

[2]Department of Economics, Texas A&M University

[3]Institute of Economic Research, Hitotsubashi University

December 2023

Abstract

To what extent are the large racial disparities in criminal justice caused by racial discrimination? We test for racial bias using over a quarter million felony cases from Texas in which grand juries decided whether or not the case should be prosecuted. We perform three tests of racial bias. Each exploits the fact that grand juries can only infer race rather than directly observe it, and that the distribution of "convictability" is similar for Black defendants with Black versus White names. All three approaches indicate an absence of racial bias. First, we show that conditional on defendant and case characteristics, there is little difference in the likelihood that juries push forward, or "true bill", Black defendants with distinctively Black names compared to Black defendants with White names. Similarly, the traditional outcome-based test of racial bias indicates that among the marginal Black defendants who were true-billed, identifiably-Black defendants were convicted at similar rates to Black defendants with White names. For the third approach, we first estimate disparate impact—defined as whether juries treat Black defendants differently from White defendants with the same underlying "convictability"—by exploiting the quasi-random assignment of cases to grand juries to purge omitted variable bias. Results indicate there is a similar disparate impact against Black defendants with White names compared to Black defendants with Black names, which indicates the absence of statistical and taste-based discrimination. Collectively, these findings indicate that juries did not engage in racial bias.

1 Introduction

Racial disparities are pervasive throughout the U.S. criminal justice system. Perhaps the most stark disparity is with respect to felony charging and conviction. Thirty-three percent of Black adult males have a felony conviction, compared to only 12.8 percent of the total adult male population (Shannon et al., 2017). Unsurprisingly, this large disparity in felony conviction manifests itself in similarly large differences in incarceration, where 32.3 percent of Black males are likely to go to state or federal prison, compared to only 5.9 percent of White males (Bonczar, 2003).

To what extent are these racial disparities due to discriminatory practices in the criminal justice system? The ideal experiment would be to compare how Black defendants are treated when their race is known, to how they would have been treated if they were believed to be White. Indeed, this is what motivates audit studies such as Bertrand and Mullainathan (2004) that experimentally alter the racial distinctiveness of names in order to test for racial discrimination. It is also the intuition underlying seminal studies like Goldin and Rouse (2000), who used blinded and unblinded comparisons to test for gender discrimination, and Grogger and Ridgeway (2006), who used the "veil of darkness" to test for racial bias in vehicle stops.[1]

In this paper, we test for racial bias in a context that we believe comes as close as possible to this ideal setting of blinded and unblinded comparisons. We do this using data on more than a quarter million felony cases heard by grand juries in Harris County, Texas, between February of 1990 and July of 2022. An important feature of this context is that while we directly observe race in the data, the grand jurors who hear a brief summary of the evidence observe only the defendant's name. This enables us to use whether the first or last name of the defendant is racially identifiable as Black as a measure of whether grand jurors believed the defendant was Black.[2] Importantly, we show that the distribution of "convictability" is similar for Black defendants with Black names, compared to Black defendants with White names.[3]

[1]There is a large literature using the veil of darkness to study racial profiling by police. See, for example, Horrace and Rohlin (2016), Pierson et al. (2020), Kalinowski et al. (2021), Worden, McLean, and Wheeler (2012), and Brewer (2023).

[2]We also show that results are robust to alternative ways of inferring perceived race.

[3]We first estimate the relationship between felony conviction and defendant and case characteristics using White

A second important feature of this setting is that because we have linked the grand jury records to whether the case ended in a felony conviction, we can use outcome-based approaches to test for racial bias. This contrasts with many contexts—including trial juries—where such tests are impossible to implement. Moreover, a unique feature of this context helps us avoid a common pitfall associated with all outcome-based tests, which is that any subsequent bias in the process can bias the test. This is of minimal concern here, however, because the *only* stage at which race is (partially) blinded is the grand jury stage. Felony conviction is determined later in a completely unblinded process, where prosecutors, judges, and trial juries directly observe the defendant's race. As a result, under the assumption that those actors do not discriminate on the basis of the racial distinctiveness of name, *separate from any discrimination on the basis of race itself*, this is not a concern in this context. Put differently, for our outcome-based tests, we need to assume that prosecutors would not treat Black defendants with Black names more harshly than otherwise-similar Black defendants with White names. We believe this assumption is reasonable, given that prosecutors directly observe race, as well as defendants' full criminal histories and the details of their alleged crimes in the case at hand.[4]

Aside from the methodological benefits of this context for assessing racial bias, there are two reasons why testing for racial bias *by juries* is of particular interest. First, criminal conviction decisions are either made directly by juries, or are negotiated by prosecutors and defendants under threat of what a jury would decide if the case went to trial—that is, "under the shadow of the law". To the extent juries are racially discriminatory, it would directly or indirectly impact every case in the criminal justice system. Second, since juries are composed from a panel of randomly selected citizens, understanding the extent to which juries are discriminatory provides a measure of whether the broader population is discriminatory. This is important not only for understanding racial bias

defendants who were true billed. We then use the resulting regression equation to predict the likelihood of conviction for each Black defendant.

[4]This contrasts with audit study settings, where race is unobserved, and therefore where a racially biased agent may well discriminate on the basis of name. In addition, in audit study settings, criminality (or productivity) is also unobserved. As a result, biased agents may well use name as a proxy for those factors (i.e., statistically discriminate), which is unlikely to be the case in our setting given that the details of the case and the full criminal history are observed by prosecutors and judges.

in the criminal justice system, but also more generally.

We implement three different empirical tests for racial bias, all of which indicate that grand juries do not discriminate against Black defendants. We begin by showing that conditional on defendant and case characteristics, there is little difference in the likelihood that grand juries true bill Black defendants with Black names, compared to Black defendants with White names. While this approach has the same limitations as any selection-on-observables approach, we get some comfort from the fact that the distributions of convictability—based on observed characteristics—are similar for Black defendants with Black versus White names.

The other two approaches are outcome-based tests that explicitly account for the possibility that there may be differences in convictability that are not reflected in data on case and defendant characteristics. This setting is well-suited for an outcome-based test because jurors are asked to consider only a single criterion: whether there is evidence of probable cause, which is a measure of “convictability". This differs from other contexts, where the objective of decision-makers is multi-faceted, and often unclear to the researcher. We first employ a traditional outcome-based test, originally proposed by Knowles, Persico, and Todd (2001). In our context featuring blinded and unblinded decisions, this means we compare the felony conviction rates of Black defendants with Black names to those of Black defendants with White names. In order to address the well-known problem of inframarginality, we focus explicitly on those Black defendants who were, based on the type of case and other characteristics, on the margin of being true-billed by the grand jury. Results indicate that among these marginal Black defendants, conviction rates were nearly identical for those with Black versus White names. In particular, we show that among those Black defendants whose predicted probability of being true-billed was in the bottom 10 percent of the distribution—the felony conviction rate was 60 percent for those with Black names, and a statistically equivalent 61 percent for those with White names.

In the third approach, we use the clever method proposed by Arnold, Dobbie, and Hull (2022) and used by Baron et al. (2023), combined with the blinded/unblinded context of grand jury cases, to estimate racial bias. Intuitively, we first estimate the underlying felony convictability of *all*

Black and White defendants by asking how often they would subsequently be convicted of a felony when facing a particularly tough grand jury that "true bills" every case that it sees. We then use this measure of true underlying felony convictability to purge omitted variable bias from the estimated racial disparity. The resulting estimate captures what Arnold, Dobbie, and Hull (2022) term disparate impact, which is the sum of racial bias, statistical discrimination, and other sources of non-race-based disparate impact.[5,6]

Results from this third approach indicate that grand jury decisions impose a disparate impact of 0.8 percentage points against Black defendants who are likely, and correctly, identified by grand jurors as such. Put differently, felony cases against Black defendants are 0.8 percentage points more likely to proceed than felony cases against White defendants with similar levels of convictability. Strikingly, however, results indicate a similar, if not slightly larger, disparate impact when comparing Black and White defendants who had similarly-White names, and were thus racially indistinguishable to grand jurors. That is, compared to racially-identifiable White defendants, Black defendants *who were likely (mis)perceived as White by grand jurors* were 0.9 percentage points more likely to have their felony cases pushed forward. The similarity of findings across the "unblinded" and "blinded" samples indicates that the disparate impact estimated between White and Black defendants is not caused by either taste-based or statistical discrimination, but rather is due to similar treatment on the basis of some other factor that differs across race. This is important because while taste-based and statistical discrimination are illegal, the legality of equal treatment on the basis of some other factor that differs across race—such as having a lower threshold for some types of cases, which may be more common among Black defendants—is more ambiguous.

This study contributes to the existing literature by combining blinded/unblinded comparisons and (two different) outcome-based tests, which are arguably the two most credible designs used to

[5]This method exploits the fact that cases are quasi-randomly assigned across grand juries, some of which are less lenient than others. As we describe later, our understanding of the assignment process is that it is quasi-random. Consistent with that understanding, we show empirically that grand jury leniency is uncorrelated with case and defendant characteristics.

[6]One advantage of this approach is that it does not require the monotonicity assumptions that are required for alternative approaches, such as Arnold, Dobbie, and Yang (2018).

assess racial bias. The advantage of exploiting blinded and unblinded comparisons in the context of the traditional outcome-based test is twofold. First, it reduces concerns that arise when the distributions of the two groups are different. In this context, we show that the characteristics and risk distributions are similar across Blacks with and without racially distinctive names, even though both are different from those of Whites. Second, it allows us to overcome concerns about bias in the test due to downstream racial bias, since downstream decisions are fully unblinded. In addition, the blinded/unblinded nature of decisions in this setting is also advantageous when purging omitted variable bias, as it allows for a simple and intuitive test to distinguish racial bias from non-race-based disparate impact, rather than impose a structural model as in Arnold, Dobbie, and Hull (2022).

In addition, the paper also contributes to the literature on jury bias. This includes the seminal paper by Anwar, Bayer, and Hjalmarsson (2012) that cleverly used random variation in jury panel composition, as well as subsequent studies that use similar approaches by Flanagan (2018) and Hoekstra and Street (2021). The advantage of this study is rather than examining the interaction of juror and defendant characteristics, as has been done in the existing literature, we are able to implement outcome-based tests in a setting in which defendant race is sometimes blinded, and sometimes not. This is the first study to use either of these approaches in the context of juries. These approaches enable us to estimate directly whether jurors treat defendants perceived as Black more harshly than those perceived as White. In addition, we are able to purge omitted variable bias and estimate disparate impact against Black defendants directly, as well as use the blinded comparison to estimate the component of disparate impact due to racial bias. Finally, our approach and sample provide us with vastly more statistical power to detect effects. In particular, we show that the minimum detectable effects in other studies from the impact of changing the race of only one in six jurors range from 23 percent to 78 percent. This amplifies the concerns expressed by Ioannidis and Doucouliagos (2017), which is that underpowered studies are more likely to be published if they happen to find a statistically significant result, as opposed to an imprecise null result. By comparison, the approach used in our study can detect an effect of only 0.17 percentage points.

Our findings have important implications. First and foremost, the absence of racial bias across three different tests suggests that racial bias by juries is not directly or indirectly responsible for the large racial disparities observed in the criminal justice system. More generally, results here provide evidence that the jurors drawn from a randomly selected panel of U.S. citizens do not engage in racial discrimination. While we do not claim that these jurors are perfectly representative of the population at large, or that none of the jurors are racially discriminatory, this does provide some comfort with respect to the amount of racial animus present in the U.S. population.

2 Background and Data

A unique feature of the criminal justice system in Texas is that every felony case must first go before a grand jury before it is prosecuted. Grand juries consist of 12 grand jurors and four alternates.[7] Each grand jury hears all types of felony cases; there is no specialization. After hearing a case, each juror must choose whether to "true bill" a case, in which the case proceeds forward through the system, or “no bill" a case, at which point the case is dropped. If at least 9 out of the 12 seated grand jurors vote to “true bill” a case, then it moves forward in the criminal justice system. If not, the case is “no billed”, which is the end of the case. As a result, at the conclusion of a grand jury, the case can i) move forward with at least one felony charge (if that charge is “true billed"); ii) move forward with at least one misdemeanor charge (if the felony charge is “no billed" but at least one misdemeanor charge is “true billed");, or iii) be closed at that point, with no further action taken. Grand jurors are asked to use a standard of “probable cause” when deciding whether to true bill a charge, which is a lower standard than the "beyond a reasonable doubt" standard used in jury trials.[8] The term for each grand jury is three months. Over that time period jurors meet twice per week, and make judgments on 50+ cases per day.

The Texas Code of Criminal Procedure mandates that only the presenting prosecutor is allowed to be in the room with the grand jurors, along with court reporters. In some cases, there may also

[7]Prior to 2015, the number of alternates was limited to two. However, with the passing of HB 2150, the number increased to four.

[8]The full ranking of the strength of evidence in the legal system, from lowest to highest, is reasonable suspicion, probable cause, preponderance of evidence, clear and convincing, and beyond a reasonable doubt.

be an expert witness, if one is needed to help explain the evidence to the grand jurors, or translators, if needed. Video and photographs are not usually shown; the exceptions are the more serious (and rare) cases.

There are significant constraints in scheduling cases before a grand jury. The main one is that Texas law requires the State to present the case within 90 days if the defendant is in custody, and all cases must be heard within 180 days. Thus, the main priority of each prosecutor is to make sure they meet these requirements. The typical process for assignment is that prosecutors who want to present one or more cases to the grand jury typically ask an administrative assistant in the Grand Jury Division for a certain day of the week that fits their schedule. The administrative assistant then randomly assigns the prosecutors to one of the two grand juries that meet on that day of the week. As a result, conditional on the timing of that case, cases could have been heard by one grand jury or another, which can differ with respect to their leniency. As a result, we control for year-by-month-by-week fixed effects throughout our analyses, though we note that quasi-random assignment of cases to grand juries is only required for the third empirical approach, which estimates disparate impact.

Another unique feature of the grand jury system is that because the defendant is not present at the hearing, the race or ethnicity of the defendant is not directly observed by the grand jurors. Similarly, while we do not have data on whether prosecutors somehow cue the race of the defendant, anecdotally prosecutors have been warned *not to* signal race or ethnicity in any way during the hearing, for fear of the case being thrown out on the basis of that later. However, grand jurors do observe the full name of the defendant. In addition, they sometimes observe the neighborhood in which the crime occurred, which may be indicative of the defendant's likely race or ethnicity. We address this in the robustness section, where we use a combination of name and residential location to measure how jurors likely perceive a defendant's race.

Our data include every grand jury case filed from February of 1990 through July of 2022. After excluding cases for defendants identified in the administrative records as neither Black nor White, there are a total of 695,500 cases.[9] We then link these cases to data on case disposition received from the Office of Harris County District Clerk. As a result, for each charge in each case—including felony charges—we observe whether that charge resulted in a conviction outcome.

These data provide three important advantages, relative to prior work. The first is that we have a measure of the outcome of interest, which is the only factor that grand jurors are told to consider in deciding how to vote on the case. That is, we observe whether defendants who were true billed were convicted of a felony later in the process. This is not possible in the setting of trial juries, because there is no objective way—apart from the decision of the jury—to know which cases cleared the "beyond a reasonable doubt" standard.[10] This means that we can implement both the standard outcome-based test proposed by Knowles, Persico, and Todd (2001), as well as the method for estimating disparate impact proposed by Arnold, Dobbie, and Hull (2022).

The second advantage of this setting is that while we directly observe race, grand jurors do not. Rather, some decisions on Black defendants are unblinded—because the Black defendant has a distinctively Black name —while others are effectively blinded because the Black defendant has a White name. This enables blinded versus unblinded comparisons, which are rare outside of experimental settings. To implement this, we use the first name and last name of the defendants in our sample to identify those who fall into one of three groups: White defendants with identifiably-White names, Black defendants with identifiably-Black names, and Black defendants with identifiably-White names. Formally, we identify the "Whiteness" and "Blackness" of names by using the "predictrace" package in R. We classify a Black defendant as having an identifiably-Black name if the probability of being Black is greater than 0.5 for either his first name or his last name. We classify defendants as having an identifiably-White name if the probability of being White is greater than 0.5 for both the first name and the last name.[11]

[9]We do not observe in the administrative data whether or not defendants are Hispanic.

[10]As a result, the literature on jury bias instead asks whether the interaction of juror and defendant race matters. The limitation of this is that there is even in the presence of a nonzero effect, it is difficult to know whether White decision-makers are biased for White defendants, against Black defendants, or if Black decision-makers are biased for Black defendants, or against White defendants, or some combination of the above.

[11]In the robustness section, we show that our findings are robust to using alternative, higher thresholds. In addition, we also report results in which the Whiteness and Blackness of names are based on last name and location, which accounts for the possibility that jurors may be able to infer race to the extent the location of the crime is demographically

Finally, the third advantage of studying racial bias in this context is practical. One challenge in assessing racial bias in other contexts—especially when using the more compelling research designs that use a fraction of the total variation in race—is statistical power. This is a particular challenge when studying trial juries since the vast majority of cases do not go to trial. This is evident in Figure 1 Panel A, which shows the number of cases studied by the seminal paper on racial bias by juries by Anwar, Bayer, and Hjalmarsson (2012), as well as subsequent papers on jury bias by Flanagan (2018) and Hoekstra and Street (2021). These papers utilized a total of 712, 737, and 1,481 cases, respectively. By comparison, even when we limit our analysis to those cases involving Black or White defendants with names that are identifiable as White or Black, we have a total of 315,995 “cases" representing a total of 350,560 felony charges.[12]

That stark difference in sample size, combined with the approach enabled by the use of outcome data, results in a vast improvement in the ability to detect effects. This is shown in Figure1 Panel B, which shows what the upper bound of the 95 percent confidence interval would be for a zero estimate in each study, relative to the mean of the outcome. For the first three studies, the treatment effect is defined as the effect of having one more juror of a different race or gender, out of six.[13] Figure 1 Panel B shows that while the seminal paper on jury bias would not be able to detect an effect of replacing one of six White jurors with a Black juror of smaller than 54.6 percent, we are able to detect effects as small as 0.12 percent. Similarly, Figure1 Panel C shows the minimum detectable

similar to one's residential neighborhood. Results from this approach are summarized in Panel C of Table 7, while the full set of results is shown in Appendix A. This approach uses the wru package in the statistical software R, which uses data in which name, race, and residential information are known to predict the probability that an individual is White, Black, Hispanic, Asian, or Other race/ethnicity. We classify someone as White if the predicted probability of White is greater than 0.5, and classify someone as having a Black-sounding name (and location) if the predicted probability of Black is greater than 0.5.

[12]Each defendant-charge in our data is given a unique case number. However, some defendants are associated with multiple cases heard by the same grand jury on the same day, which arose from the same event. To make an apples-to-apples comparison to the three other studies, which have multiple charges per defendant-case, For the purpose of Figure 1 we count situations in which a defendant has multiple charges heard by the same grand jury on the same day as a single “case". Throughout our analyses we keep our data at the defendant-charge level (for which we observe a unique “case" number in the administrative data), and weight each observation by the inverse of the number of charges heard on that defendant on that particular day.

[13]One could argue that to make those studies comparable to this one, which simply compares Black and White defendants, one should rescale the MDE for these other studies to be the effect of going from 0 to 6 out of 6 jurors of a different race or gender. Doing so would require multiplying each of those estimates by 6, which would obviously make the difference in MDE even more stark.

effect (MDE) at 80 percent power for all four studies.[14] The minimum detectable effect in the Anwar, Bayer, and Hjalmarsson (2012), Flanagan (2018), and Hoekstra and Street (2021) papers are 78 percent, 23 percent, and 26 percent, respectively. By comparison, the minimum detectable effect in this study is 0.17 percent. This has important implications, as Ioannidis and Doucouliagos (2017) question whether underpowered papers would publish if they were to find an imprecise, null effect, rather than a very large, statistically significant one. In addition, we note that even with a sample of our size, statistical power can be an issue using other approaches. For example, if we use the methodology proposed by Arnold, Dobbie, and Yang (2018), we estimate a statistical zero with a standard error of 8 percentage points. We view this as an uninformative, imprecise zero, and one that we suspect would go unpublished.

Summary statistics are shown in Table 1. Column (1) shows results for the full sample of 695,500 grand jury cases filed between February of 1990 and July of 2022. It shows that 53 percent of defendants are White, while 47 percent are Black. Eighty-two percent of defendants are male, with an average age of 32. With respect to charge characteristics, 12.1 percent of charges are 1st degree felonies, 22.5 percent are 2nd degree felonies, and 29.6 percent are 3rd degree felonies. The average number of charges per case is 1.15.

With respect to outcomes, 97.1 percent of the cases are true billed. Sixty-two percent of cases result in a felony conviction. Conditional on a true bill from the grand jury, 64 percent of cases result in a felony conviction.

Column (2) of Table 1 shows summary statistics for White defendants with racially identifiable White names, Black defendants with racially identifiable Black names, and Black defendants with racially identifiable White names. This is the sample used throughout the analysis. Defendants in this sample are slightly older (33.9 versus 32.4 years old for the full sample), and are slightly more likely to have had a prior offense (70.1 versus 66.8 percent). The seriousness of the felony charges faced is similar to that of the full sample, as is the number of charges. In addition, the true bill rate for this sample of 35,560 defendants is identical to the full sample (97.1 percent).

[14] This is defined as 2.8 times the standard error, as shown in Ioannidis and Doucouliagos (2017).

Column (3) of Table 1 shows summary statistics for White defendants with racially identifiable White names. It shows that the predicted likelihood of being Black for these individuals is only 15 percent. Column (4) shows summary statistics for Black defendants with racially identifiable Black names. The predicted likelihood of being Black for these individuals is 70 percent. Thus, it is clear that for the set of White and Black defendants shown in Columns (3) and (4), race is highly identifiable even without directly observing the individual or their race. This is important, because a comparison of these two samples enables us to assess the extent to which identifiably-Black defendants are treated differently than identifiably-White defendants for whom felony convictability is similar.

Column (5) shows summary statistics for Black defendants with racially identifiable White names. For these individuals, the predicted likelihood of being Black is only 26 percent, which is substantively lower than for the Black defendants with identifiably-Black names (70 percent), and only somewhat higher than for White defendants with identifiably-White names (15 percent).

Columns (3) and (4) of Table 1 also show that there are substantive differences between defendants that jurors would likely identify as Black and White. Black defendants shown in Column (3) are somewhat younger (32.4 versus 34.6 years old), and are more likely to have a prior offense (74 versus 62 percent). In addition, the identifiably-Black defendants in Column (4) face more serious charges; they are more likely to face a 1st degree felony charge (12.6 versus 8.1 percent), slightly more likely to face a second degree felony charge (22.6 versus 20.5 percent), and less likely to face a 3rd degree felony charge (26.9 versus 31.9 percent). Identifiably-Black defendants are true billed at somewhat higher rates than White defendants (97.4 versus 96.6 percent), and are more likely to be convicted of a felony conditional on a true bill outcome (64.9 versus 64.5 percent).

In contrast to the substantive differences between the cases of identifiably-White and identifiably-Black defendants, there are at most few such differences between the identifiably-Black defendants in Column (4), and the Black defendants in Column (5) who are likely to be be (incorrectly) perceived as White. Of cases with identifiably-Black (identifiably-White) defendants, 12.6 (12.7), 22.6 (22.8), and 26.9 (27.1) percent are 1st, 2nd, and 3rd degree felonies, respectively.

In addition, we can also show the similarity in the distributions of risk between Black defendants with Black versus White names. To do so, we first predict the likelihood of a felony conviction outcome for each Black defendant in our sample. We do so by first estimating the relationship between felony conviction and all case and defendant characteristics for *White* defendants who were true-billed. We then use that out-of-sample equation to predict the likelihood of a felony conviction for the *Black* defendants in our sample.

Results are shown in Figure 2. Consistent with Columns (4) and (5) of Table 1, Figure 2 Panel A shows that the densities are nearly identical. We also fail to reject the null hypothesis that the distribution functions underlying the two samples are identical.[15] Indeed, Figure 2 looks exactly how one would expect if the "Blackness" of names were randomly assigned among grand jury defendants. This similarity among Black defendants with and without racially distinctive names is helpful for all three of our empirical approaches. For the selection-on-observables approach comparing true-bill rates, it gives us some comfort that if the observable characteristics and distributions of risk are similar across Black defendants with and without distinctively Black names, perhaps unobserved characteristics are also similar. Similarly, the traditional outcome-based test can fail if the distributions of the two groups are different. Figure 2 suggests this is unlikely to be the case, especially—and most importantly—when it comes to the left tail of the distribution, which are the defendants most likely to be on the margin of being true-billed. Finally, this is also helpful for our third approach, which estimates the disparate impact against Black defendants, relative to White ones. In particular, in comparing disparate impacts against Black defendants with Black versus White names to distinguish racial bias from non-race-based disparate impact, we assume that the latter would be similar across both groups. Columns (4) and (5) of Table 1, and Figure 2, suggest that this is a reasonable assumption.

[15] We do so using the Epps-Singleton test, which is preferable to the Kolmogorov-Smirnov test because the latter is unable to address ties, which are relatively common in our setting.

3 Methodology

3.1 Blinded/Unblinded Comparison of True Bill Rates

We begin by comparing the true bill rates of Black defendants with and without racially distinctive names. In the absence of racial bias, we expect to see little difference.

To implement this test, we estimate the following for the sample of Black defendants:

$$Truebill_i = \alpha + \beta IdentifiablyBlack_i + \rho_t + u_i \tag{1}$$

Where *Truebill* is an indicator equal to one if the grand jury voted to push the case forward for prosecution, *IdentifiablyBlack* is an indicator equal to one if the defendant has a Black name, and ρ is a year-month-week fixed effect. In addition, we also show results when we control for individual charge fixed effects as well as defendant gender, age, and criminal history, as measured by whether the defendant is observed in a previous case in the data.

3.2 Traditional Outcome-Based Test

In addition, we also perform a traditional outcome-based test. The intuition of the test is that if grand juries are racially discriminatory against Black defendants, then they will use a lower threshold for true billing those defendants, compared to otherwise-similar White defendants. This racially biased behavior would then manifest itself through lower subsequent felony conviction rates for (marginal) true-billed Black defendants, compared to Whites. Put differently, a racially biased grand jury will true bill more Black defendants for whom the evidence—which is observed to them, but not to us—is flimsy. This will become evident later in the process when true-billed Black defendants are convicted at lower rates.

We implement a modified version of this traditional test by comparing entirely within Black defendants. In particular, we compare Black defendants who are likely perceived as Black by grand jurors, to Black defendants who have White names and are thus likely perceived as White. We do so for two reasons. The first is to address the so-called “inframarginality" problem. This refers to the fact that the traditional test can fail to the extent the distributions of the two groups differ. While we can address that in part by focusing on defendants with the lowest predicted conviction rates, we also address it by exploiting the fact that as shown in Figure 2 and Columns (4) and (5) of Table 1, the conviction risk distribution is similar across these two groups of Black defendants, even while it is substantively different for White defendants.

In addition, outcome-based tests can yield incorrect conclusions to the extent that the outcome used is itself impacted by racial bias. For example, in a typical comparison between Black and White defendants, one might be worried that the conviction rate of Black defendants is “too high" due to racial bias by prosecutors, judges, or trial juries. This would potentially disguise any racial bias by grand juries at that stage in the process. Put differently, if Black defendants with flimsy evidence are both true-billed, and convicted, due to racial bias, then the outcome-based test will fail to detect it in the traditional comparison of Black and White defendants.

We overcome this by exploiting the fact that while grand juries can only infer race based on defendant name, subsequent conviction decisions are made in a completely unblinded setting. Thus, under the assumption that prosecutors exhibit equal bias against defendants *they know to be Black*, regardless of the racial distinctiveness of their name, any artificial inflating of conviction rates would be the same for both groups of Black defendants in our analysis. We view this assumption as reasonable. In particular, we are unaware of any models of racial discrimination in which agents discriminate on the basis of name *even when* they directly observe defendant race, full criminal history, and the full details of the case. Moreover, we would expect this type of bias to be especially unlikely in this context, given that the two populations of Black defendants have nearly identical observable characteristics and distributions of risk, as shown in Figure 2 and Columns (4) and (5) of Table 1.

To implement the traditional outcome-based test, we estimate the following for Black defendants whose predicted convictability risk puts them in the left tail of the distribution:

$$Conviction_i = \alpha + \beta IdentifiablyBlack_i + \gamma X_i + \rho_t + u_i \tag{2}$$

The coefficient of interest is beta, which captures the difference in conviction rates between Black defendants perceived as Black, and those perceived as White. In particular, to the extent beta is negative, it suggests that marginal defendants perceived as Black are being held to a lower "true bill" standard than marginal defendants perceived as White. Under this test, this would be interpreted as evidence of racial bias against Black defendants.

3.3 Disparate Impact

In our third empirical approach, we first estimate disparate impact against identifiably-Black defendants relative to identifiably-White defendants. To do so, we follow the methodology proposed by Arnold, Dobbie, and Hull (2022), who test for disparate impact in the context of bail judges. The estimate from this method captures the sum of racial bias and statistical discrimination as well as disparate impact caused by equal treatment on the basis of a factor that is more prevalent for defendants of one race than the other. Intuitively, following Arnold, Dobbie, and Hull (2022), we can estimate the disparate impact by rescaling the true bill rate of each race with the observed felony conviction outcome and mean underlying felony convictability of each race. The hurdle is that we do not normally observe mean underlying felony convictability. Rather, we only observe the felony conviction outcome of defendants who were true billed.

To overcome this issue, we exploit the fact that there is quasi-random assignment of cases to grand juries, and that some grand juries are less lenient than others. This enables us to estimate the mean underlying felony convictability of each race by looking at the outcome of cases handled by a hypothetical grand jury that true bills every case. In practice, we do this by extrapolating the trend observed among the different grand juries we observe to estimate what the average felony conviction rate would be for Black and White defendants if their cases were handled by a grand jury that true billed every case. In this way, we estimate the underlying felony convictability of each racial group.

One advantage of doing this in our setting is that it requires much less extrapolation than it does in other settings. In particular, the mean true bill rate in our data is 97.1 percent. By comparison, in Arnold, Dobbie, and Hull (2022) extrapolated in a setting where the mean rate of interest was 73 percent.

Below, we formalize and explain the methodology in more detail.

Defining disparate impact

In this section, we formally define disparate impact in a similar fashion as in Arnold, Dobbie, and Hull (2022). Let $D_{ij} \in \{0,1\}$ denote the grand jury j's true bill decision on defendant i. $R_i \in \{w,b\}$ is the race of defendant i. $Y_i^* \in \{0,1\}$ signifies underlying felony convictability of defendant i. In other words, Y_i^* indicates whether defendant i would be convicted of a felony if they were true billed by the grand jury. $Y_i \in \{0,1\}$ is the observed felony conviction outcome of defendant i, which is always 0 when the case was not true billed $(D_{ij} = 0)$ and equals to Y_i^* when the case is true billed $(D_{ij} = 1)$.

We define the disparate impact in grand jury j's decisions Δ_j as the difference in their true bill rates of black and white defendants with the same underlying level of felony convictability. Δ_{j0} is the true bill rates disparity between black and white defendants without underlying felony convictability $(Y_i^* = 0)$. And Δ_{j1} is the true bill rates disparity between black and white defendants with underlying felony convictability $(Y_i^* = 1)$.

$$\Delta_{j0} = E[D_{ij}|R_i = b, Y_i^* = 0] - E[D_{ij}|R_i = w, Y_i^* = 0] \tag{3}$$

$$\Delta_{j1} = E[D_{ij}|R_i = b, Y_i^* = 1] - E[D_{ij}|R_i = w, Y_i^* = 1] \tag{4}$$

Grand jury j's average disparate impact Δ_j is then the weight average of Δ_{j0} and Δ_{j1}.

$$\Delta_j = \{1 - Pr(Y_i^* = 1)\}\Delta_{jo} + Pr(Y_i^* = 1)\Delta_{j1} \tag{5}$$

Since $\bar{\mu} = E[Y_i^*] = Pr(Y_i^* = 1)$,

$$\Delta_j = (1 - \bar{\mu})\Delta_{jo} + \bar{\mu}\Delta_{j1} \tag{6}$$

To estimate disparate impact Δ_j, we would need $\delta_{jry} \equiv E[D_{ij}|R_i = r, Y_i^* = y]$ and $\bar{\mu}$. This presents a challenge because underlying felony convictability (Y_i^*) is not observed for those who were not true billed.

Estimating disparate impact

As shown earlier, we need $\delta_{jr0} = E[D_{ij}|R_i = r, Y_i^* = 0]$ and $\delta_{jr1} = E[D_{ij}|R_i = r, Y_i^* = 1]$ to estimate disparate impact Δ_j. Following Arnold, Dobbie, and Hull (2022), we show below that we can estimate these two terms with the observed felony conviction outcome Y_i, true bill decision D_i, and the race-specific underlying felony convictability μ_r.

First, using the law of iterated expectation, we show that

$$\begin{aligned}
E[D_{ij}(1-Y_i)|R_i = r] &= Pr(R_i = r, D_{ij} = 1)E[D_{ij}(1-Y_i)|R_i = r, D_{ij} = 1] \\
&= Pr(R_i = r, D_{ij} = 1)E[D_{ij}(1-Y_1^*)|R_i = r, D_{ij} = 1] \\
&= E[D_{ij}(1-Y_i^*)|R_i = r] \\
&= Pr(R_i = r, Y_i^* = 0|R_i = r)E[D_{ij}(1-0)|R_i = r, Y_i^* = 0] \\
&= (1-\mu_r)E[D_{ij}|R_i = r, Y_i^* = 0]
\end{aligned} \tag{7}$$

Therefore, with quasi-random assignment of cases to grand juries,

$$\begin{aligned}
\delta_{jr0} &= E[D_{ij}|R_i = r, Y_i^* = 0] \\
&= \frac{E[D_{ij}(1-Y_i)|R_i = r]}{1-\mu_r} = \frac{E[D_i(1-Y_i)|R_i = r, Z_{ij} = 1]}{1-\mu_r}
\end{aligned} \tag{8}$$

Likewise,

$$
\begin{aligned}
E[D_{ij}(Y_i)|R_i = r] &= Pr(R_i = r, D_{ij} = 1)E[D_{ij}(Y_i)|R_i = r, D_{ij} = 1] \\
&= Pr(R_i = r, D_{ij} = 1)E[D_{ij}(Y_1^*)|R_i = r, D_{ij} = 1] \\
&= E[D_{ij}(Y_i^*)|R_i = r] \\
&= Pr(R_i = r, Y_i^* = 1|R_i = r)E[D_{ij}|R_i = r, Y_i^* = 1] \\
&= \mu_r E[D_{ij}|R_i = r, Y_i^* = 1]
\end{aligned}
\tag{9}
$$

Therefore,

$$
\begin{aligned}
\delta_{jr1} &= E[D_{ij}|R_i = r, Y_i^* = 1] \\
&= \frac{E[D_{ij}Y_i|R_i = r]}{\mu_r} = \frac{E[D_iY_i|R_i = r, Z_{ij} = 1]}{\mu_r}
\end{aligned}
\tag{10}
$$

where Z_{ij} is a binary variable indicating that defendant i was assigned to grand jury j.

Substitute δ_{jr0} from Equation 8 and δ_{jr1} from Equation 10 into Equation 6, we get

$$
\begin{aligned}
\Delta_j &= \bar{\mu}\Delta_{j1} + (1-\bar{\mu})\Delta_{j0} \\
&= \bar{\mu}\{\delta_{jb1} - \delta_{jw1}\} + (1-\bar{\mu})\{\delta_{jb0} - \delta_{jw0}\} \\
&= E\left[\left(\frac{\bar{\mu}Y_i}{\mu_b} + \frac{(1-\bar{\mu})(1-Y_i)}{1-\mu_b}\right)D_i|R_i = b, Z_{ij} = 1\right] \\
&- E\left[\left(\frac{\bar{\mu}Y_i}{\mu_w} + \frac{(1-\bar{\mu})(1-Y_i)}{1-\mu_w}\right)D_i|R_i = w, Z_{ij} = 1\right] \\
&= E[\Omega_i D_i|R_i = b, Z_{ij} = 1] - E[\Omega_i D_i|R_i = w, Z_{ij} = 1]
\end{aligned}
\tag{11}
$$

where

$$
\Omega_i = \left(\frac{\bar{\mu}Y_i}{\mu_r} + \frac{(1-\bar{\mu})(1-Y_i)}{1-\mu_r}\right)
\tag{12}
$$

Since $\bar{\mu} = Pr(R_i = w)\mu_w + Pr(R_i = b)\mu_b$ and we observe felony conviction outcome Y_i and true bill decision D_i, we only need to know μ_r to estimate disparate impact Δ_j.

Extrapolations of underlying felony convictability

In most settings, μ_r, which is the average underlying felony conviction rate of race r cannot be estimated because we do not observe the underlying felony convictability (Y_i^*) of defendants who were not true billed. However, with the random assignment of cases to grand juries, the defendants assigned to each grand jury have the same underlying felony convictability on average. Specifically, the average underlying felony conviction rate of defendants of each race assigned to a grand jury (j) would be the same as the average underlying felony conviction rate of defendants of that race μ_r.

$$\mu_r = E[Y_i^*|R_i = r] = E[Y_i^*|R_i = r, Z_{ij} = 1] = E[Y_i^*|R_i = r, Z_{ij^\star} = 1] \tag{13}$$

In the case of the maximally tough grand jury $j^\star$ that true billed every case, we would be able to observe the underlying felony convictability Y_i^* of all defendants assigned to that jury. This gives us an estimate of the average underlying felony convictability of *all* defendants within a given racial group. This is because true bill decision D_{ij} would be equal to 1 and $Y_i = Y_i^*$ for all defendants of this maximally tough grand jury $j^\star$.

$$E[Y_i^*|Z_{ij^\star} = 1, R_i = r] = E[Y_i|D_{ij} = 1, Z_{ij^\star} = 1, R_i = r] \tag{14}$$

Combining Equations 13 and 14, we show that in a setting featuring the random assignment of cases to grand juries, we can estimate the race-specific underlying felony convictability μ_r from the conviction outcomes of the defendants who were true billed by the maximally tough grand jury that true billed all cases $E[Y_i|D_{ij} = 1, Z_{ij^\star} = 1, R_i = r]$.

$$\mu_r = E[Y_i|D_{ij} = 1, Z_{ij^\star} = 1, R_i = r] \tag{15}$$

Similar to Arnold, Dobbie, and Hull (2022), there are two issues that we need to address in order to estimate the race-specific underlying felony convictability μ_r. First, we do not have a maximally tough grand jury in each time period of our data that true billed every case. Moreover, we want to minimize any sampling error that would arise given the finite sample of cases heard by a maximally tough grand jury. We, therefore, obtain the race-specific average underlying felony convictability μ_r by statistical extrapolation. In practice, this requires little extrapolation because the average true bill rate in our sample is 97.1 percent. Second, in our setting, grand juries were only quasi-randomly assigned conditional on year-month-week of the hearing. We have to take out these time effects from our variables. We explain the procedure used to take out the time effects as well as the extrapolation process below.

Step 1: regress true bill decision D_i and felony conviction outcome Y_i on year-month-week fixed effects, grand jury fixed effects, and the interactions of black defendant dummy and grand jury fixed effects.

$$D_i = \Sigma_j \alpha_j Z_{ij} + \Sigma_j \beta_j Blk_i Z_{ij} + \gamma_t + u_i \tag{16}$$

$$Y_i = \Sigma_j \rho_j Z_{ij} + \Sigma_j \omega_j Blk_i Z_{ij} + \gamma_t + u_i \tag{17}$$

Step 2: construct the time-effects adjusted $E[D_i|Z_{ij} = 1, R_i = r]$ and $E[Y_i|D_i = 1, Z_{ij} = 1, R_i = r]$ from the estimated coefficients as follow:

$$\begin{aligned} E[D_i|Z_{ij} = 1, R_i = w] &= \alpha_j \\ E[D_i|Z_{ij} = 1, R_i = b] &= \alpha_j + \beta_j \\ E[Y_i|D_i = 1, Z_{ij} = 1, R_i = w] &= \rho_j \\ E[Y_i|D_i = 1, Z_{ij} = 1, R_i = b] &= \rho_j + \omega_j \end{aligned} \tag{18}$$

Step 3: Regress and plot $E[Y_i|D_i = 1, Z_{ij} = 1, R_i = r]$ on $E[D_i|Z_{ij} = 1, R_i = r]$ and extrapolate the average underlying felony convictability μ_r by predicting the value of $E[Y_i|D_i = 1, Z_{ij} = 1, R_i = r]$ at $E[D_i|Z_{ij} = 1, R_i = r] = 1$.

Once we extrapolate μ_r, we can estimate the average underlying felony convictability of the whole defendant population $\bar{\mu}$, and then use Equation 11 and Equation 12 to obtain an unbiased estimate the disparate impact Δ_j. Standard errors are two-way clustered at the defendant and grand jury levels.

3.4 Mechanism behind disparate impact

In addition, as alluded to above, one advantage of the context we study is that in contrast to both trial juries and judges in other contexts, grand juries do not directly observe race. This provides an opportunity to us to employ a strategy similar to the “veil of darkness" literature following the seminal paper by Grogger and Ridgeway (2006), or even other studies that have compared across blinded and unblinded settings. In this way, our study deviates from the approach used by Arnold et al. (2022), who use a structural model to tease out the extent to which disparate impact is due to racial bias, versus other factors.

In particular, we use the method explained in the earlier subsection to estimate the disparate impact against Black defendants who have Black names (i.e., those who are likely perceived to be Black). The comparison group is White defendants with identifiably-White names. This is our main specification, as it captures the sum of taste-based bias, statistical discrimination, and non-race-based disparate impact against defendants perceived by jurors to be Black.

Next, we estimate the disparate impact against Black defendants with identifiably White names– that is, Black defendants likely to be perceived by the jurors as White. Again, the comparison group is White defendants with identifiably-White names. This estimate of disparate impact captures only the effect of non-race-based disparate impact against defendants perceived by jurors to be Black. In contrast, it does not capture either taste-based or statistical discrimination based on race, since grand jurors would not be able to identify that the defendants are Black.

Under the assumption that the non-race-based disparate impact against identifiably-Black defendants is the same as that against Black defendants perceived to be White, the difference in the two measures of disparate impact captures the combined effect of taste-based and statistical discrimination based on race. We view this assumption as reasonable because as shown in Table 1, the case characteristics of Black defendants with identifiably-White names are similar to those with identifiably-White names (i.e., Column (3) versus Column (4)), even while both are substantively different from White defendants. It is these substantive differences across races that provide scope for non-race-based disparate impact. For example, grand jurors could use a lower true-bill threshold for certain crimes that are disproportionately committed by Black defendants.

4 Results

4.1 Results from Blinded/Unblinded Comparisons of True Bill Rates

We begin by showing results from comparing the true bill rates of Black defendants with Black versus White names. Results are shown in Table 2. Column (1) shows results for the full sample of 215,279 defendants, controlling only for year-month-week fixed effects. Results indicate that Black defendants who are perceived as Black are 0.15 percentage points *less* likely to be true billed than Black defendants perceived as White. The estimate is significant at the 10 percent level.

Column (2) shows results from the same specification, using only the sample of 198,824 defendants for whom all control variables are observed. The estimated true bill disparity is similar at -0.16 percentage points, which is also significant at the 10 percent level.

Column (3) shows results when we control for offense code fixed effects as well as defendant age, gender, and criminal history. The estimate is reduced somewhat to -0.09 percentage points, which indicates that there is still a modest, though statistically insignificant, reduced likelihood of being true billed for defendants perceived as Black. Importantly, based on the specification in Column (3), we are able to rule out increases in the true bill rate of more than 0.07 percentage points.

Overall, results in Table 2 provide no evidence that grand juries discriminate against defendants perceived to be Black. Rather, if anything these defendants are true billed at very slightly lower rates than observably-similar Black defendants likely perceived as White.

4.2 Results from the Traditional Outcome-Based Test

Results from the traditional outcome-based test are shown in Figure 3 and Table 3. Figure 3 shows a binscatter of the guilty rate for true billed defendants. It does so separately for White defendants with White names (dotted line), Black defendants with Black names (solid line), and Black defendants with White names (dashed line). To minimize the inframarginality problem, we do so for defendants who are most likely to be on the margin of being true billed, based on having predicted conviction likelihoods that would place them in the bottom quartile.[16]

Results in Figure 3 show little difference in the subsequent conviction rates of these three groups of defendants. To better quantify any differences that are present, we show regression results in Table 3. Estimates in Columns (1) and (2) show results without and with controls for Black defendants in the bottom 25 percent of the conviction risk distribution. The estimate in Column 2 indicates that Black defendants who are perceived as Black have conviction rates that are 0.12 percentage points lower than Black defendants perceived as White. This is small, and is statistically indistinguishable from zero. The lower bound of the 95 percent confidence interval indicates we would be able to detect a difference of more than -1.1 percentage points. Estimates in columns (3) through (6) are also economically small and statistically insignificant.

Overall, Figure 3 and Table 3 provide additional evidence that grand juries do not discriminate on the basis of race. In particular, these results show that grand juries are not more likely to push forward weak cases featuring defendants perceived as Black, compared to similar cases featuring defendants perceived as White.

4.3 Grand jury leniency is uncorrelated with case and defendant characteristics

Prior to showing estimates of disparate impact between Black and White defendants, we first provide empirical evidence that the leave-one-out measure of grand jury leniency is uncorrelated

[16]For the purpose of constructing this graph, we exclude those with predicted conviction probabilities that would rank below the first percentile in the distribution for Black defendants.

with defendant and case characteristics. This is important because random assignment is required in order to obtain an unbiased estimate of the likelihood of felony convictability for the full populations of Black and White defendants. We do this by using the variation in grand jury leniency to perform a (slight) extrapolation to estimate the conviction rate for defendants if they were put before a hypothetical grand jury that true bills every case.

Results are shown in Table 4, in which the leave-out-out measure of grand jury leniency was regressed on all defendant and case characteristics. Column (1) shows results for the full set of White and Black defendants. Results for the sample of White defendants with racially-identifiable White names, Black defendants with racially-identifiable Black names, and Black defendants with racially-identifiable White names are shown in Columns (2), (3), and (4), respectively. At the bottom of each column, we report the results from an F-test of joint significance for all of the coefficients on defendant and case characteristics. Each regression also includes year-by-month-by-week fixed effects. Intuitively, this test answers the following thought experiment: Conditional on being heard during the same week, are cases that are heard by the less lenient grand jury observably similar to those that are heard by the more lenient grand jury?

Results are consistent with random assignment. Of the 33 estimates shown, three are significant at the five percent level. While this is somewhat more than one would expect due to chance, it is telling that the magnitude of even the statistically significant estimates is very small. In particular, the absolute magnitude of all three statistically significant estimates is less than 0.001, which indicates a tiny correlation between factors such as male and felony type and grand jury leniency. Moreover, despite the large number of observations, one cannot reject the null hypothesis that the coefficients are all equal to zero for any of the four samples shown.

4.4 Disparate impact on Black defendants with Black names

Next, we turn to our main estimate of disparate impact between Black and White defendants. As noted earlier, however, because actual race is not directly observed by the grand jury, we focus first on comparing White defendants who are likely identified as White, to Black defendants who are likely identified as Black.

Results are shown in Table 5. Panel A shows the estimated mean risk by race for both White and Black defendants. Intuitively, these estimates measure the fraction of White and Black defendants who would be convicted of a felony if every defendant were to be true-billed. Results indicate that across the entire population of identifiably-White and identifiably-Black defendants, Black defendants are somewhat more "convictable" than White defendants (65 percent versus 63.5 percent).

Panel B shows the estimate of disparate impact between White and Black defendants. We show results using three different methods: linear extrapolation, local linear extrapolation with an Epanechnikov kernel with a rule-of-thumb bandwidth, and local linear with a Gaussian kernel with rule-of-thumb bandwidth.[17] Estimates are 0.0076, 0.0084, and 0.0085, respectively, all of which are statistically significant at the one percent level using bootstrapped standard errors. This indicates that grand juries are approximately 0.8 percentage points, or 0.8 percent relative to the mean of 0.971, more likely to true bill a case with a Black defendant than with a White defendant whose underlying felony convictability is similar.

4.5 Disparate impact on Black defendants with White names

While the results in the previous section provide strong evidence that Black defendants are subjected to a small, but statistically significant, disparate impact, it is less clear *why* this is the case. Potential explanations include behavior that is specifically targeted on the basis of race, such as taste-based or statistical discrimination. However, while both of these underlying sources of the disparate impact are illegal, other sources of disparate impact against Black defendants may not be. As a result, while understanding the extent to which there is disparate impact is useful, some important questions are left unanswered.

In order to examine the extent to which the disparate impact is caused by taste-based or statistical

[17]We do not show results from a quadratic extrapolation given how little we are extrapolating (i.e., mean true bill rate = 0.971).

discrimination, rather than differential treatment on the basis of something correlated with race, we borrow a methodology employed by both the “veil of darkness" literature as well as tests of blinded and unblinded behavior. The logic of the test is straightforward: If the disparate impact documented in Table 5 is due entirely to taste-based or statistical discrimination based on race, then we should see no effect *when race cannot be inferred by the grand jurors*.

To implement this test, we show results when we compare White defendants with White names to Black defendants who also have White names. As a result, jurors are unlikely to infer, consciously or otherwise, that these two groups of jurors differ with respect to race. As a result, any nonzero disparate impact estimate will be due to non-race-based disparate impact.

Results are shown in Table 6, which takes the same form as Table 5. Strikingly, results indicate that estimates of disparate impact are similar to those shown in Table 5 and, if anything, are slightly larger. Estimates across the three extrapolation methods are 0.98, 0.96, and 0.94 percentage points. In no case is the estimate in Table 6 smaller than the corresponding estimate in Table 5. In contrast, if anything disparate impact estimates against Black defendants who are likely inferred as White are slightly larger (e.g., 0.0098 versus 0.0076 in Column (1)), which is the opposite of what we would expect if some or all of the disparate impact in Table 5 were due to taste-based or statistical discrimination based on race.

In short, results in Table 5 indicate there is a similar disparate impact against Black defendants *even when the race is unobserved by the grand jurors*. This indicates that whatever the cause of the disparate impact against identifiably-Black defendants documented in Table 5, it is unlikely to be taste-based racial bias or statistical discrimination.

4.6 Rescaling estimates and robustness to an alternative method of measuring juror perception of race

One potential concern with the estimates from Tables 5 and 6 is that jurors are unlikely to perceive that every Black defendant with a Black name is Black, or that every Black defendant with a White name is White. Indeed, Table 1 shows that the predicted likelihood of being Black is

70.2 percent for Black defendants with Black names, and 26.1 percent for Black defendants with White names. In short, while our proxy for the perceived “Blackness" of a name clearly captures differences, it is not perfect. In this way, estimates in Tables 5 and 6 capture reduced-form effects.

Panel A of Table 7 shows the estimates when they are rescaled to account for the imperfect nature of our proxy for the Blackness (or Whiteness of names). The first row shows the difference between the reduced-form estimates shown in Tables 5 and 6. Under the assumption that the non-race-based “disparate impact" component of the estimates in Table 5 is the same as in Table 6, the difference captures the sum of statistical and taste-based discrimination against perceived-Black defendants. Positive estimates indicate the presence of race-based bias against Black defendants, while negative estimates indicate race-based bias against Whites, relative to Black defendants.

The second row shows this same difference, except rescaled to account for the imperfect proxies of Blackness we use.[18] None of the estimates is positive, and thus none suggests the presence of taste-based racial bias or statistical discrimination against identifiably-Black defendants. Rather, estimates in Columns (2) and (3) are close to zero and statistically insignificant. The estimate in Column (1), from linear extrapolation, is negative and significant, and is thus the opposite of what one would expect in the presence of taste-based or statistical discrimination based on race against Black defendants.

Next, we show differences in disparate impacts using an alternative approach for predicting the “Whiteness" and “Blackness" of names. In particular, we predict perceived race using last name and Census Block Group using the “wru" package in R. The intuition for doing so is that while jurors do not directly observe an individual’s neighborhood, to the extent the crime location is demographically similar to one’s neighborhood, or to the extent that witness names or characteristics are correlated with the demographics of one’s own neighborhood, jurors may be able to infer race based on that. Moreover, this method enables us to test whether using a substantively different approach provides similar answers. As with our main approach, we use a

[18]We rescale by dividing the difference in the first row by (0.702 - 0.261), or 0.441. Equivalently, the difference in predicted race between White defendants with White predictions and Black defendants with Black predictions is (0.702-0.151), while the difference between White defendants and Black defendants with White predictions is (0.261-0.151). The difference between these two is 0.441.

50 percent threshold for determining what jurors would infer with respect to the defendant's race.

The full set of results that mirror the main results reported above are shown in Appendix A. Results are largely similar to those presented in the main analysis. Estimates of disparate impact are slightly larger at one percent, but are again the same both for Black defendants with identifiably-Black names, and for Black defendants with White names.

Panel B of Table 7 shows the estimated difference in disparate impacts using that approach, as well as the rescaled difference in disparate impacts. As in Panel A, the differences in disparate impact estimates are small. The rescaled differences are 0.0013, -0.0011, and -0.0011, none of which are statistically significant. Moreover, the magnitude is very small. Taken literally, and under the assumption that non-race-based disparate impact against Black defendants with identifiable names is the same as against Black defendants with White names, the estimates would imply that the sum of statistical and taste-based bias is 0.13 percentage points against Black defendants, or 0.11 percentage points against White defendants.

In summary, results in Table 7 indicate that grand jurors do not engage in statistical or taste-based discrimination against Black defendants. Rather, the small but statistically significant 0.8 percent disparate impact shown in Table 5 appears to be due to jurors applying a slightly lower threshold to cases that tend to be slightly more common among Black defendants, compared to White ones.

5 Robustness

As described above, all three of the empirical analyses used first name and last name, along with a 50 percent threshold, to estimate racial bias. In this section, we show the robustness of our findings to thresholds ranging from 50 to 80 percent, as well as to making this classification based on the median “Blackness" of names, and thus including all Black defendants in the sample. In addition, we show robustness to predicting perceived race using last name and residential address, rather than first and last name, in order to address concerns that jurors may sometimes be able to infer race based on the neighborhood of the alleged offense. Finally, we also show robustness to the inclusion or exclusion of various controls, or to the sample for which there are no missing values with respect to controls.

Figure 4 shows the specification curve for the first approach, where true bill rates are compared across defendants perceived as Black versus White. As shown there, it is clear that there is no difference in the likelihood of being true billed. Rather, the specification curve is both flat and centered at zero. Few, if any, of the estimates are statistically different from zero. This supports our conclusion that defendants perceived as Black are true-billed at the same rate as observably-similar defendants perceived as White.

Figure 5 shows the specification for the traditional outcome-based test. It shows robustness to the same factors as Figure 4, plus robustness to our method of defining Hispanic defendants for the purpose of predicting risk probabilities for Black defendants.[19] Again, while there is some divergence from zero in the tails, the distribution of estimates is centered at zero, and even most of those estimates in the tails are not statistically different from zero. This supports our conclusion that the traditional outcome-based test also indicates that grand juries are not biased against defendants they perceive as Black.

Finally, Figure 6 shows the specification curve for the method based on computing disparate impact, and then comparing disparate impacts in order to estimate the component due to racial bias.[20] Again, the curve is both relatively flat, and is centered at zero. This suggests that our conclusion of no racial bias is robust to alternative decisions regarding prediction method, classification threshold, extrapolation functional form, and the identification of non-Hispanic Whites.

Overall, Figures 4 - 6 indicate that the results from the three empirical tests are robust to a

[19] In our main analysis, we predict conviction risk using all White defendants, and then use that equation to estimate conviction risks for Black defendants. We note that this “White" group includes some Hispanics, as ethnicity is not observed in the administrative data. Our specification curve shows robustness to alternative methods of identifying and excluding Hispanics from the sample used to predict conviction probabilities.

[20] In this case, we test the robustness to alternative ways of identifying Hispanic defendants because this method entails first comparing Whites perceived as White non-Hispanic to Black defendants perceived as Black, and then to compare Whites perceived as White non-Hispanic to Black defendants perceived as Whites. Because we do not observe Hispanic ethnicity in the data, we use alternative ways of identifying Hispanics in the data.

wide range of alternative assumptions about how grand jurors perceive defendant race, and how to estimate racial bias.

6 Conclusion

There has been much interest in the extent to which implicit or explicit racial discrimination is responsible for racial disparities in the criminal justice system. In this paper, we test for racial bias in the context of grand juries. This setting provides three important advantages. The first is that we can perform outcome-based tests, because we observe whether the cases that are pushed forward by grand juries end in felony conviction. This is not feasible in many contexts, including trial juries. Moreover, the subsequent conviction outcomes, combined with the quasi-random assignment of cases across grand juries that differ in leniency, enable us to recover an unbiased estimate of disparate impact by purging omitted variable bias from the estimated Black-White disparity.

The second advantage is that because grand juries do not directly observe the race of the defendant, we can compare outcomes from cases in which jurors likely correctly perceived race (i.e., Black defendants with Black names), to those in which jurors likely (mis)perceived Black defendants as White (i.e., Black defendants with White names). These blinded and unblinded comparisons, which occur in the natural setting of grand jury decision-making, enable us to overcome several issues associated with cross-sectional comparisons of true-bill rates, the traditional outcome-based test, and the disparate impact analysis. Perhaps most notably, this feature of the data, along with the fact that subsequent decisions are made in a fully unblinded setting in which race is directly observed, enable us to overcome concerns that subsequent racial bias by prosecutors, judges, or trial juries would bias the test. In addition, the blinded and unblinded comparisons allow us to perform a simple, intuitive test of whether any disparate impact against identifiably-Black defendants is due to taste-based racial bias or statistical discrimination—both of which are illegal—or if it is due to similar treatment on the basis of something correlated with race, the legality of which is less clear.

Finally, the setting of grand juries enables us to test whether a representative set of U.S. citizens engages in racial bias using a data set of cases that is vastly larger than any previous jury data set. This has important practical implications. For example, the minimum detectable effect (MDE) for our disparate-impact-based approach is 0.17 percentage points. By comparison, the MDE of the next-most-powered study of jury bias is over 130 times larger, and would be even larger than that if we were to scale it to account for the fact the authors estimate the effect of changing the race of only one of six jurors. Thus, the context of grand juries gives us the statistical power to provide a precise and informative—and thus, we believe, publishable—answer, even if that answer is statistically indistinguishable from zero.

Results from three different approaches all indicate the absence of racial bias against defendants perceived to be Black. This includes a comparison of true bill rates of Black defendants with White vs. Black names, an outcome-based test that compares the conviction rates of marginal Black defendants with White vs. Black names, and the disparate impact analysis that uses the blinded/unblinded comparisons to isolate the racial bias component of disparate impact. Moreover, the large sample size enables all three approaches to rule out even small amounts of racial bias. As a result, we conclude that at least in this setting, there is little evidence that American jurors engage in racial bias against Black defendants.

References

Shamena Anwar, Patrick Bayer, and Randi Hjalmarsson. The impact of jury race in criminal trials. *Quarterly Journal of Economics*, 127(2):1017–1055, 2012.

David Arnold, Will Dobbie, and Crystal S Yang. Racial bias in bail decisions. *Quarterly Journal of Economics*, 133(4):1885–1932, 05 2018. ISSN 0033-5533. doi: 10.1093/qje/qjy012. URL `https://doi.org/10.1093/qje/qjy012`.

David Arnold, Will Dobbie, and Peter Hull. Measuring racial discrimination in bail decisions. *American Economic Review*, 112(9):2992–3038, 2022.

E. Jason Baron, Joseph J. Doyle Jr., Natalia Emanuel, Peter Hull, and Joseph P. Ryan. Racial discrimination in child protection, 2023.

Marianne Bertrand and Sendhil Mullainathan. Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94 (4):991–1013, 2004.

Thomas P. Bonczar. Prevalence of imprisonment in the U.S population, 1974-2001, 2003.

Ben Brewer. Is there racial bias in the enforcement of primary seat belt laws? Evidence from veil of darkness tests. *The BE Journal of Economic Analysis & Policy*, (0), 2023.

Francis X Flanagan. Race, gender, and juries: Evidence from North Carolina. *Journal of Law and Economics*, 61(2):189–214, 2018.

Claudia Goldin and Cecilia Rouse. Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90(4):715–741, 2000.

Jeffrey Grogger and Greg Ridgeway. Testing for racial profiling in traffic stops from behind a veil of darkness. *Journal of the American Statistical Association*, 101(475):878–887, 2006.

Mark Hoekstra and Brittany Street. The effect of own-gender jurors on conviction rates. *The Journal of Law and Economics*, 64(3):513–537, 2021.

William C Horrace and Shawn M Rohlin. How dark is dark? Bright lights, big city, racial profiling. *Review of Economics and Statistics*, 98(2):226–232, 2016.

T.D. Stanley Ioannidis, John P.A. and Hristos Doucouliagos. The power of bias in economics research. *Economic Journal*, 127(605):236–265, 2017.

Jesse Kalinowski, Matthew B Ross, and Stephen L Ross. Endogenous driving behavior in tests of racial profiling. Technical report, National Bureau of Economic Research, 2021.

John Knowles, Nicola Persico, and Petra Todd. Racial Bias in Motor Vehicle Searches: Theory and Evidence. *Journal of Political Economy*, 109(1):203–232, February 2001. doi: 10.1086/318603. URL `https://ideas.repec.org/a/ucp/jpolec/v109y2001i1p203-232.html`.

Emma Pierson, Camelia Simoiu, Jan Overgoor, Sam Corbett-Davies, Daniel Jenson, Amy Shoemaker, Vignesh Ramachandran, Phoebe Barghouty, Cheryl Phillips, Ravi Shroff, and Sharad Goel. A large-scale analysis of racial disparities in police stops across the United States. *Nature Human Behaviour*, 4(7):736–745, 2020.

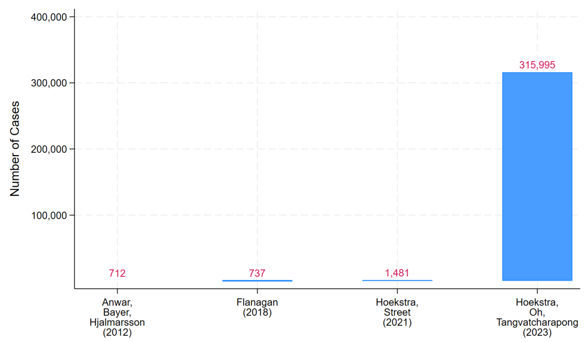
Sarah KS Shannon, Christopher Uggen, Jason Schnittker, Melissa Thompson, Sara Wakefield, and Michael Massoglia. The growth, scope, and spatial distribution of people with felony records in the united states, 1948–2010. *Demography*, 54(5):1795–1818, 2017.

Robert E Worden, Sarah J McLean, and Andrew P Wheeler. Testing for racial profiling with the veil-of-darkness method. *Police Quarterly*, 15(1):92–111, 2012.

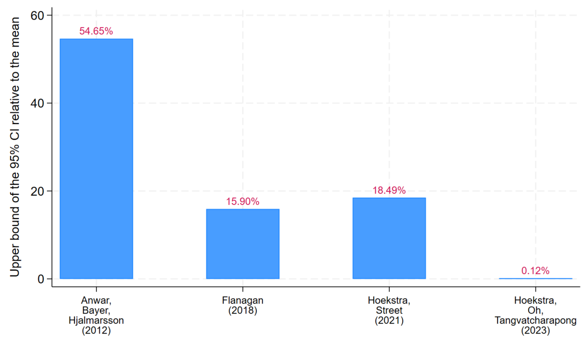
Figures

Figure 1: Sample size (# Cases) and Minimum Detectable Effects in Jury Bias Literature

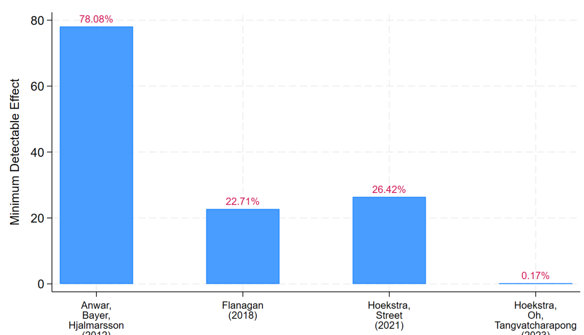
Panel A: Sample Size



Panel B: Upper bound of 95% Confidence Interval



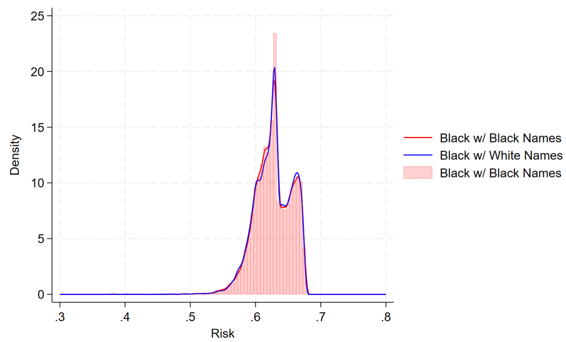
Panel C: Minimum Detectable Effect



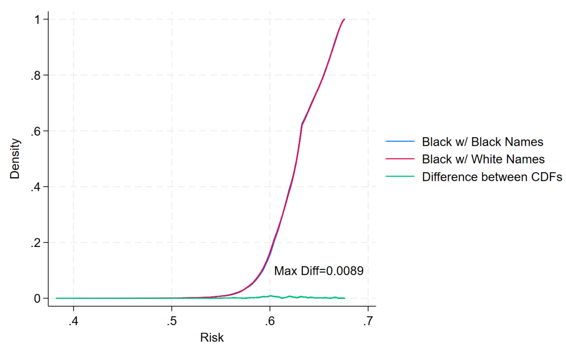
Notes: Panel A displays the total number of criminal cases examined by each study, with the fourth panel representing this study. Panel B illustrates the upper/lower bounds of the 95% confidence interval relative to the mean. Panel C shows the minimum detectable effect (MDE) at 80 percent power for each study, calculated as 2.8 times the standard error from the main estimate of the average effect. For the first three papers, the treatment effect is defined as the impact of having an additional juror (out of six) of a different race or gender. In the current study, depicted in the fourth bar, treatment is defined as the disparate impact between Black and White defendants.

Figure 2: Distribution of Convictability

Panel A: Probability Distribution Function

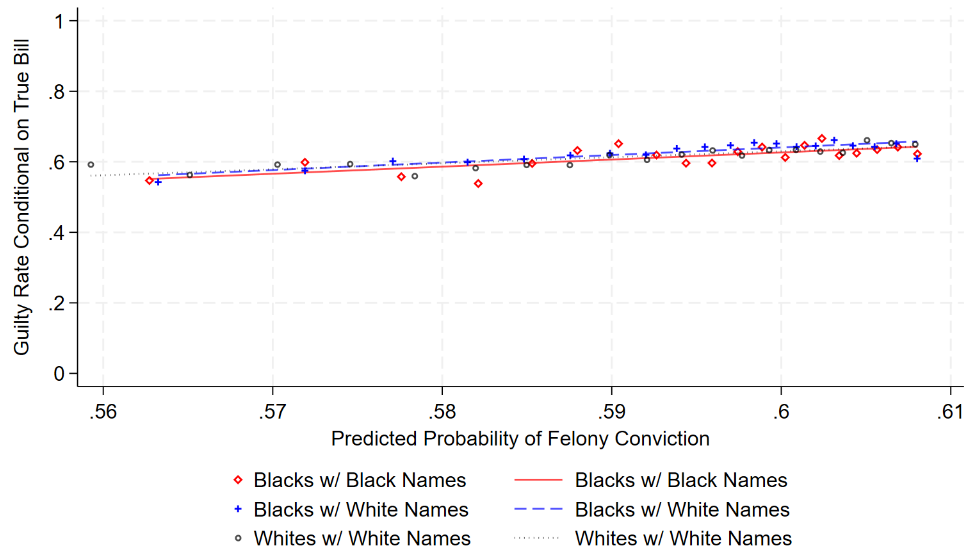


Panel B: Cumulative Distribution Function



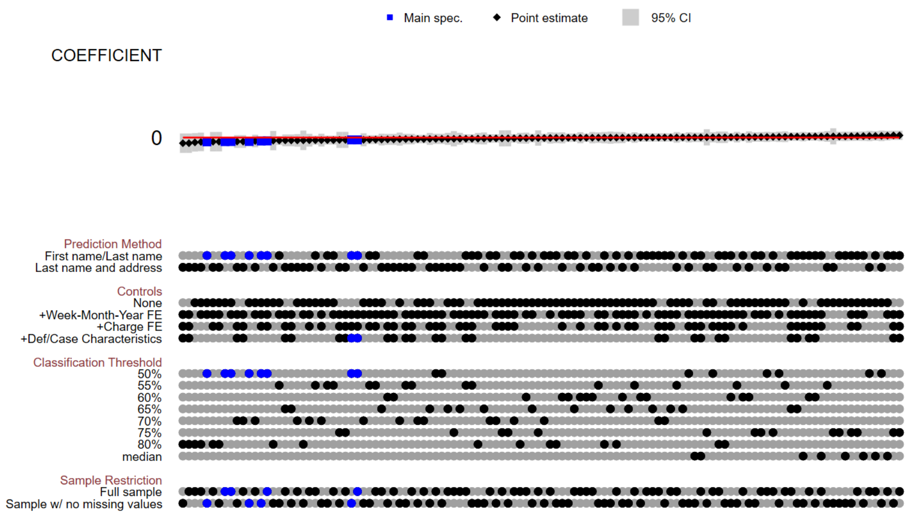
Notes: Panel A shows the probability density function of the predicted probability of felony conviction. Panel B illustrates the cumulative density function of the predicted probability of felony conviction. According to the Epps-Singleton two-sample empirical characteristics function test, the distributions of identifiably black defendants and non-identifiably black defendants show no significant difference, with a p-value of 0.25213. The Epps-Singleton test assesses the hypothesis that the distribution functions underlying two samples are identical. In our context, the ES test is deemed more appropriate than a Kolmogorov-Smirnov two-sample test, as the KS test is unable to address ties.

Figure 3: Outcome-based test



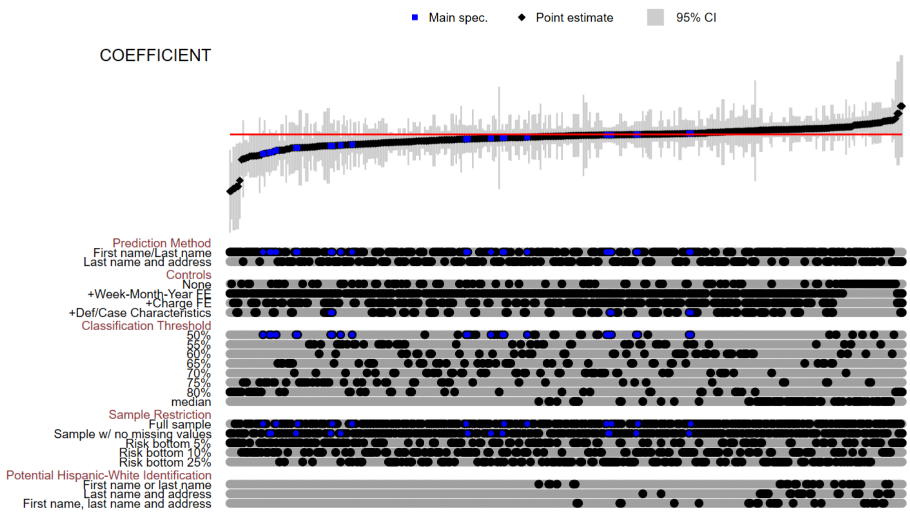
Notes: This figure shows the result from the outcome-based test, where we regress the felony conviction on an identifiably Black indicator among Black defendants, conditional on the true bill.

Figure 4: Specification curve for the OLS estimates of the disparity in true bill rates



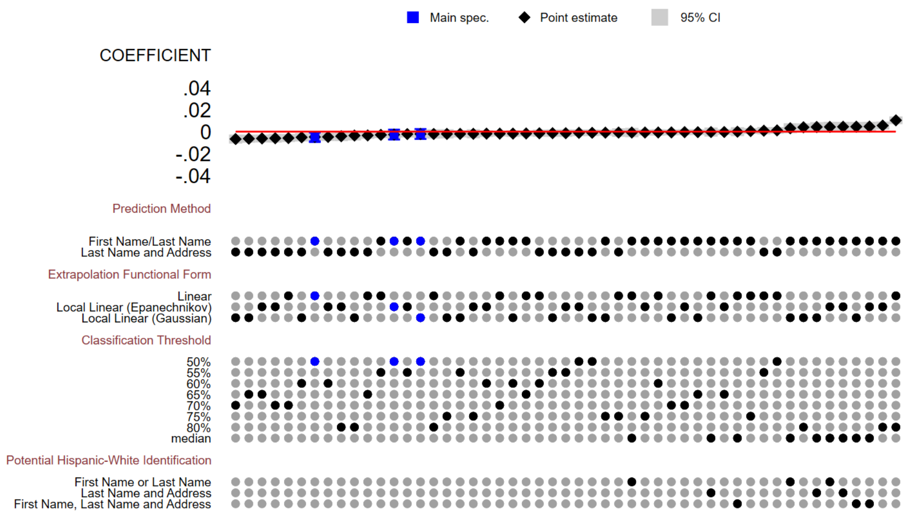
Notes: This specification curve illustrates the OLS estimates of the true bill rates disparity between Blacks perceived as Black vs. White across various settings.

Figure 5: Specification curve for the outcome-based test



Notes: This specification curve illustrates estimates of the outcome-based test across various settings, where we regress felony conviction on an identifiably Black indicator among Black defendants, conditional on the true bill.

Figure 6: Specification curve for the rescaled difference in disparate impact estimates



Notes: This specification curve illustrates the rescaled difference in disparate impact across various settings.

Tables

Table 1: Summary Statistics

	All defendants	Sample w/ racially identifiable names	White w/ White Prediction	Black w/ Black Prediction	Black w/ White Prediction
	(1)	(2)	(3)	(4)	(5)
Panel A: Race Characteristics					
White	0.533	0.387	1.000	0.000	0.000
Black	0.467	0.613	0.000	1.000	1.000
Predicted White	0.433	0.860	1.000	0.000	1.000
Predicted Black	0.074	0.140	0.000	1.000	0.000
Black Prediction Value	0.197	0.280	0.151	0.702	0.261
Panel B: Defendant Characteristics					
Male	0.822	0.823	0.785	0.814	0.858
Age at grand jury hearing	32.398	33.860	34.733	32.547	33.532
Age at filing	32.269	33.736	34.611	32.418	33.409
Prior Offense	0.668	0.701	0.622	0.738	0.754
Panel C: Charge Characteristics					
Felony 1st degree	0.121	0.109	0.081	0.126	0.127
Felony 2nd degree	0.225	0.219	0.205	0.226	0.228
Felony 3rd degree	0.296	0.289	0.319	0.269	0.271
Felony Capital degree	0.003	0.003	0.002	0.004	0.004
Felony State degree	0.354	0.379	0.392	0.375	0.370
Offense degree in linear	3.243	3.324	3.419	3.276	3.262
Number of Charges	1.149	1.151	1.144	1.152	1.156
Panel D: Case outcomes					
Truebill	0.971	0.971	0.966	0.974	0.975
Felony Guilty Conviction	0.621	0.634	0.621	0.631	0.644
Any Guilty Conviction	0.745	0.748	0.747	0.740	0.750
Felony Guilty Conviction when True-billed	0.641	0.653	0.645	0.649	0.661
Any Guilty Conviction when True-billed	0.768	0.770	0.775	0.760	0.770
Observations	695,500	350,560	134,903	49,033	166,624

Notes. This table summarizes the main analysis sample using the inverse of the number of charges as weights. The sample consists of grand jury hearings that were quasi-randomly assigned to grand jury panels between February 1990 and July 2022. Race prediction is based on the likelihood of an individual being White or Black, computed by the R package 'predictrace,' with a threshold of 50%.

Table 2: OLS estimates of the disparity in true bill rates between Black defendants perceived as Black vs. White

	(1)	(2)	(3)
Outcome: True Bill			
Black perceived as Black	-0.00149*	-0.00160*	-0.000933
	(0.000859)	(0.000859)	(0.000844)
Observations	215279	198824	198824
Outcome Mean	0.975	0.976	0.976
Week-Month-Year FE	Y	Y	Y
Case Characteristics	N	N	Y
Offense Code FE	N	N	Y
Non-missing sample	N	Y	Y

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3: Outcome-based test: OLS estimates of the disparity in felony conviction rates conditional on true bill between Black defendants perceived as Black vs. White

	Bottom 25%		Bottom 10%		Bottom 5%	
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome: Felony Conviction						
Black perceived as Black	-0.00537	-0.00121	-0.00726	-0.00162	-0.0183	-0.00864
	(0.00557)	(0.00536)	(0.0100)	(0.00949)	(0.0162)	(0.0159)
Observations	48499	48499	19268	19268	9482	9482
Outcome Mean	0.635	0.635	0.604	0.604	0.560	0.560
Week-Month-Year FE	Y	Y	Y	Y	Y	Y
Case Characteristics	N	Y	N	Y	N	Y
Offense Code FE	N	Y	N	Y	N	Y

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 4: Balance Tests

	(1) All defendants	(2) White defendants w/ White Prediction	(3) Black defendants w/ Black Prediction	(4) Black defendants w/ White Prediction
Black	0.000			
	(0.000)			
Male	0.000	-0.000	0.001**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Age at grand jury hearing	0.000	0.000	0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Felony 1st degree	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Felony 2nd degree	-0.000**	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Felony 3rd degree	-0.000**	-0.000*	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Felony Capital degree	0.000	-0.000	0.000	0.000
	(0.000)	(0.001)	(0.001)	(0.001)
Prior Offense	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Number of Charges	-0.000*	-0.000	0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Joint F-test	0.895	0.848	1.032	1.276
p-value	0.529	0.560	0.410	0.253
Observations	642,652	122,653	45,293	153,711

Notes. This table reports OLS estimates of regressions of grand jury leniency on defendant and case characteristics using an inverse of the number of charges as weights. Each specification controls for year-by-month-by-week-of-year fixed effects. Grand jury leniency is estimated using data from other cases assigned to a given grand jury with weights. The p-values reported at the bottom of each column are from F-tests of the joint significance of the variables. Robust standard error, two-way clustered at the defendant and the grand jury panel level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Mean Risk and Total Disparate Impact Estimates: Black Defendants with Black Predictions vs. White Defendants with White Predictions

	Linear Extrapolation	Local Linear Extrapolation Epanechnikov (ROT)	Local Linear Extrapolation Gaussian (ROT)
Panel A: Mean Risk by Race	(1)	(2)	(3)
White Defendants	0.6240	0.6356	0.6361
	(0.0007)	(0.0013)	(0.0013)
Black Defendants	0.6455	0.6533	0.6541
	(0.0009)	(0.0016)	(0.0016)
Panel B: Total Disparate Impact			
Mean Across Cases	0.0076***	0.0084***	0.0085***
	(0.0006)	(0.0006)	(0.0006)
Juries	646	646	646

Notes. Panel A shows the estimated mean likelihood of being found guilty of a felony for the full population of White and Black defendants. This is computed using variation in leniency across grand juries to extrapolate (slightly) to what the felony guilt rate would be if a grand jury were to true bill every defendant. Panel B shows the estimated total disparate impact, which is the sum of taste-based discrimination, statistical discrimination, and non-race-based disparate impact. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Mean Risk and Total Disparate Impact Estimates: Black Defendants with White Predictions vs. White Defendants with White Predictions

	Linear Extrapolation	Local Linear Extrapolation Epanechnikov (ROT)	Local Linear Extrapolation Gaussian (ROT)
Panel A: Mean Risk by Race	(1)	(2)	(3)
White Defendants	0.6391	0.6357	0.6335
	(0.0003)	(0.0045)	(0.0046)
Black Defendants	0.6584	0.6588	0.6571
	(0.0005)	(0.0027)	(0.0027)
Panel B: Total Disparate Impact			
Mean Across Cases	0.0098***	0.0096***	0.0094***
	(0.0004)	(0.0006)	(0.0006)
Juries	646	646	646

Notes. Panel A shows the estimated mean likelihood of being found guilty of a felony for the full population of White and Black defendants. This is computed using variation in leniency across grand juries to extrapolate (slightly) to what the felony guilt rate would be if a grand jury were to true bill every defendant. Panel B shows the estimated total disparate impact. In this case, because the Black defendants are not identifiable as Black by the grand jury, disparate impact captures only non-race based disparate impact, such as using a lower threshold of guilt for some types of cases that are more common among Black defendants. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Taste-based racial bias and statistical discrimination against identifiably-Black defendants

	Linear Extrapolation	Local Linear Extrapolation Epanechnikov (ROT)	Local Linear Extrapolation Gaussian (ROT)
Panel A: Predicting race using either first name or last name			
Difference in disparate impacts	-0.0022***	-0.0012	-0.0009
	(0.0007)	(0.0008)	(0.0008)
Rescaled Difference in disparate impacts	-0.0050***	-0.0027	-0.0020
	(0.0016)	(0.0019)	(0.0019)
Panel B: Predicting race using either first name or last name - Median			
Difference in disparate impacts	-0.0002	0.0012*	0.0009
	(0.0006)	(0.0007)	(0.0006)
Rescaled difference in disparate impacts	-0.0007	0.0041*	0.0031
	(0.0022)	(0.0024)	(0.0022)
Panel C: Predicting race based on last name and address together			
Difference in disparate impacts	0.0008	-0.0007	-0.0007
	(0.0010)	(0.0010)	(0.0010)
Rescaled difference in disparate impacts	0.0013	-0.0011	-0.0011
	(0.0016)	(0.0016)	(0.0016)

Notes. The first row in each panel is the disparate impact estimate against Black defendants with Black names minus the disparate impact against Black defendants with White names. Under the assumption that the non-race-based disparate impact is similar across both groups of Black defendants, this difference captures the sum of taste-based and statistical discrimination against identifiably-Black defendants. The second row shows the same difference in disparate impacts, except rescaled by the difference in the likelihood of being perceived as Black for Black defendants with Black versus White names. Panel B and C show results when using the same prediction method but a median as a threshold instead of 50%, and results when using last name and address to predict the perceived race of the defendant, respectively. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix A: Results from using race predictions based on surname and address

Table A.1: Summary Statistics

	All defendants	Sample w/ racially identifiable names	White Defendants w/ White Prediction	Black Defendants w/ Black Prediction	Black Defendants w/ White Prediction
	(1)	(2)	(3)	(4)	(5)
Panel A: Race Characteristics					
White	0.533	0.255	1.000	0.000	0.000
Black	0.467	0.745	0.000	1.000	1.000
Predicted White (50%)	0.172	0.315	1.000	0.000	1.000
Predicted Black (50%)	0.404	0.685	0.000	1.000	0.000
Likelihood of being White	0.196	0.281	0.765	0.068	0.659
Likelihood of being Black	0.391	0.610	0.087	0.830	0.195
Panel B: Defendant Characteristics					
Male	0.822	0.798	0.771	0.806	0.820
Age at grand jury hearing	32.398	32.771	34.327	32.244	32.151
Age at filing	32.269	32.640	34.198	32.113	32.012
Prior Offense	0.668	0.711	0.605	0.752	0.691
Panel C: Charge Characteristics					
Felony 1st degree	0.121	0.113	0.075	0.127	0.118
Felony 2nd degree	0.225	0.222	0.202	0.228	0.223
Felony 3rd degree	0.296	0.287	0.327	0.272	0.283
Felony Capital degree	0.003	0.003	0.002	0.004	0.002
Felony State degree	0.354	0.375	0.394	0.369	0.374
Offense degree in linear	3.243	3.305	3.438	3.258	3.291
Number of Charges	1.149	1.154	1.141	1.158	1.157
Panel D: Case outcomes					
Truebill	0.971	0.972	0.964	0.975	0.974
Felony Guilty Conviction	0.621	0.624	0.604	0.632	0.613
Any Guilty Conviction	0.745	0.739	0.737	0.741	0.727
Felony Guilty Conviction when True-billed	0.621	0.624	0.604	0.632	0.613
Any Guilty Conviction when True-billed	0.745	0.739	0.737	0.741	0.727
Observations	695,500	278,688	70,371	191,709	16,608

Notes. This table summarizes the main analysis sample using the inverse of the number of charges as weights. The sample consists of grand jury hearings that were quasi-randomly assigned to grand jury panels between February 1990 and July 2022. Race prediction is based on the likelihood of an individual being White or Black, computed by the R package 'wru,' with a threshold of 50%.

Table A.2: Balance Tests

	(1) All defendants	(2) White defendants w/ White Prediction	(3) Black defendants w/ Black Prediction	(4) Black defendants w/ White Prediction
Black	0.000			
	(0.000)			
Male	0.000	-0.000**	0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.001)
Age at grand jury hearing	0.000	-0.000	-0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Felony 1st degree	-0.000	0.000	-0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.001)
Felony 2nd degree	-0.000**	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.001)
Felony 3rd degree	-0.000**	-0.000	-0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.001)
Felony Capital degree	0.000	-0.001	-0.001	-0.004
	(0.000)	(0.003)	(0.001)	(0.005)
Prior Offense	0.000	0.000	0.000	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)
Number of Charges	-0.000*	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Joint F-test	0.895	1.582	1.405	0.869
p-value	0.529	0.127	0.191	0.542
Observations	642,652	63,852	177,437	15,157

Notes. This table reports OLS estimates of regressions of grand jury leniency on defendant and case characteristics using an inverse of the number of charges as weights. Each specification controls for year-by-month-by-week-of-year fixed effects. Grand jury leniency is estimated using data from other cases assigned to a given grand jury with weights. The p-values reported at the bottom of each column are from F-tests of the joint significance of the variables. Robust standard error, two-way clustered at the defendant and the grand jury panel level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Mean Risk and Total Disparate Impact Estimates: Black Defendants with Black Predictions vs. White Defendants with White Predictions

	Linear Extrapolation	Local Linear Extrapolation Epanechnikov (ROT)	Local Linear Extrapolation Gaussian (ROT)
Panel A: Mean Risk by Race	(1)	(2)	(3)
White Defendants	0.6140	0.6093	0.6099
	(0.0013)	(0.0016)	(0.0016)
Black Defendants	0.6442	0.6422	0.6413
	(0.0009)	(0.0010)	(0.0010)
Panel B: Total Disparate Impact			
Mean Across Cases	0.0104***	0.0098***	0.0099***
	(0.0005)	(0.0005)	(0.0005)
Juries	646	646	646

Notes. Panel A shows the estimated mean likelihood of being found guilty of a felony for the full population of White and Black defendants. This is computed using variation in leniency across grand juries to extrapolate (slightly) to what the felony guilt rate would be if a grand jury were to true bill every defendant. Panel B shows the estimated total disparate impact, which is the sum of taste-based discrimination, statistical discrimination, and non-race-based disparate impact. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

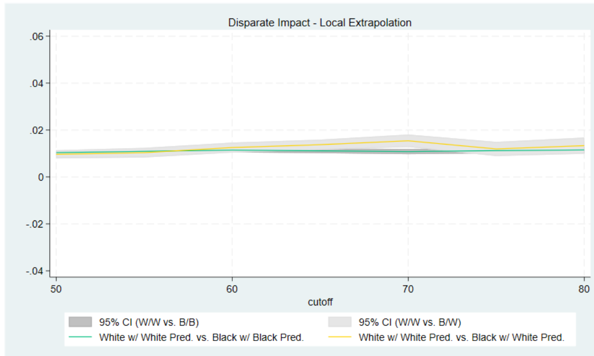
Table A.4: Mean Risk and Total Disparate Impact Estimates: Black Defendants with White Predictions vs. White Defendants with White Predictions

	Linear Extrapolation	Local Linear Extrapolation Epanechnikov (ROT)	Local Linear Extrapolation Gaussian (ROT)
Panel A: Mean Risk by Race	(1)	(2)	(3)
White Defendants	0.6040	0.6116	0.6119
	(0.0013)	(0.0019)	(0.0019)
Black Defendants	0.6292	0.6374	0.6380
	(0.0018)	(0.0024)	(0.0024)
Panel B: Total Disparate Impact			
Mean Across Cases	0.0097***	0.0105***	0.0105***
	(0.0009)	(0.0009)	(0.0009)
Juries	646	646	646

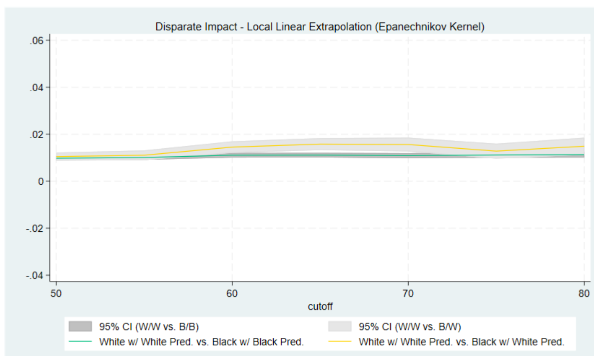
Notes. Panel A shows the estimated mean likelihood of being found guilty of a felony for the full population of White and Black defendants. This is computed using variation in leniency across grand juries to extrapolate (slightly) to what the felony guilt rate would be if a grand jury were to true bill every defendant. Panel B shows the estimated total disparate impact. In this case, because the Black defendants are not identifiable as Black by the grand jury, disparate impact captures only non-race based disparate impact, such as using a lower threshold of guilt for some types of cases that are more common among Black defendants. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A.1: Robustness to Alternative Thresholds

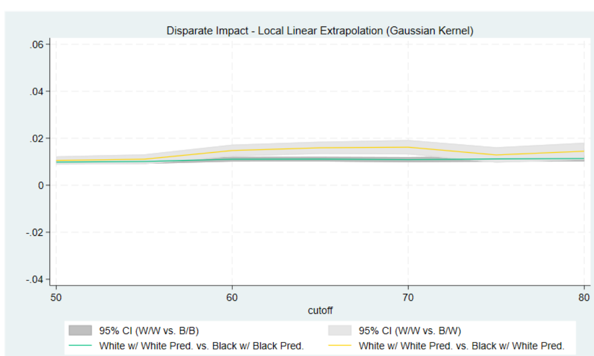
Panel A: Linear Extrapolation



Panel B: Local Linear Extrapolation Epanechnikov



Panel C: Local Linear Extrapolation Gaussian



Notes: This figure provides disparate impacts and 95% confidence intervals using linear extrapolation for thresholds ranging from 50% (shown in main results) to 80%. Estimates in green are disparate impact estimates against Black defendants with identifiably-Black names/locations, while estimates in yellow are disparate impact estimates against Black defendants with White names/locations who are therefore likely perceived as White by grand jurors.

Appendix B: Results from using race predictions based on first name and surname with a median threshold

Table B.1: Summary Statistics

	All defendants	Sample w/ racially identifiable names	White w/ White Prediction	Black w/ Black Prediction	Black w/ White Prediction
	(1)	(2)	(3)	(4)	(5)
Panel A: Race Characteristics					
White	0.533	0.208	1.000	0.000	0.000
Black	0.467	0.792	0.000	1.000	1.000
Black Prediction Value	0.197	0.277	0.086	0.475	0.182
Panel B: Defendant Characteristics					
Male	0.822	0.810	0.796	0.818	0.809
Age at grand jury hearing	32.398	32.768	34.399	32.399	32.280
Age at filing	32.269	32.640	34.276	32.270	32.151
Prior Offense	0.668	0.714	0.618	0.746	0.733
Panel C: Charge Characteristics					
Felony 1st degree	0.121	0.119	0.082	0.128	0.130
Felony 2nd degree	0.225	0.224	0.205	0.228	0.229
Felony 3rd degree	0.296	0.280	0.319	0.269	0.272
Felony Capital degree	0.003	0.004	0.002	0.004	0.004
Felony State degree	0.354	0.373	0.392	0.371	0.366
Offense degree in linear	3.243	3.288	3.417	3.260	3.248
Number of Charges	1.149	1.155	1.145	1.158	1.158
Panel D: Case outcomes					
Truebill	0.971	0.973	0.965	0.975	0.974
Felony Guilty Conviction	0.621	0.629	0.617	0.636	0.628
Any Guilty Conviction	0.745	0.741	0.745	0.742	0.738
Felony Guilty Conviction when True-billed	0.641	0.647	0.641	0.653	0.645
Any Guilty Conviction when True-billed	0.768	0.762	0.773	0.762	0.758
Observations	695,500	412,089	85,056	162,526	164,507

Notes. This table summarizes the main analysis sample using the inverse of the number of charges as weights. The sample consists of grand jury hearings that were quasi-randomly assigned to grand jury panels between February 1990 and July 2022. Race prediction is based on the likelihood of an individual being White or Black, computed by the R package ‘predictrace,’ with a median as a threshold.

Table B.2: Balance Tests

	(1) White defendants w/ White Prediction	(2) Black defendants w/ Black Prediction	(3) Black defendants w/ White Prediction
Male	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Age at grand jury hearing	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Felony 1st degree	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Felony 2nd degree	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Felony 3rd degree	-0.000	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)
Felony Capital degree	0.002	0.000	-0.001
	(0.001)	(0.001)	(0.001)
Prior Offense	0.000***	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Number of Charges	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Joint F-test	1.809	0.568	1.221
p-value	0.072	0.805	0.284
Observations	77,427	150,375	152,127

Notes. This table reports OLS estimates of regressions of grand jury leniency on defendant and case characteristics using an inverse of the number of charges as weights. Each specification controls for year-by-month-by-week-of-year fixed effects. Grand jury leniency is estimated using data from other cases assigned to a given grand jury with weights. The p-values reported at the bottom of each column are from F-tests of the joint significance of the variables. Robust standard error, two-way clustered at the defendant and the grand jury panel level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Mean Risk and Total Disparate Impact Estimates: Black Defendants with Black Predictions vs. White Defendants with White Predictions

	Linear Extrapolation	Local Linear Extrapolation Epanechnikov (ROT)	Local Linear Extrapolation Gaussian (ROT)
Panel A: Mean Risk by Race	(1)	(2)	(3)
White Defendants	0.6259	0.6233	0.6232
	(0.0013)	(0.0018)	(0.0017)
Black Defendants	0.6475	0.6471	0.6467
	(0.0009)	(0.0010)	(0.0010)
Panel B: System-Wide Discrimination			
Mean Across Cases	0.0096	0.0093	0.0093
	(0.0005)	(0.0005)	(0.0005)
Juries	646	646	646

Notes. Panel A shows the estimated mean likelihood of being found guilty of a felony for the full population of White and Black defendants. This is computed using variation in leniency across grand juries to extrapolate (slightly) to what the felony guilt rate would be if a grand jury were to true bill every defendant. Panel B shows the estimated total disparate impact, which is the sum of taste-based discrimination, statistical discrimination, and non-race-based disparate impact. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: Mean Risk and Total Disparate Impact Estimates: Black Defendants with White Predictions vs. White Defendants with White Predictions

	Linear Extrapolation	Local Linear Extrapolation Epanechnikov (ROT)	Local Linear Extrapolation Gaussian (ROT)
Panel A: Mean Risk by Race	(1)	(2)	(3)
White Defendants	0.6361	0.6122	0.6152
	(0.0003)	(0.0027)	(0.0027)
Black Defendants	0.6472	0.6362	0.6366
	(0.0005)	(0.0015)	(0.0013)
Panel B: System-Wide Discrimination			
Mean Across Cases	0.0098	0.0081	0.0084
	(0.0004)	(0.0005)	(0.0004)
Juries	646	646	646

Notes. Panel A shows the estimated mean likelihood of being found guilty of a felony for the full population of White and Black defendants. This is computed using variation in leniency across grand juries to extrapolate (slightly) to what the felony guilt rate would be if a grand jury were to true bill every defendant. Panel B shows the estimated total disparate impact, which is the sum of taste-based discrimination, statistical discrimination, and non-race-based disparate impact. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$