



A Test of Racial Bias in Capital Sentencing

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters

Citation	Alesina, Alberto, and Eliana La Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing." <i>American Economic Review</i> 104 (11) [November]: 3397–3433. doi:10.1257/aer.104.11.3397.
Published Version	doi:10.1257/aer.104.11.3397
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:30752840
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

A Test of Racial Bias in Capital Sentencing[†]

By ALBERTO ALESINA AND ELIANA LA FERRARA*

We collect a new dataset on capital punishment in the United States and we propose a test of racial bias based upon patterns of sentence reversals. We model the courts as minimizing type I and II errors. If trial courts were unbiased, conditional on defendant's race the error rate should be independent of the victim's race. Instead we uncover 3 and 9 percentage points higher reversal rates in direct appeal and habeas corpus cases, respectively, against minority defendants who killed whites. The pattern for white defendants is opposite but not statistically significant. This bias is confined to Southern states. (JEL J15, K41, K42)

One of the arguments against the death penalty in the United States is that it is applied with a racial bias against minorities. Consider for example the following statement, taken from the opening paragraph of a document by one of the most vocal organizations opposing capital punishment:

African Americans are disproportionately represented among people condemned to death in the USA. While they make up 12 percent of the national population, they account for more than 40 percent of the country's current death row inmates, and one in three of those executed since 1977.¹

While factually correct, statements like these can hardly be interpreted as evidence of racial bias, because violent crime rates are higher amongst minorities than whites. Accounting for differences in patterns of crime—and more generally in unobservables which may be correlated with race—is crucial if one wants to test rigorously for racial bias. We propose a test of racial bias in capital sentencing that allows for the possibility that members of different racial groups differ along

*Alesina: Department of Economics, Harvard University, Littauer Center 210, 1805 Cambridge Street, Cambridge, MA 02138, IGER, and NBER (e-mail: aalesina@harvard.edu); La Ferrara: Department of Economics and IGER, Università Bocconi, via Rontgen 1, 20136 Milano, Italy (e-mail: eliana.laferrara@unibocconi.it). The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. We thank three anonymous referees, Shamena Anwar, Abhijit Banerjee, Katherine Barnes, Francesco Corielli, Richard Dieter, John Donohue, Jeffrey Fagan, James MacKinnon, Nicola Persico, Stephen Ross, Andrei Shleifer, and seminar participants at CIAR, Duke, Warwick, Royal Holloway, IIES Stockholm University, Queen's University, University of Connecticut, NBER Political Economy Conference, CEPR Public Policy Conference, and Villa La Pietra Conference on New Directions in Applied Microeconomics for helpful comments. Giulia La Mattina, Lucia Rizzica, Erika Deserranno, and Mariagiovanna Di Feo provided invaluable help with data collection. We also thank Damiano Briguglio, Giovanni Rizzo, Matteo Fiorini, Marta Barazzetta, Matteo Libergolis, and Nicola Fontana for excellent research assistance. This paper was written while Alesina was visiting IGER Bocconi. He thanks this institution for hospitality. La Ferrara acknowledges financial support from the European Research Council grant ERC-2007-StG-208661. The usual disclaimer applies.

[†]Go to <http://dx.doi.org/10.1257/aer.104.11.3397> to visit the article page for additional materials and author disclosure statement(s).

¹Amnesty International (2003, p. 1), *USA Death by Discrimination—The Continuing Role of Race in Capital Cases*.

observable and unobservable dimensions, and we collect a large original dataset to implement this test.

We develop a model where courts minimize the probability of making judicial errors and we derive a simple test for racial bias. We build upon the work of Anwar and Fang (2006) and propose a test based upon the same insight. Even if we do not observe in the data all the elements that trial courts consider when imposing a death verdict, if the judicial process is unbiased, *ex post* we should not observe judicial errors more frequently in cases involving certain combinations of defendant and victim's race. We exploit a feature of the capital sentencing process in the United States, namely that all first degree capital sentences are automatically appealed (so we have no selection bias), and we focus upon errors of first degree courts reversed by higher courts. Our test rests on the assumption that superior courts can only improve upon the accuracy of first sentencing and therefore remove part or all racial bias.

Our model allows for the possibility that racial groups differ in their propensity to commit crimes, in the quality of legal assistance they have access to, and in other unobserved dimensions. This implies that a simple test comparing errors in judgements against minority defendants with errors against white defendants is inconclusive, as differences in error rates may reflect differences in unobservables that are correlated with defendants' race. Under the assumption that for given defendant's race the distribution of these unobservables does not vary with the race of the victim, we can build a test based on pairs of victim/defendant races.² This test relies upon the idea that the ranking of first degree mistakes depending upon these pairs should not violate certain patterns that are consistent with unbiased courts. For example, if courts commit more errors on minority defendants who killed white victims than on those who killed nonwhite victims, they should also commit more errors on *white* defendants who killed white victims than on those who killed nonwhite ones. In other words, for each defendant's race the ranking of error rates across victims' race must be the same. Failure to satisfy this condition implies the presence of racial bias in our model. We also discuss an extension of the model that allows for bias in the work of the police, the prosecutor, or the defense, which implies that distribution of the evidence may depend on both victim's and defendant's races. In this case the results of our rank-order test cannot be interpreted as bias attributable exclusively to the court, but can be viewed more generally as a bias in the criminal justice system: that is, the combination of court, police, and trial procedures.

In order to implement our test we embarked on a challenging data collection project. We started from the data on capital appeals assembled by Liebman, Fagan, and West (2000) for the period 1973–1995. From their data we take information on reversal of first degree death sentences at two stages of appeal: the direct appeal, which is the first stage of appeal that every death sentence undergoes and is ruled by state high courts, and the habeas corpus, which is the final stage of appeal and is ruled by federal courts. We supplement this data with information we collected on a case by case basis. An especially difficult variable to reconstruct was the race of the victim for each case (neither defendant's nor victim's race are available in Liebman,

²The death penalty applies almost exclusively to homicides so there is always at least one well-identified victim.

Fagan, and West's data). As a result, our study is the first to provide even descriptive information on the racial composition of victims in capital cases for the entire United States in the period under study. As we report below, 51 percent of defendants in the first stage of appeal are white and 41 percent are African American. On the other hand, 78 percent of these cases involve at least one white victim, and 17 percent at least one African American victim.

When we implement our test we find results consistent with the presence of racial prejudice: *ceteris paribus*, first degree courts are more severe (i.e., they tend to give more death sentences which are then reversed) against cases involving a minority defendant killing one or more white victims. For habeas corpus cases involving a minority defendant, the error rate was 37.6 percent if the victim was white, and 28.4 percent if it was nonwhite, with a statistically significant difference of 9 percentage points. For cases involving a white defendant the difference indicates higher reversal rates when the victim is nonwhite, but it is not significant. In the direct appeal sample, cases involving a minority defendant had an error rate of 37.7 percent if the victim was white and 34.7 percent if the victim was a minority, with a statistically significant difference of 3 percentage points. In cases involving a white defendant the difference is again in the opposite direction, comparable in magnitude and not significant. This pattern of results is consistent with racial bias according to our rank order test. When we disaggregate the results by region, we find that the effect is driven exclusively by Southern states. The difference in error rates in these states is large: in habeas corpus cases, the error rate is 15.6 percentage points higher for minority defendants with white victims as compared to minority defendants with nonwhite victims (*p*-value 0.01). For the direct appeal sample the corresponding difference is 3.3 percentage points (*p*-value 0.13). For white defendants in the South, the difference is 10 percentage points in the habeas corpus and 4 percentage points in the direct appeal, in both cases in the opposite direction (higher reversals for whites killing minorities) but not statistically significant. It should be clear that the above results cannot be interpreted as a bias against minority defendants *per se*, but only against minority defendants who killed white victims relative to those who killed nonwhites, and only in the South.

The validity of our test relies upon several assumptions. The first is an assumption about the behavior of the higher courts. If these courts are unbiased and make mistakes uncorrelated with the race of defendant and victim our tests are exactly specified. If higher courts are also racially biased in the same direction of the lower courts but less so, our test *underestimates* the amount of racial bias of first degree courts. Our test would overestimate the level of bias if higher courts actively discriminated in favor of minority defendants who killed white victims.³ Note that our test would *not* fail if higher courts simply discriminated in favor of minorities, say because of lower quality of their legal counsel: to invalidate our test the reverse discrimination should be targeting very specifically the minority defendant/white victim pair. We assess the plausibility of this interpretation empirically, exploiting differences in ideology across appeal courts. We build upon the premise that the judges who would be most likely to reverse discriminate in favor of minority defendants who killed

³ See Argys and Mocan (2004) on the issue of reverse discrimination in executions and sentence commutations.

white victims would be those more left-leaning. Using various measures of political orientation of higher courts' judges we do not find any evidence of this effect: both left-wing and right-wing leaning judges exhibit the same pattern of reversal of first degree sentences. Note that if, as our model assumes, the supreme courts are totally unbiased, in the end the legal system produces racially unbiased final sentences, namely the errors of the lower courts are corrected. However, even in this scenario the racial bias of lower courts entails substantial costs. First, the years spent on death row by inmates whose sentence will be reversed: Liebman, Fagan, and West (2000, iii) report that "catching so much error takes time—a national average of 9 years from death sentence to the last inspection and execution." Second, not all cases reach the habeas corpus final procedure, which according to our results is where it appears that a significant fraction of the bias is corrected. Thus, a fraction of the errors may still go undetected.

Another assumption upon which our test rests is that possible unobservables, such as characteristics of the crime or quality of the evidence, are not systematically different across victims' races, for given defendant's race. To assess the plausibility of this assumption, we test whether the distribution of observable characteristics correlated with the nature of the crime (e.g., situation, weapon, etc.) or with its severity (e.g., aggravating and mitigating factors) differs systematically across racial pairs, and we find that for the most part it does not. We also investigate whether the results of our test are robust to conditioning on a number of observables and find that by and large they are. While this does not rule out possible differences in unobservables, the outcome of this battery of tests increases our confidence in the interpretation of our results as consistent with racial bias.

As mentioned above, our model builds upon the literature on racial bias in motor vehicle searches and in particular on recent work by Knowles, Persico, and Todd (2001); Anwar and Fang (2006); and Antonovics and Knight (2009). However, our model differs from those papers in several ways. In models of car searches the issue is which car to stop and then, with certainty, either contraband is found or not. In our model the courts have to evaluate guilt or innocence based on a noisy signal and there is a review of the first decision. Guilt or innocence cannot be decided for sure like in a car search. The objective function of our courts is therefore different from that of a trooper stopping cars, in that it trades off the extent of type I and type II errors.⁴ We share with Gennaioli and Shleifer (2007) an interest in the effect of bias in judges' decisions. These authors however address a different research question, namely how common law and the accumulation of precedents leads towards an equilibrium without judicial bias. We do not pursue this type of dynamic analysis of bias. Two recent papers exploit random variation to test if there are systematic differences across judges in the racial gap in sentencing for felony crimes (Abrams, Bertrand, and Mullainathan 2012) or if there is evidence of bias in felony trial outcomes depending on the racial composition of the jury (Anwar, Bayer, and Hjalmarsson 2012). This type of random variation is not available for death penalty cases, so our test is built on different grounds, exploiting a prediction on the equilibrium behavior of the court. We should also clarify that, although it would

⁴Both our paper and the literature of motor vehicle searches owe a lot to the pathbreaking work by Becker (1957) on rational models of crime.

be extremely interesting to disentangle the behavior of different actors (prosecutor, jurors, judge), our data does not allow us to advance in this direction and will only allow us to detect a *combined* bias of the criminal justice system up to the trial stage.

Our paper is also related to the literature on the death sentence and its usefulness. We do not touch upon this issue. We focus only on the question of whether or not the death penalty is applied with a racial bias.⁵ There are several early contributions in the law literature on the role played by race in capital sentencing and execution. The stylized facts described in this literature include: (i) the disproportionate execution of blacks compared to whites and (ii) the higher likelihood that the death penalty is imposed when the victim is white. Most of these studies rely on small samples and can be criticized on the grounds that important factors affecting the decision of the court may not be observable in the data. This is almost inevitably the case when a direct test of discrimination at the sentencing stage is attempted. Even the most comprehensive data source, in fact, will not possibly include all the information that was available to the court at the moment when the sentence was imposed. One of the most influential early attempts at controlling for observable factors is a study by Gross and Mauro (1984). They constructed an index of aggravating factors and found that, after controlling for them, the race of the victim was still a strong predictor of capital sentencing (the likelihood of a death sentence being higher when the victim was white), but the race of the defendant had no residual explanatory power. In a famous study, Baldus, Pulanski, and Woodworth (1983) collected a rich dataset on murder cases tried in Georgia during 1973–1978 and used a variety of methods to identify cases which were similar in terms of factual characteristics but were judged differently by various courts. We share with this study our interest in excessive penalties imposed in the first trial and in subsequent reversal decisions. We also obtain qualitatively similar findings regarding the role of victim's race. However, we differ on methodological grounds. While Baldus, Pulanski, and Woodworth (1983) control for a rich set of observable characteristics, their approach cannot deal with unobserved heterogeneity: it is this shortcoming that our analysis tries to address. More recently, Blume, Eisenberg, and Wells (2004) combine data on death row cases for eight US states with homicide data for the same states over the period 1976–1998 and find that murders involving black defendants and white victims are significantly more likely to result in death sentences than white defendant-white victim murders. On the other hand, they find that black defendant-black victim cases are significantly under-represented on death row. Compared to this literature, our test is not subject to the omitted variable bias critique (under the assumption of the model). At the same time, our test has a more limited scope, in that it applies to cases that have received the death sentence in the first trial, and cannot estimate bias occurring from exclusion errors, i.e., cases that should have received a death sentence and did not.

The paper is organized as follows. Section I offers a brief synthesis of the institutional details useful to understand judicial errors in capital cases in the United States. Section II describes our model of behavior of the court and derives our test of racial bias. Section III describes the data. Section IV presents our empirical methodology and results. The last section concludes.

⁵ On the deterrence effect see among others, Ehrlich (1975); Katz, Levitt, and Shustorovich (2003); and Donohue III and Wolfers (2005) for a review.

I. Institutional Background

Today, 34 states in the United States allow capital punishment.⁶ Each state has its own statute but much similarity exists among them. Most statutes are in fact modelled around the Georgia one approved by the supreme court in *Gregg v. Georgia*, 428 U.S. 153 (1976). That statute prescribed: (i) an independent trial of guilt or innocence; (ii) a second hearing solely to determine the sentence; (iii) a finding of at least one aggravating circumstance; (iv) an automatic review by the Georgia Supreme Court; and (v) the comparison to similar cases. Even though the statutes are similar, the actual application of the death penalty varies greatly across states.

First Trial and Sentencing.—Trials for capital crimes embed two stages: the guilt phase, where the jury deliberates whether the defendant is guilty or not, and the sentencing stage where, if the defendant is found guilty, the jury (or in some states and until 2002, the judge) weighs the aggravating and mitigating factors presented by the prosecutor and the defense and determines the sentence. The supreme court has ruled that no statute can prescribe mandatory capital punishment; that is, no one found guilty of a capital crime can be automatically sentenced to death.⁷ This implies that the jury always has discretion in choosing between a death sentence or imprisonment, if the defendant is found guilty. A death sentence requires the existence of at least one aggravating circumstance and the consideration of applicable mitigating factors. What constitutes both vary from state to state. Certain aggravating circumstances or mitigating factors are very clear, like killing a police officer (aggravating) or killing under a certain age (mitigating). But other factors are much less clear cut, like a murder being “in cold blood and pitiless” (aggravating) or “acting under duress” (mitigating). The supreme court has struggled with unclear and vague definitions of aggravating circumstances and mitigating factors but quite a large latitude remains. About 1 percent of the murders committed in a year end up in a death sentence.⁸

The Appeal Process.—The most important aspect of capital punishment procedural rules for the purpose of our study is that all capital sentences, with no exceptions, are automatically appealed in state high courts. During our study period, in all but two states the appeal runs directly from the trial court to the state court of last resort (typically the state supreme court), while in Alabama and Ohio it went through an intermediate court of criminal appeals before reaching the highest court. Sentences which survive state direct appeals are then inspected by state post-conviction courts and, if they survive this stage too, they can be reviewed in federal habeas corpus petitions. The process often lasts several years. At each stage, the appeal court can overturn the sentence if “serious error” is found: i.e., “error that substantially undermines the reliability of the guilt finding or death sentence imposed at trial” (Liebman, Fagan, and West 2000). When all appeals are exhausted, the only hope left for the defendant is an act of clemency from the state governor.

⁶The Federal Government has two death penalty statutes, one for the military and the other for nonmilitary crimes. This section draws on Coyne and Entzeroth (2006).

⁷*Roberts v. Louisiana*, 428 U.S. 325 (1976) and *Woodson v. North Carolina*, 428 U.S. 280 (1976).

⁸See Barnes, Sloss, and Thaman (2009) for a recent discussion of criteria according to which prosecutors pursue the death penalty in Missouri.

Liebman, Fagan, and West (2000) conducted a study of all 4,578 state capital appeals in the period 1973–1995, plus 248 state post conviction reversals and 599 capital sentences reviewed by federal habeas corpus courts in the same period. Their findings were striking: between 1973 and 1995, the proportion of fully reviewed capital judgments in which *serious error* was found and which were overturned at one of the three stages was 68 percent. This is what happened to the sample of overturned cases that they analyze at retrial: 7 percent were found to be innocent, 75 percent were resentenced to less than death, and 18 percent were resentenced to death. These sentence reversals will play a key role in our empirical test.

II. The Model

A. Setup

We consider a defendant whom a court can condemn to the death penalty or to a lesser penalty (which includes the case in which the defendant is set free). If the court decides for the death penalty, there is an appeal. In case of a lesser sentence or a no guilt verdict there is no appeal and the decision stands. In appeal, the superior court can either confirm the death penalty or reverse the decision of the lower court because of errors. An error can occur in establishing the guilt of the defendant or in sentencing the death penalty for a crime that did not warrant it. In the theoretical model we assume that there are only two courts: a trial court that can make mistakes and can be biased, and a superior court that is unbiased and—in the benchmark version of our model—makes no mistakes (we relax the latter assumption below).

Before formally presenting the setup of the model, it is useful to introduce two definitions that we will use to simplify the exposition and that constitute a slight abuse of terminology.

DEFINITION 1: *We label “guilty” a defendant who has committed a crime deserving of the death penalty according to the law. We define “innocent” a defendant who has committed no crime or has committed a crime which does not deserve the death penalty.*

Thus sentencing an “innocent” may mean sentencing to the death penalty someone for a crime which instead deserved, say, a life term. Empirically, as we shall see below, most of the errors imply sentencing to death somebody guilty of a crime not deserving death.

DEFINITION 2: *We denote as “court” the combination of judge and jury.*

While in reality judge and jury are obviously distinct, our data does not allow to separately test for bias among the two; hence, we consider them jointly in the model because this matches the empirical test we conduct.⁹

⁹This implies that when we test for racial bias we observe a combination of possibly two biases: of the jury and of the judge. If these go in opposite directions, our test may fail to detect bias even when such bias exists in the individual components of the court.

We assume that each crime involves one defendant and one victim. Defendants are characterized by their race and by a set of characteristics of the person or the crime or the relationship between the person and crime. Let $r \in \{w, m\}$ be the race of the defendant, where w stands for “white” and m for “minority.” Let’s define the race of the victim as $R \in \{W, M\}$ where W stands for white and M for minority. The court observes various signals about the characteristics of the defendant and of the crime, provided by the police and the trial proceedings and summarizes them in a single dimension which we denote with x and that we can denote as “the evidence.” We normalize the support of x to $x \in [0, 1]$. The distribution of evidence can depend upon the race of the defendant, for instance because of the quality of his/her legal assistance: if minority defendants (on average poorer) have a lower quality defense, they may carry a less precise signal and face more errors against them in the first trials.¹⁰ While the quality of defense is not explicitly modeled in our framework, it can be incorporated in a different distribution of evidence faced by the courts for minority and white defendants. Let g stand for “guilty” and n stand for “not guilty.” We denote with $F_g^r(x)$ the distribution of evidence if the defendant is guilty and with $F_n^r(x)$ the one for non-guilty defendants. In the benchmark derivation of our test we allow $F_g^r(x)$ and $F_n^r(x)$ to depend on the race of the defendant but not on the combination of defendant and victim’s race. In Section IIC we discuss possible ways of relaxing this hypothesis and the implications for interpreting our test.

We assume that the signal is informative for the court and the densities $f_g^r(x)$ and $f_n^r(x)$ satisfy the strict monotone likelihood ratio property (MLRP), that is:

$$\text{MLRP: } f_g^r(x)/f_n^r(x) \text{ is strictly increasing in } x, \text{ for } r \in \{w, m\}.$$

This property implies that higher values of the signal x are associated with a relatively higher probability of guilt. We also assume $f_g^r(x)/f_n^r(x) \rightarrow +\infty$ as $x \rightarrow 1$.

The Problem of the Court.—An individual considering whether to commit a capital crime (crime in short) compares the costs and benefits of it. In the online Appendix we model this problem, but for the purpose of our test we can equivalently assume that there exists an exogenous probability π^r that an individual of race r commits a crime. We allow this probability to vary across defendant races. In Section IIC we discuss whether π is likely to also depend on victim’s race and how this may affect our results.

The trial court wants to sentence guilty defendants to the death penalty, but wants to avoid the mistake of sentencing defendants who do not deserve the death penalty (labeled as *innocent*).¹¹ These considerations can be summarized by assuming that the court minimizes a weighted average of the probability of condemning an

¹⁰Results by Iyengar (2007) indeed suggest that this may be the case. When comparing the effectiveness of two types of defense lawyers provided for indigent defendants, namely public defenders or court private lawyers compensated by the hours, she finds that the former perform better and minority defendants are disproportionately represented by the latter.

¹¹Note that under the assumption that the higher courts never make mistakes the costs of sentencing an innocent for the lower courts are the moral costs of inflicting high costs to innocent defendants and the costs of reputation losses of having made mistakes. In the more general version of the model in which even superior courts can make mistakes, lower courts also have to worry about the possibility that innocent people sentenced to death are never released.

innocent (type I error) and the probability of letting a guilty person free (type II error). Therefore the court chooses the optimal x_{rR} as the value of x which solves

$$(1) \quad \min_{x_{rR}} \{ \alpha_{rR} [1 - \pi^r] [1 - F_n^r(x_{rR})] + (1 - \alpha_{rR}) \pi^r F_g^r(x_{rR}) \},$$

with $0 < \alpha_{rR} < 1$, $r \in \{w, m\}$, and $R \in \{W, M\}$. The first and second term in (1) are, respectively, the type I and type II errors. The parameter α_{rR} represents the relative weight given by the court to type I error and will be crucial for defining our test of racial bias. The court chooses to sentence the defendant to death if the evidence is above a certain threshold. We define the threshold x_{rR} , which as indicated in the notation could vary with the race of the defendant and of the victim, allowing the court to choose four potentially different thresholds. The probability that an individual of race r killing an individual of race R is sentenced to death is: $p_r(x) = \Pr(x \geq x_{rR})$. The optimal decision of the court in a case involving a defendant of race r and a victim of race R is to impose a death sentence if and only if the signal x exceeds the threshold x_{rR}^* given by the court's first-order condition

$$(2) \quad \frac{f_g^r(x_{rR}^*)}{f_n^r(x_{rR}^*)} = \frac{\alpha_{rR}}{1 - \alpha_{rR}} \frac{1 - \pi^r}{\pi^r}.$$

The cutoff value x_{rR}^* is thus the *standard of proof* applied by the court. Inspection of (2) immediately reveals that x_{rR}^* is increasing in α_{rR} , i.e., the higher the relative concern about condemning an innocent, the higher will be the standard of proof required before imposing a death sentence.

Note that even if $\alpha_{rR} = \alpha$ for any r and R , the equilibrium cutoff point chosen by the court can be, and in general will be, different for white and minority defendants. In fact, the left-hand side of (2) depends on the race of the defendant r , because f_g^r and f_n^r depend on r ; and the right-hand side of (2) depends on r through π^r . For example, if minorities have a higher probability of committing crimes, *ceteris paribus* the court will choose $x_{mR}^* < x_{wR}^*$. On the other hand, under the assumptions of our model the only way the race of the victim R can enter (2) is through the parameter α_{rR} . This is the key insight upon which our test is based.

B. A Test of Racial Bias

The spirit of the test we develop can be summarized as follows. Bias in the adjudication system induces differences in the strength of the evidence (the standard of proof x^*) beyond which a death verdict is imposed. In the absence of bias, this threshold may depend on defendant's race r but not on defendant and victim's race combined (r, R). In other words, the standard of proof should not systematically differ for—say—minority defendants who kill whites and minority defendants who kill nonwhites. However, the threshold x^* is not observable, hence we construct a mapping from this threshold to average overturn rates and exploit the fact that, under the assumptions of our model, this mapping is independent of the particular combination of defendant and victim's race (r, R)—though it may depend on defendant's

race alone. Hence we can use differences in average overturn rates across victim's races—conditional on defendant's race—to infer bias. In what follows we formally define bias in the context of our model and derive our test.

In our model a court could be biased in different ways.

DEFINITION 3 (Bias only on the race of the defendant): $\alpha_{rR} = \alpha(r)$.

α depends upon the race of the defendant (r), but not the race of the victim (R). A bias against minority defendants is represented by $\alpha_{mM} = \alpha_{mW} < \alpha_{wM} = \alpha_{wW}$, as the court places less weight on the possibility of wrongly condemning a minority defendant.

DEFINITION 4 (Bias only on the race of the victim): $\alpha_{rR} = \alpha(R)$.

α depends on the race of the victim (R), but not on the race of the defendant (r). A bias in favor of white victims is represented by $\alpha_{mW} = \alpha_{wW} < \alpha_{mM} = \alpha_{wM}$, as the court places less weight on the possibility of wrongly condemning someone who has killed a white victim.

DEFINITION 5 (Bias on both the race of the victim (R) and the race of the defendant (r)): $\alpha_{rR} = \alpha(r, R)$.

The court applies a differential treatment on the basis of the race of the victim, but it further differentiates depending on who has killed that victim. For example, a situation where the court treats minority defendants who have killed white victims more harshly than whites who have killed whites, but is relatively and equally lenient on both if the victim is nonwhite, can be represented as: $\alpha_{mW} < \alpha_{wW} < \alpha_{mM} = \alpha_{wM}$.

We now derive a test for racial bias. Define the capital sentencing rate for r, R pairs as:

$$(3) \quad \gamma(r, R) = \pi^r [1 - F_g^r(x_{rR}^*)] + [1 - \pi^r] [1 - F_n^r(x_{rR}^*)].$$

The equilibrium error rate on r, R pairs, which we denote as $E(r, R)$, is

$$(4) \quad E(r, R) = 1 - \frac{\pi^r [1 - F_g^r(x_{rR}^*)]}{\gamma(r, R)},$$

with $\gamma(r, R)$ given by (3). As shown in the Appendix, under MLRP $E(r, R)$ is monotonically decreasing in x_{rR}^* : the higher the standard of proof, the lower the error rate.

The parameter α_{rR} enters the error rate (4) through the optimal threshold x_{rR}^* derived from (2), allowing us to build a test of racial bias. The race of the victim R only affects x_{rR}^* if α_{rR} depends on R . Therefore if the court does not discriminate over the race of the victim, the error rate should be independent of it, even though it could depend upon the race of the defendant. If for given race of the defendant we find higher error rates in cases involving white victims compared to minority victims,

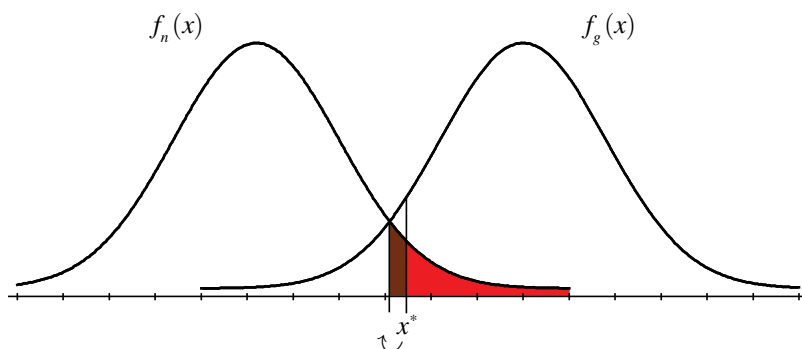


FIGURE 1. ERROR RATE FOR GIVEN DEFENDANT'S RACE

this is evidence of racial bias in favor of white victims. In the benchmark version of our model, this bias originates from a lower weight given to type I errors in cases involving a white victim ($\alpha_{rW} < \alpha_{rM}$), which led the court to apply a lower standard of proof, *ceteris paribus*. The intuition for this result is illustrated in Figure 1, which shows the density functions of the signal x for non-guilty and guilty defendants, holding constant the race of the defendant.

The type I error is the shaded area to the right of the threshold x^* . When we hold constant the race of the defendant and vary the race of the victim, the assumptions of our model imply that the distributions f_n^r and f_g^r do not shift. Hence the only way one could obtain different error rates is through a shift in the value of x^* , which in turn must reflect different values of α . Figure 1 shows an example where α decreases (e.g., less weight is given to type I errors against defendants who killed white victims), the threshold x^* moves left, and the error rate increases.

On the other hand, a similar inference cannot be made if one holds constant the race of the victim and compares errors against defendants of different races. In other words, we cannot derive from our model the implication that if the court does not discriminate across defendants we should find the same error rate for white and minority defendants when we hold constant the race of the victim. Figure 2 illustrates the point.

The top and bottom panels of Figure 2 display, respectively, the signal distributions for white and minority defendants, holding constant the race of the victim. Because our model allows f_n^r and f_g^r to differ according to the defendant's race r , error rates (the shaded areas in Figure 2) may differ even when the court is unbiased and selects the same threshold x^* .¹²

Thus we cannot test for the presence of bias only on the race of the defendant, while we can test for bias which depends on the race of the victim (bias type 2 and 3 in the definition above). We now derive a test for the relatively more conservative definition that bias is purely a function of victim's race (bias of type 2). This

¹²The fact that the threshold x^* is the same for both defendant races is neither a necessary nor a sufficient condition for unbiasedness of the court. It is used in Figure 2 purely for illustrative purposes. Also, the fact that we allow the probability of guilt π^r to differ across races implies that we could not make inference on bias by comparing errors across defendant races even if the signal distributions were the same.

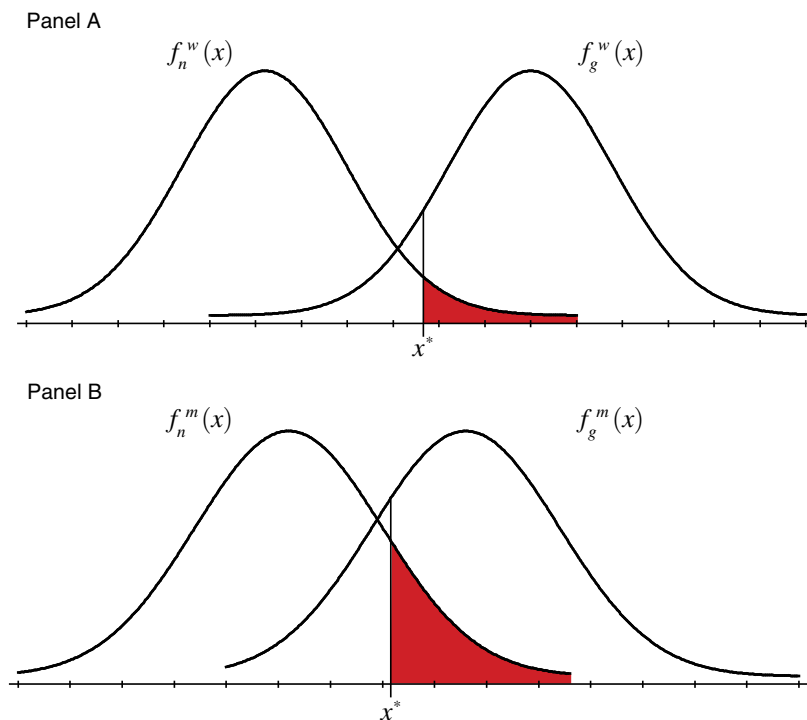


FIGURE 2. ERROR RATE FOR GIVEN VICTIM'S RACE

amounts to asking the question: in the presence of bias related to the victim's race, does this bias affect defendants of different races in the same way? In the empirical section we will show that the answer is no, and that according to our results the behavior of the court is consistent with a bias that depends on the particular combination of defendant and victim race (bias type 3).

PROPOSITION 1: *If $\alpha_{rR} = \alpha_R$ independent of r , then the ranking of average error rates $E(r, M)$ and $E(r, W)$ should not depend on r , for $r \in \{m, w\}$.*

PROOF:

Suppose without loss of generality that $\alpha_{mW} = \alpha_{wW} < \alpha_{mM} = \alpha_{wM}$. Consider first minority defendants. Because x_{rR}^* is strictly increasing in α_{rR} , $\alpha_{mW} < \alpha_{mM}$ implies $x_{mW}^* < x_{mM}^*$, which in turn implies $E(m, W) > E(m, M)$ due to the fact that $E(r, R)$ is strictly decreasing in x_{rR}^* . The same reasoning applies to white defendants, with $\alpha_{mW} = \alpha_{wW} < \alpha_{wM} = \alpha_{mM}$ implying $E(w, W) > E(w, M)$.

Thus the following condition must hold if the court discriminates on the basis of victims' race but treats defendants of different races in an unbiased way:¹³

$$(5) \quad E(m, W) > E(m, M) \Leftrightarrow E(w, W) > E(w, M).$$

¹³ Expression (5) refers to the case of bias in favor of white victims. Obviously for bias in favor of minority victims the inequalities should be reversed.

Expression (5) says that if we find a higher error rate on minority defendants who killed white victims, compared to minority defendants who killed minority victims, then we should also find a higher error rate on white defendants who killed white victims than on white defendants who killed minority victims, and vice versa. This forms the basis for our rank-order test.

Our use of rank order tests is in the same spirit of Anwar and Fang (2006), with the difference that in their case one of the two dimensions over which troopers' success rates are computed pertains to the behavior of the agent who may be discriminating (i.e., the police officer), while in our case the two dimensions pertain to offender and victim characteristics, and not features of the court.

C. Discussion and Extensions

In this section we discuss some important assumptions underlying our model, possible extensions, and how they affect the interpretation of our results. We start with the probability of guilt π_r , then move to the shape of the signal distributions $F_n^r(x)$ and $F_g^r(x)$, and finally discuss the behavior of superior courts and the role of plea bargain.

Different Probabilities of Guilt.—It is well known that the frequency of homicides varies depending on the combination of defendant and victim's races: namely, that intraracial homicides are more frequent than interracial ones. However, the parameter π_r in our model does not represent the *frequency* of crimes but the probability of *guilt*, or more precisely the likelihood that the defendant has actually committed a crime that deserves the death penalty. The question then is: does this probability depend on the four combinations of defendant and victim race in a way that would invalidate our test, leading us to attribute to bias differences in errors which arise from differences in π_{rR} ? To answer this question we first present some available empirical evidence and then discuss from a theoretical point of view under what configurations differences in π_{rR} would create a problem for interpreting our results.

We are not aware of a comprehensive dataset including information on guilt rates by race of defendant and victim for the period under study. So we resorted to two alternative sources.¹⁴ The first is a representative sample of murders adjudicated in 1988 in 33 of the largest counties in the United States (US Department of Justice 1996). This dataset includes information on race of the defendant and of the victim, as well as the final disposition outcome of the case. We restricted the sample to first degree murders, which are the ones potentially eligible for the death penalty, and to cases that underwent trial, to provide as close a benchmark as possible to our dataset. To investigate whether the differential incidence of intraracial murders may be indicative of differences in π_{rR} , we regressed the average guilt rate in the county on the fraction of murders in which defendant and victim belonged to the same race group. The estimated coefficient was -0.05 , with a standard error of 0.31. Figure 3

¹⁴ Other than these two sources, to the best of our knowledge existing public use data typically include defendant's race (and in some cases the disposition of the case) but not victims' race. Datasets with information on defendant's and victim's race typically do not have information on the disposition of the case, which means we could not calculate guilt rates.

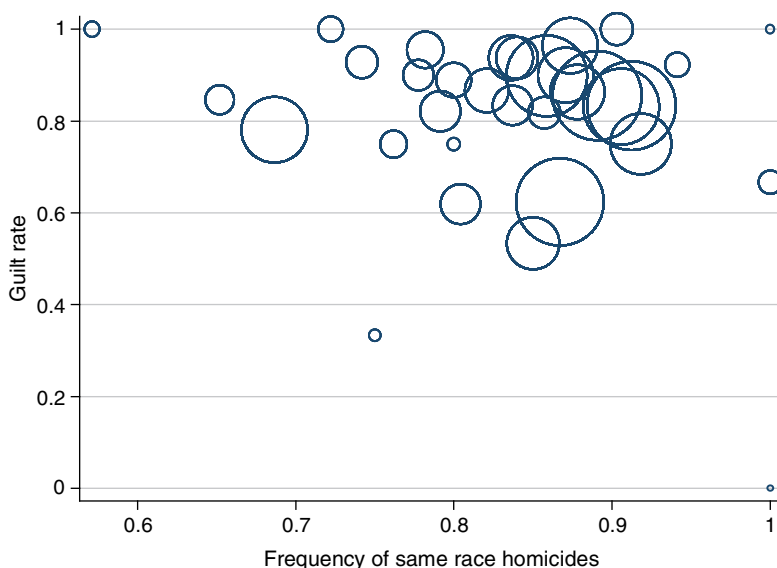


FIGURE 3. GUILT RATE AND FREQUENCY OF HOMICIDES

plots the raw data (with the size of the circles proportional to the number of homicides in the county) and clearly shows that there is no correlation between the two variables. We also exploited the individual level data and regressed the probability that the defendant was found guilty on the race of the defendant, the race of the victim, and an interaction of the two. The coefficient on the latter term was again not statistically different from zero.¹⁵

A second source of empirical evidence is the work of Blume, Eisenberg, and Wells (2004). They calculate death sentence rates for eight US states over the period 1977–2000 as the ratio of number of death sentences and number of murders in the state. While this is not the empirical equivalent of our guilt rate (among other things, because the numerator includes biased decisions of trial courts), a rough correction can be applied by subtracting state and race-pair specific relief rates (which we computed in our data) and assuming that the share of guilt sentences that were affirmed are correct. Again, looking at the pattern of guilt rates computed in this way no systematic correlation emerges between same race pairs and guilt rates.¹⁶

Even though differences in π_{rR} do not seem warranted on the basis of the above evidence, it is useful to discuss in what direction our results would be affected if such differences were present. This amounts to assessing the sign of the derivative of the error rate (4) with respect to π . In the online Appendix we show that this derivative can be decomposed in two parts. The first includes a *direct effect* of π on

¹⁵Specifically, our estimated linear probability model is

$$\text{Guilty} = \frac{0.755}{(0.089)} + \frac{0.013}{(0.091)} ND + \frac{0.116}{(0.086)} WV - \frac{0.049}{(0.086)} (ND \times WV),$$

where ND is a dummy for nonwhite defendant, WV is a dummy for white victim, and standard errors (in parentheses) are clustered at the county level. Results are very similar if we run the regression on the full sample of first degree murders (including cases that did not go to trial and guilty pleas) or if we include county fixed effects.

¹⁶Results available from the authors. Note that we do not have access to the raw data used by Blume, Eisenberg, and Wells (2004), so we conducted this meta-analysis using the published data in their paper.

the error E that would obtain if x were treated as exogenous. This effect is negative and captures the fact that, for given standard of proof x , higher values of π imply that fewer of the people condemned will be innocent, hence the error rate is lower. The second part of the derivative is an *indirect effect* that works through endogenous changes in x . To see how, notice that x is decreasing in π (from (2) and MLRP), and E is decreasing in x , hence the indirect effect pushes in the direction of $\frac{\partial E}{\partial \pi} > 0$. Intuitively, this occurs because when π increases the court chooses a lower standard of proof, which—for given π —increases the size of the type I error. Which of the two effects dominates depends on the value of π and on the shape of the functions $F_n^r(x)$ and $F_g^r(x)$. In the online Appendix we provide some simulations where the direct effect dominates when the f_n distribution is relatively more dispersed than f_g .¹⁷

In summary, what matters for our test is the cross pattern of guilt rates for the four pairs of defendant and victim races. Contrary to perceptions based on the higher *frequency* of intraracial homicides, the (admittedly scant) evidence that we are aware of does not allow to establish any systematic pattern in this respect. In the remainder of the paper we maintain the assumption that π does not systematically vary with the combination of defendant and victim race.

Bias in the Collection of Evidence.—A second important assumption underlying our test is that the functions $F_n^r(x)$ and $F_g^r(x)$ do not depend *jointly* on the race of the defendant and of the victim.¹⁸ Given our test the critical question is whether, in particular, $F_n^{mW}(x)$ differs from $F_n^{mM}(x)$ and $F_g^{mW}(x)$ differs from $F_g^{mM}(x)$. The two pairs of functions may not be the same for two reasons. One is the nature of the crime: the types of crimes involving defendant/victim pairs of different race may be *objectively* different and lead to different type I errors for any given level of evidence x . The second reason is that there may be no intrinsic difference in the nature of the crime but the combination of police work, prosecutor work, and defense attorneys' work leads the court to face different distributions as a function of the defendant/victim pair.¹⁹

Let's begin with the second case. More precisely let's assume that if the police, prosecutor, and defense attorney were unbiased, the distributions would indeed not depend on the race of the victim. We denote these as the *unbiased* distributions $F_n^r(x)$ and $F_g^r(x)$. Suppose now that a biased "police, prosecutor, or defense" (PPD) distorts the information available to the court so that the distributions the court uses in its optimization problem, hence in (2), are instead $\tilde{F}_n^{rR}(x)$ and $\tilde{F}_g^{rR}(x)$. How might these differ from the true ones and what would be the implications for our test? To answer this question, let us consider the example of a murder involving a minority defendant and a white victim. Figure 4 reports the unbiased distributions for this case, $f_n(x)$ and $f_g(x)$ (for simplicity we omit superscripts) and denotes as x^* the solution of an unbiased court in this case.

¹⁷ High relative dispersion of f_n implies that decreases of x in response to increases of π translate into relatively small increases in the area corresponding to the type I error as compared to the decreases in the type II error, resulting in a decrease of the overall error rate. Conversely, in cases where f_g has much fatter tails than f_n the opposite effect may prevail.

¹⁸ A similar restriction is common to Anwar and Fang's (2006) test of prejudice.

¹⁹ Empirically, Radelet and Pierce (1985) analyzed a sample of 1,017 homicides in Florida in the period 1973–1977 and compared the descriptions of the homicides in police reports with the (later) descriptions given by courts. They found that homicides involving African American suspects and white victims were more likely to be described as *felony* by prosecutors.

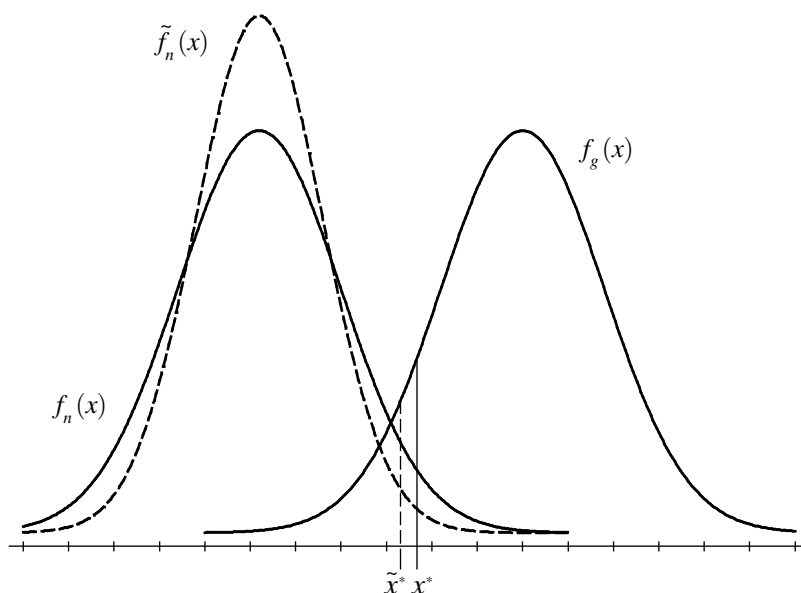


FIGURE 4. BIAS IN THE COLLECTION OF EVIDENCE

A biased PPD may want to convince the court that for a given threshold of evidence the probability of convicting an innocent is (artificially) low compared to the risk of acquitting a guilty person. In the figure this is represented by the function $\tilde{f}_n(x)$. In correspondence of x^* the size of the *perceived* type I error has now decreased relative to type II (which is unchanged in the example). An unbiased court would thus adjust the standard of proof downwards and choose a threshold $\tilde{x}^* < x^*$. This would result in a higher type I error ex post when the appeal court has available the unbiased distribution $f_n(x)$. In other words, the court would commit more type I error given the *true* distributions while acting on the *wrong* ones. In this case the observed higher error rate for the minority/white pair would be the result of a PPD bias, not a court bias. Obviously, if the court were also biased the observed error rate would be the combination of the two. What this discussion suggests is that if we allowed the PPD to manipulate the distribution of the signal differentially depending on the combination of defendant and victim race, we would need to qualify the interpretation of our test. The observed error would not necessarily derive from a bias in the parameter α (i.e., it may not necessarily be attributable to the court). But one would still consider this error as resulting from racial bias in the criminal justice system.

Consider now the other case, namely PPD is unbiased but $F_n^{mW}(x)$ differs from $F_n^{mM}(x)$ and/or $F_g^{mW}(x)$ differs from $F_g^{mM}(x)$ because the type of crime is objectively different. For instance murders with minorities killing whites may involve worse crimes or cases where the evidence is less clear-cut, something which would go against the validity of our test. In this case more errors might be made by the courts simply because of the nature of the crime. One way of addressing this issue is to examine empirically whether observable indicators of the nature of the crime (e.g., aggravating and mitigating factors) differ systematically across racial pairs, and test if our results are robust to conditioning on observable characteristics. This should

give some insights about the potential bias due to differences in unobservables. This is what we do in Section IVB.

To sum up, we have discussed two reasons why the assumption that $F_n^r(x)$ and $F_g^r(x)$ do not depend on R may be violated. One reason relates to the behavior of the police, prosecutor, or defense attorney, who may distort the *true* distribution of the signal in a way that produces more type I errors. This possibility requires that we interpret the higher error not as resulting from a bias of the court, but from a bias of another part of the criminal justice system. But it is bias nonetheless. The second potential violation occurs if the severity of the crime is objectively different in inter-racial homicides. This case could generate higher error even if all parts of the criminal justice system are unbiased. We cannot rule out this possibility on theoretical grounds, but we assess empirically how serious a concern it may be and find results that we believe are supportive of our interpretation.

Bias in the Decisions of Superior Courts.—Another hypothesis in our model is that superior courts are racially unbiased. We can relax this assumption in one direction. Suppose that superior courts were racially biased in the same direction of lower courts. This would go against finding higher error rates on certain racial pairs because the superior courts would simply reaffirm the first sentence. That is, if we did not find evidence of racial bias based upon our test it could mean that the same bias applies to all levels of courts. Thus not finding a bias could be inconclusive but finding it would not. Note that if the racial bias declines with subsequent stages of revision (from state courts to federal habeas corpus courts) then we should find that the difference in errors rates across pairs of defendant's and victim's race should become larger in later stages of appeal. This is what we find below.

What we cannot allow in our model is that superior courts are biased in the opposite direction to lower courts, because in this case higher error rates may be interpreted as *reverse discrimination* rather than evidence of mistakes by lower courts. We are not aware of a literature that documents such bias in opposite directions, and at the same time the pattern of inequalities that we find in our tests (higher error rates on cases in which defendant and victim are from different racial groups) would require a particular pattern of bias by superior courts, not in favor of a particular group but of specific *pairings* of races. Having said this, in the empirical part of the paper we try to address the possibility of bias by superior courts by testing if our results depend on certain characteristics of the appeal court (e.g., political orientation). We do not find evidence that the pattern of reversal we uncover is driven by the ideology of the appeal judges.

Unbiased but Imperfectly Informed Superior Courts.—We can also relax the assumption of infallibility of the superior court. More specifically, we assume that the latter has the same information available to the trial court: i.e., it does not receive additional signals and it has the same distributions $F_n^r(x)$ and $F_g^r(x)$, but it is unbiased.²⁰ The appeal court minimizes

$$(6) \quad \min_{x_{rR}} \{ \alpha^s [1 - \pi^r] [1 - F_n^r(x_r)] + (1 - \alpha^s) \pi^r F_g^r(x_r) \},$$

²⁰ An intermediate assumption which we do not develop formally here is that the superior court is unbiased and has superior (but still imperfect) information relative to the trial court.

where S stands for “superior” and α^s is the relative weight placed by the superior court on type I errors. This weight is independent of race because the superior court is assumed to be unbiased. If we want to test whether minorities who kill white victims receive harsher treatment, the empirically relevant scenario is one in which $\alpha^s \geq \max\{\alpha_{mW}, \alpha_{wW}\}$: i.e., the superior court cares about possibly condemning innocent defendants in cases with white victims at least as much as the lower court does. This is because if α^s were lower than the trial court’s α s, the superior court would not overturn any sentence (see below).²¹

Let us denote with $x_{r,2}^*$ the evidentiary threshold chosen by the superior court: i.e., the solution to (6), and with $x_{rR,1}^*$ the threshold chosen by the trial court.²² From our earlier results we know that x^* is increasing in α , hence under our assumptions we have $x_{r,2}^* \geq x_{rW,1}^*, \forall r$: in cases involving white victims the superior court requires (weakly) stronger evidence than the trial court to convict a defendant (affirm a sentence). This implies that the superior court will reverse sentences based on evidence $x > x_{rW,1}^*$ (i.e., sentences imposed by the trial court because the evidence exceeded their standard of proof), but with $x \leq x_{r,2}^*$ (i.e., the evidence falls short of the stricter standards used by the appeal court). In terms of Figure 5, these reversals will comprise the areas below f_n^r and f_g^r in the interval $(x_{rW,1}^*, x_{r,2}^*]$. In reversing these sentences, the superior court *corrects* some of the type I errors, but makes more type II errors.

The error rate in our data is the fraction of death sentences imposed by the trial court that are reversed by the superior court:

$$(7) \quad E(r, R) = 1 - \frac{\pi^r[1 - F_g(x_{r,2}^*)] + (1 - \pi^r)[1 - F_n(x_{r,2}^*)]}{\pi^r[1 - F_g(x_{rR,1}^*)] + (1 - \pi^r)[1 - F_n(x_{rR,1}^*)]}.$$

Expression (7) differs from the earlier error rate (4) because the numerator now captures the fallibility of the superior court: not all type I errors are corrected, but only those with $x \leq x_{r,2}^*$, and some type II errors are now committed.

However, one can easily see that, for given $x_{r,2}^*$, (7) is decreasing in $x_{rR,1}^*$. Thus, the lower is the threshold of evidence set by the trial court the larger the error rate, namely the larger the fraction of cases reversed by the superior court. This is the condition used in the proof of Proposition 1, which means that our rank-order test still holds in this case. Intuitively, if the trial court is relatively biased against the mW combination of defendant/victim, it will choose an especially low $x_{mW,1}^*$; this will generate a high error rate, that is, the superior court—even when imperfectly informed—will reverse a relatively large fraction of cases with this combination of defendant and victim.

Plea Bargain.—Many potential capital cases are plea bargained. The strength of the evidence against the defendant and the severity of the crime are critical factors in determining the incentive for defense and prosecution to pursue a plea bargain.

²¹ We do not need to impose conditions on α_{mM} and α_{wM} because, if we want to test for bias in the direction we are interested in (higher error rates with white victims), we can allow the superior court to make zero reversals in cases involving minority victims.

²² Note that since the $F(\cdot)$ functions and π depend on the defendant’s race, the threshold of evidence chosen by an unbiased superior court (i.e., with identical α for all race combinations) still depends on the race of the defendant.

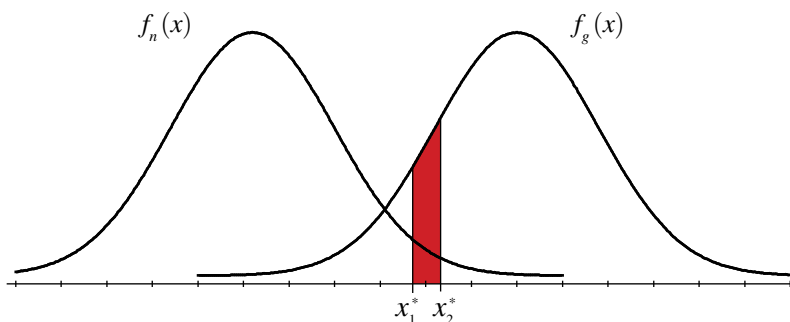


FIGURE 5. IMPERFECTLY INFORMED APPEAL COURT

Comprehensive empirical studies of the nature and characteristics of plea agreements are hard to come by due to data limitations. There is evidence that minority defendants and defendants with previous criminal history receive a harsher plea bargained prison term (Humphrey and Fogarty 1987). Models of plea bargain typically involve asymmetric information between prosecutor and defendant about the strength of the case, as in for instance Grossman and Katz (1983) and Reinganum (1988). In our model we do not have this asymmetry and only with this extension (which we leave for future research) we could incorporate plea bargain in a meaningful way. As far as our empirical test is concerned, if the likelihood that a case is plea bargained were uncorrelated to the races of the racial pair defendant/victim, our test would be unaffected. If it were not, then this correlation might introduce a bias, but the direction of the bias is unclear, as it would depend among other things on the shape of the signal distribution.

We tried to empirically assess the potential relevance of this source of bias using data on a representative sample of murders adjudicated in 1988 in 33 of the largest counties in the United States (US Department of Justice 1996). This dataset includes information on race of the defendant and of the victim, as well as the final disposition outcome of the case (among which guilty plea). When we regress the likelihood of guilty plea on race of the defendant, race of the victim, and the interaction of the two, the coefficient on the interaction is not statistically different from zero.²³

III. The Data

To implement our test of racial bias we could not rely on any readily available dataset. In fact all existing datasets containing information on the race of the defendant and of the victim in capital cases have limited geographical and temporal coverage and—most importantly for our purposes—do not contain information on whether the capital sentence was reaffirmed in appeal. The only comprehensive dataset containing information on judicial errors in capital cases, that is, the one used in Liebman,

²³Specifically, our estimated linear probability model is

$$Plea = \frac{0.346}{(0.060)} + \frac{0.017}{(0.068)} ND + \frac{0.046}{(0.064)} WV - \frac{0.005}{(0.075)} (ND \times WV),$$

where ND is a dummy for nonwhite defendant, WV is a dummy for white victim, and standard errors are clustered at the county level. The results are similar if we include county fixed effects.

Fagan, and West's (2000) study and compiled by Fagan and Liebman (2002)—henceforth, FL—does not contain information on the race of the defendant nor on the race of the victim. We therefore constructed our dataset by examining each individual record in FL's data and searching for information on the race of the defendant and of the victim. For a detailed description of FL's data collection methodology and variable definition we refer the reader to Liebman, Fagan, and West (2000). In what follows we start by briefly reviewing the characteristics and scope of FL's data, then discuss our search methodology and present some descriptive statistics.

Data Coverage.—FL's data is the first systematic collection of information on capital appeals in the modern death penalty era in the United States. We use two datasets originally compiled by FL.²⁴

- Direct appeal dataset (DA): 4,546 state capital cases whose direct appeal decisions became final between January 1, 1973 and December 31, 1995.²⁵ This is the universe of all capital sentences that were reviewed on direct appeal by a state high court.
- Habeas corpus dataset (HC): 557 capital cases whose review was finalized by a federal habeas corpus court between January 1, 1973 and December 31, 1995. This is the universe of all capital sentences that were finally reviewed over this period.

After eliminating cases for which the name of the defendant could not be identified, we are left with a pool of 4,416 observations in DA and 531 observations in HC.²⁶

Definition of Error.—In FL's data, "error" is defined as such only if it led to the reversal of a capital conviction or sentence. If an error was discovered that did not result in a reversal, this is not coded as "error" in the database. For DA cases, a "serious error" that warrants reversal must have three characteristics. First, it must be "prejudicial," in the sense of affecting the outcome of the case (harmless errors do not lead to reversals). Second, it must have been "properly preserved," in the sense that the claim must have been asserted at the time and in the way required by the law. Third, obviously the error must have been discovered. At the federal HC stage,

²⁴ FL also compiled a *post-conviction* dataset which, however, is incomplete due to the fact that state post-conviction decisions are often not published and includes a selected subset of cases, *all of which resulted in a reversal*. In our analysis we therefore only employ the DA and HC datasets, which comprise the universe of available cases at those stages. Notice that there is a discrepancy between the number of observations which we cite in Section I (taken from Liebman, Fagan, and West 2000, p. 3) and the number of observations listed in this section, which corresponds to the data that is available through ICPSR (Interuniversity Consortium for Political and Social Research). We are constrained by the latter, hence the starting point for our analysis is 4,546 DA cases and 557 HC cases.

²⁵ "Became final" should be understood as "the highest state court with jurisdiction to review capital judgments in the relevant state must have taken one or two actions during the study period: (i) affirmed the capital judgment or (ii) overturned the capital judgment (either the conviction or the sentence) on one or more grounds" (Liebman, Fagan, and West 2000, p. 126).

²⁶ For 26 of the 557 cases in HC, either the sentence indicated in FL's data or the name of the defendant could not be found in LexisNexis, hence we drop those cases. In the DA dataset, the sentence could not be found for 84 of the 4,546 available cases. Also, because some observations in the DA dataset correspond to multiple sentences for the same first degree trial and we want to record error once for each trial, we use one observation per appeal-trial pair and attribute an error if it was found in the first stage appeal (the one automatically granted by all states).

a serious error is reversible if, in addition to satisfying the three conditions required for DA, it violates the federal Constitution.²⁷

Collection of the Race Variables.—FL's data does not contain any information on the race of the defendant, nor of the victim. To collect such information, we relied on a number of sources including the LexisNexis database, the quarterly publication "Death Row USA" issued by the NAACP Legal Defense Fund, information from the Department of Corrections of several states, FBI UCR Supplementary Homicide Files, the CDC National Death Index, a number of websites specialized in death penalty issues, plus communications with police officers and defense lawyers.²⁸ We assembled an almost complete dataset for HC and a very extensive one for DA. To the best of our knowledge, this is the only dataset currently spanning two decades of trials for the entire United States containing information on race of defendant and victim. From the same sources we also collected information on the weapon used.

Table A1 of the online Appendix reports a tabulation of cases with missing information on the race of defendant and victim for the HC (panel A) and the DA (panel B) datasets. In the HC data, we achieved almost full coverage of the defendant's race (3 missing cases out of 531), but we are missing information on the race of the victim in 20 cases out of the 528 for which we have the race of the defendant. Thus we have a usable sample of 508 cases out of 531. In the DA data, we have information on defendants' race for 4,146 cases out of 4,416 (94 percent of the sample), and on victims' race for 3,717 cases out of these 4,146 (90 percent). Online Appendix Tables A2 and A3 contain summary statistics on the share of missing observations for victim's race by state and by year, conditional on defendant's race being known.

In online Appendix Table A4 we try to gauge the extent of possible selection in the pattern of missing data for victims' race for all cases in which we have information on the defendant's race. We report the means of several variables related to defendant, victim, and crime characteristics for the subsample in which we have information on victims' race (column 1) and the cases in which we don't (column 2). We conduct a *t*-test for the equality of means and report the *p*-values in column 3. Overall we do not find statistically significant differences across the two samples, with the exception of victim's gender in the habeas corpus data, but we show below that our results are robust to excluding female victims. This increases our confidence that there may not be a significant degree of selection on unobservables in the cases for which we have information on victim race.²⁹

A. Descriptive Statistics

Table 1 reports summary statistics on the main variables of interest in the HC (panel A) and DA (panel B) datasets.

The error rate in the HC data, measured by the variable "relief" as the fact that relief is granted at some stage of the review process, is 0.36. Regarding the race of

²⁷ Some additional technical rules for reversibility at the HC stage are listed in Liebman, Fagan, and West (2000, p. 130).

²⁸ A detailed description of the search procedure and of the sources is posted in our data .zip on the AER website.

²⁹ One possible reason for the unbalance in the gender variable is greater media coverage of murders involving women, since media coverage makes it easier for us to find information on the race of the victim.

TABLE 1—ERROR RATE FOR GIVEN DEFENDANT'S RACE

	Panel A. Habeas corpus			Panel B. Direct appeal		
	Observations	Mean	SD	Observations	Mean	SD
Relief	508	0.36	0.48	3,717	0.37	0.48
African American defendant	508	0.44	0.50	3,717	0.41	0.49
White defendant	508	0.51	0.50	3,717	0.51	0.50
African American victim	508	0.13	0.34	3,717	0.17	0.38
White victim	508	0.83	0.37	3,717	0.78	0.41
White defendant, nonwhite victim	508	0.03	0.18	3,717	0.03	0.17
Nonwhite defendant, white victim	508	0.36	0.48	3,717	0.30	0.46
White defendant, white victim	508	0.48	0.50	3,717	0.48	0.50
Nonwhite defendant, nonwhite victim	508	0.13	0.34	3,717	0.19	0.39

the defendant, 51 percent of the cases involve white defendants, 44 percent African Americans, with the remaining fraction mostly constituted by Hispanics. In contrast to the relatively even split between white and African American in the defendant's race, 83 percent of the cases involve a white victim, and only 13 percent an African American victim. Cases in which a nonwhite defendant killed a white victim constitute 36 percent of the total, as opposed to 3 percent for the cases in which a white defendant killed a nonwhite victim. The remaining cases are split between nonwhites who killed nonwhites (13 percent) and whites who killed whites (48 percent). The proportions are fairly similar for the DA sample: 37 percent of the sentences are overturned; 51 percent of the defendants are white, 41 percent are African American; 78 percent of the cases involve a white victim, as opposed to 17 percent with an African American victim. In this sample the share of nonwhite defendants who killed a white victim is 0.30.

IV. Results

The test for racial bias we derived in Section II required that, in the absence of bias against particular defendant/victim pairs, a difference in error rates for defendants of a given race depending on the victim's race should be maintained in the same direction for defendants of a different race. To implement this test we use a rank order test reminiscent of Anwar and Fang's (2006) test for prejudice.

We hold constant the defendant's race r , and compare error rates across victim's race, $R \in \{W, M\}$. Let us denote with $\overline{E(r, R)}$, the average error rate for cases in which a defendant of race r killed a victim of race R . We test the null $\overline{E(r, W)} = \overline{E(r, M)}$ (absence of racial bias) against the alternative $\overline{E(r, W)} > \overline{E(r, M)}$ (racial bias in favor of white victims) using the Z-statistic:

$$(8) \quad Z = \frac{\overline{E(r, W)} - \overline{E(r, M)}}{\sqrt{\frac{SVar_{rW}}{n_{rW}} + \frac{SVar_{rM}}{n_{rM}}}},$$

where $r \in \{w, m\}$; $SVar_{rR}$ is the sample variance of the error variable in the cases involving a defendant of race r and a victim of race R ; and n_{rR} is the number of cases involving a defendant of race r and a victim of race R , with $R \in \{W, M\}$. Under

TABLE 2—ERROR RATES BY DEFENDANT AND VICTIM'S RACE

Defendant's race	Panel A. Habeas corpus: victim's race				Panel B. Direct appeal: victim's race			
	White	Nonwhite	<i>p</i> -values	Observations	White	Nonwhite	<i>p</i> -values	Observations
White	0.358 (0.031)	0.471 (0.125)	0.809	260	0.373 (0.011)	0.395 (0.046)	0.673	1,908
Nonwhite	0.376 (0.036)	0.284 (0.055)	0.082	248	0.377 (0.015)	0.347 (0.018)	0.097	1,809
Observations	424	84			2,911	806		

Note: Standard errors of the means in parentheses.

the null hypothesis and given our large sample, Z has a standard normal distribution by the central limit theorem. We will thus reject the null in favor of the alternative if expression (8) exceeds a threshold value z_α , where α is the significance level of the test and $\Phi(z_\alpha) = 1 - \alpha$. Performing this test separately for each defendant race allows us to test the prediction of our model, expression (5).

A. Main Results

Table 2 contains the outcome of our test for the HC (panel A) and the DA (panel B) datasets and the main result of the paper. Each cell reports the average probability of error (“Relief”) for a given combination of defendant’s and victim’s race, $E(r, R)$, and the associated standard error (in parentheses). The p -values reported at the end of each row are those associated with test statistic (8) for the one-sided test of no difference in error rates against positive difference for white versus nonwhite victim. They represent the probability that, for a given defendant’s race reported in that row, a difference in the error rates between white and minority victims at least as large as the one reported can be found, given that the null (of no racial bias against defendant/victim pairs) is true.

The first row of panel A of Table 2 shows that in cases involving a white defendant the average error rate is 36 percent if the victim is white and 47 percent if it is nonwhite, with a difference of -11 percentage points.³⁰ On the other hand, in cases involving a minority defendant, the error rate is 37.6 if the victim is white, and 28.4 percent if it is nonwhite, with a difference of $+9$ percentage points (or a 32 percent increase over the nonwhite/nonwhite error rate). The differences in error rates across victim’s race thus go in opposite directions depending on the defendant’s race. For the cases involving minority defendants, we reject the null of no difference against the alternative of a positive difference in error rates (white minus nonwhite victim) with a p -value of 0.08 ; for cases involving white defendants we fail to reject the null against the same alternative (p -value 0.81). Based on our rank order test, we therefore reject the hypothesis of no racial bias on defendant/victim racial pairs on behalf of trial courts.

³⁰Note that, compared to other combinations, the number of cases involving white defendants and minority victims is quite small.

TABLE 3—ERROR RATES BY RACE AND REGION

Defendant's race	South: victim's race				Other regions: victim's race			
	White	Nonwhite	<i>p</i> -values	Observations	White	Nonwhite	<i>p</i> -values	Observations
<i>Panel A. Habeas corpus</i>								
White	0.350 (0.034)	0.455 (0.157)	0.741	208	0.391 (0.073)	0.500 (0.224)	0.678	52
Nonwhite	0.388 (0.039)	0.232 (0.057)	0.012	216	0.286 (0.101)	0.545 (0.157)	0.917	32
Observations	357	67			67	17		
<i>Panel B. Direct appeal</i>								
White	0.397 (0.014)	0.443 (0.056)	0.785	1,305	0.322 (0.020)	0.286 (0.077)	0.324	603
Nonwhite	0.409 (0.017)	0.376 (0.024)	0.134	1,234	0.288 (0.026)	0.304 (0.028)	0.657	575
Observations	2,048	491			863	315		

Note: Standard errors of the means in parentheses.

In panel B we show the same result for the DA sample. In the case of white defendants there is a -2 percentage points difference in error rates between white and nonwhite victim, though not statistically significant. In the case of minority defendants the difference is $+3$ percentage points (a 9 percent increase over the nonwhite/nonwhite error rate of 0.35) and is significant at the 10 percent level. Again, we reject the null of no racial bias on defendant/victim racial pairs.

The rank-order test implies that the difference in error rates across columns should go in the same direction for both rows in the previous tables (and in all those that follow). We shall see that for the case of minority defendants (second row) the first entry is always larger than the second entry almost always in a statistically significant way, while for white defendants (first row) the pattern of relative sizes of error rates typically goes in the opposite direction. Note also that the fact that the difference in errors is larger for the HC sample is consistent with the possibility that racial bias is eliminated in steps, that is, the DA courts may be less biased than the first degree courts but still biased relative to the final federal panels.

In Table 3 we find that the pattern of racial bias we uncovered is driven by Southern states.³¹ In the HC sample when we restrict the sample to sentences imposed by Southern courts we find a very large and statistically significant difference in errors for minority defendants who killed whites compared to minorities who killed nonwhites: the difference is striking at 15.5 percentage points (a 67 percent increase over the nonwhite/nonwhite error rate of 0.23), with a *p*-value of 0.01. The difference goes in the opposite direction and is not significant for white defendants. A similar pattern emerges for DA cases in the South (panel B), but with a smaller difference (3.3 percentage points for minority defendants, *p*-value 0.13). Again, the corresponding difference for white defendants has the opposite sign and is not significant.

³¹ We use the census definition of the South, thus we include Texas, Oklahoma, Arkansas, Louisiana, Mississippi, Alabama, Tennessee, Kentucky, West Virginia, Virginia, Washington D.C., Maryland, Delaware, North Carolina, South Carolina, Georgia, and Florida.

When we conduct analogous tests for other regions we fail to reject the null for both HC and DA. In HC the error rate is higher with nonwhite than with white victims both if the defendant is white and if he is not. In DA the sign pattern in the differences is reversed compared to the South, but none of these differences is statistically significant. One caveat about the results for regions other than the South in the HC sample, however, is that they cover a substantially smaller number of cases compared to those for the South.

One should note that the pattern of errors may suggest that white defendants killing minorities receive a harsh treatment in the first trial. This may suggest a bias against interracial homicides in both directions. However there are relatively few cases in the sample of whites killing minorities, which makes our estimates statistically insignificant. This paucity of data may not be casual. Perhaps such cases are not sentenced to death in the first place, hence they are not in our sample. On the contrary, while in the period of our data about 6 to 8 percent of homicides are by minority killers of white victims, these cases represent 35 percent of our sample. In other words, although we have no test or proof for this, it is possible that white on minority homicides are not treated relatively harshly if they are excluded from the start from death sentencing.

In summary it is important to be clear about what our data say and do not say. Our results can be summarized in three points. First, in the national sample both the HC and the DA data reveal that error rates for minority defendants killing whites are higher than for minority defendants killing nonwhites (second row of Table 2). The differences for white defendants are not significant (first row), but the overall pattern in inequalities, which have the opposite sign across defendant races, is indicative of racial bias according to our test. In the national DA sample some readers may find the difference across the cells of panel B to be small, as overturn rates range between 0.35 and 0.40. Notice, however, that in our sample of 3,717 cases, a 5 percent difference corresponds to roughly 186 human lives. Second, this pattern is region-specific and depends on Southern states. In particular, panel A of Table 3 shows that the error correction rate in the HC sample is very large in the South: 39 percent of the sentences on minorities killing whites are reversed, compared to 23 percent for minorities killing nonwhites (the difference is significant with p -value 0.012). Third, the error correction rate in the HC sample is larger than in the DA sample. This suggests to us that state supreme courts may retain some bias which is then eliminated by federal courts. This is consistent with the fact that in many states the selection of judges into state high courts still makes them responsive to the political preferences—and the possible bias—of their constituencies.

B. *Potential Confounding Factors*

We have suggested that our results on the rank-order tests show a racial bias on behalf of the criminal justice system up to the level of the trial court. An alternative interpretation would be that the pattern of inequalities in error rates is generated by unobserved characteristics of the crime that are systematically correlated with different combinations of defendant and victim's races. In the notation of our model, this would imply that the distribution of the evidence depends on the combination of races; i.e., $F_n^{r,R}(x)$ and $F_g^{r,R}(x)$. Although we cannot test for this possibility explicitly,

in this section we aim at providing evidence on the importance of potentially omitted factors in two ways. First, we test whether the distribution of observable crime characteristics is systematically correlated with defendant/victim pairs. Second, we perform our rank order test conditioning on a set of available characteristics that might be correlated with the severity of the crime or the quality of available evidence.

Balance Tests on Crime Characteristics.—As we discussed in Section I, the choice between a death sentence and life imprisonment often rests in the relative weight given to aggravating and mitigating circumstances associated with the defendant and/or the crime. We start by examining whether the description of aggravating and mitigating circumstances differs systematically between murders involving different race combinations. Information on a rich set of aggravating and mitigating factors put forward in the trial is available from FL for the HC dataset (not for DA) and includes the following categories:

- Aggravating: heinous and atrocious crime; pecuniary gain motive; attempt to avoid arrest and hinder law enforcement; murder during a violent felony; murder by a person under prison sentence; previous felony convictions; killed a police, fireman, guard, or other public official; multiple victims; young or old victim; great risk of death to many people; cold, calculated, premeditated.
- Mitigating: young age; no prior record; extreme emotional distress; intoxication; mental retardation and limited capacity; deprived or abused background; good prison record; lack of intent; duress.

Table 4 reports the fraction of cases for which a given aggravating or mitigating circumstance is recorded when the defendant is white (columns 1–2) or nonwhite (columns 4–5) and the victim is white (columns 1, 4) or nonwhite (columns 2, 5). For each defendant's race, the p -value associated with the difference in average prevalence of a circumstance between the two races of victims is reported in columns 3 and 6. Most important, the difference-in-differences and its associated p -value are reported in columns 7 and 8. The latter is useful to test whether, in case a certain aggravating or mitigating circumstance is more common in murders involving, say, a white victim, this difference is similar across defendant races. For the purpose of our analysis, we would like the differences—if any—to be similar across defendant races because this would imply that the cases we are considering are relatively comparable in terms of aggravating or mitigating circumstances. Columns 9 to 16 in Table 4 report similar statistics, but for the Southern subsample.

As we can see in Table 4, all of the mitigating circumstances are balanced across defendant/victim pairs (p -values in the range of 0.4–0.9 in column 8), and this is true also of most aggravating circumstances. In particular, the most common aggravating circumstances—"Heinous or atrocious crime," "Murder during a violent felony," and "Previous felony conviction"—are highly balanced (p -values of 0.98, 0.83, and 0.40, respectively in the full sample and similarly in the South subsample). The only exceptions are the following. "Murder by a person under prison sentence" is more common among whites killing nonwhites in the full sample, but is balanced in the South. "Great risk of death to many people" is more common on same-race

TABLE 4—BALANCE TEST ON AGGRAVATING AND MITIGATING CIRCUMSTANCES

	White defendant			Nonwhite defendant			Diff-in-diff	
	White victim (1)	Nonwhite victim (2)	Diff=0 (<i>p</i> -val) (3)	White victim (4)	Nonwhite victim (5)	Diff=0 (<i>p</i> -val) (6)	$\Delta\Delta$ (7)	(<i>p</i> -val) (8)
<i>Panel A. Full sample</i>								
Aggravating circumstances								
Heinous, atrocious	0.44	0.41	0.79	0.36	0.33	0.60	0.00	0.98
Pecuniary gain	0.16	0.18	0.86	0.13	0.13	0.88	−0.01	0.93
Avoid arrest, hinder law enforcement	0.15	0.06	0.29	0.12	0.13	0.70	0.11	0.26
Murder during violent felony	0.40	0.29	0.39	0.46	0.33	0.06	−0.03	0.83
Murder by person under prison sentence	0.09	0.29	0.01	0.08	0.09	0.87	−0.20	0.02
Previous felony conviction	0.23	0.29	0.58	0.22	0.18	0.47	−0.10	0.40
Killed police, fireman, or guard	0.01	0.00	0.65	0.06	0.04	0.63	0.00	0.94
Killed other public official	0.06	0.06	0.98	0.06	0.04	0.63	−0.02	0.80
Multiple victims	0.02	0.00	0.51	0.02	0.03	0.73	0.03	0.46
Young or old victim	0.00	0.00	—	0.01	0.01	0.81	0.00	0.86
Great risk of death to many people	0.13	0.00	0.11	0.05	0.10	0.12	0.19	0.03
Cold, calculated, premeditated	0.04	0.12	0.15	0.06	0.01	0.17	−0.12	0.05
Mitigating circumstances								
Age	0.04	0.00	0.42	0.06	0.06	0.89	0.04	0.49
No prior record	0.05	0.06	0.80	0.05	0.06	0.76	0.00	0.96
Extreme emotional distress	0.03	0.06	0.57	0.01	0.01	0.81	−0.02	0.62
Intoxication	0.02	0.00	0.60	0.01	0.01	0.46	0.03	0.41
Mental retardation, limited capacity	0.03	0.06	0.49	0.01	0.00	0.54	−0.04	0.35
Deprived/abused background	0.02	0.00	0.51	0.01	0.01	0.46	0.03	0.34
Good prison/jail record	0.004	0.00	0.79	0.01	0.01	0.80	0.01	0.76
Lack of intent	0.004	0.00	0.79	0.00	0.00	—	0.00	0.75
Duress	0.004	0.00	0.79	0.00	0.00	—	0.00	0.75

(Continued)

pairs, while “Cold, calculated, and premeditated” is more common among mix-race pairs. The fact that the unbalance goes in opposite directions in the latter two cases makes it difficult to assess what the net effect may be. Overall, based on the results in Table 4, there does not seem to be a systematic pattern of aggravating or mitigating circumstances being more likely associated with cross racial defendant/victim cases, which increases our confidence in the interpretation of our test as detecting bias.³²

We also collected information on other crime characteristics, related to the victim (gender and number) and the weapon used (knife, handgun, shotgun, rifle, or strangulation) for the HC and DA datasets. Furthermore, for HC cases information on the following circumstances is also available: defendant connected to the community

³²Note that since there is some arbitrariness in defining what constitutes, say, an atrocious murder, the attribution of this characteristic may itself incorporate racial bias (a point made by the Supreme Court about *Turner v. Murray*, 476 U.S. 28 (1986)). In this case, if we condition on aggravating or mitigating circumstances we purge out one of the channels of discrimination by juries.

TABLE 4—BALANCE TEST ON AGGRAVATING AND MITIGATING CIRCUMSTANCES (*Continued*)

	White defendant			Nonwhite defendant			Diff-in-diff	
	White victim (9)	Nonwhite victim (10)	Diff=0 (<i>p</i> -val) (11)	White victim (12)	Nonwhite victim (13)	Diff=0 (<i>p</i> -val) (14)	$\Delta\Delta$ (15)	(<i>p</i> -val) (16)
<i>Panel B. South</i>								
Aggravating circumstances								
Heinous, atrocious	0.44	0.55	0.48	0.38	0.34	0.58	−0.15	0.38
Pecuniary gain	0.15	0.27	0.26	0.11	0.11	0.99	−0.12	0.29
Avoid arrest, hinder law enforcement	0.15	0.09	0.61	0.11	0.14	0.55	0.09	0.46
Murder during violent felony	0.44	0.36	0.61	0.47	0.36	0.15	−0.03	0.85
Murder by person under prison sentence	0.07	0.09	0.81	0.07	0.09	0.62	0.00	0.99
Previous felony conviction	0.20	0.27	0.58	0.19	0.16	0.59	−0.10	0.46
Killed police, fireman, or guard	0.02	0.00	0.68	0.05	0.05	0.92	0.02	0.76
Killed other public official	0.07	0.09	0.75	0.07	0.04	0.37	−0.06	0.50
Multiple victims	0.01	0.00	0.74	0.01	0.04	0.11	0.04	0.29
Young or old victim	0.00	0.00	—	0.01	0.02	0.77	0.01	0.85
Great risk of death to many people	0.14	0.00	0.18	0.04	0.11	0.05	0.21	0.04
Cold, calculated, premeditated	0.04	0.18	0.02	0.06	0.02	0.19	−0.19	0.01
Mitigating circumstances								
Age	0.03	0.00	0.53	0.06	0.07	0.82	0.04	0.56
No prior record	0.04	0.09	0.43	0.05	0.07	0.55	−0.03	0.70
Extreme emotional distress	0.03	0.09	0.21	0.01	0.02	0.77	−0.06	0.23
Intoxication	0.02	0.00	0.68	0.01	0.02	0.44	0.03	0.48
Mental retardation, limited capacity	0.03	0.09	0.28	0.00	0.00	—	−0.06	0.17
Deprived/abused background	0.02	0.00	0.64	0.01	0.02	0.44	0.03	0.44
Good prison/jail record	0.01	0.00	0.81	0.01	0.02	0.77	0.01	0.76
Lack of intent	0.00	0.00	—	0.00	0.00	—	0.00	—
Duress	0.01	0.00	0.81	0.00	0.00	—	0.01	0.76

where the crime occurred; murder occurred during a burglary, theft, robbery, kidnapping, rape, or institutional killing. Online Appendix Table A5 reports a series of balance tests like the ones in Table 4 for these other variables. We find that in HC cases all the above crime characteristics are balanced. In the DA dataset, the type of weapon used is balanced, while variables related to the victim are not: same race pairs are more likely to involve multiple victims and women. While *ex ante* it is not obvious in which direction this may bias our test, in the next section we show that our results are largely robust to excluding these categories of crimes.

Robustness: Crime Characteristics.—We replicate our test for racial bias conditioning on several observable variables that characterize the crime. While this does not constitute direct evidence against the possibility that differences in unobservables are driving our results, it does shed some light on how important a similar concern may be.

In panel A of Table 5 we start from the HC sample, for which relatively detailed information on the crime was collected by FL. The left-most part of the table uses

the full sample, while the right-most part restricts the sample to Southern states. First we test whether the gender of the victim is a significant factor in our results. In the first section of panel A we restrict the attention to cases in which none of the victims were female. We find higher error on minority defendants who killed white men than on those who killed nonwhite men (the difference is 13.8 percentage points, p -value 0.06 in the full sample, and 19.3 percentage points, p -value 0.02 in the South). The corresponding difference for white defendants is -10.6 and -13.2 , not significant.

An aggravating factor that may be responsible for the results we find is the presence of multiple victims. Restricting the analysis to homicides with only one victim shows a difference of 9 percentage points (15 in the South) for nonwhite defendants who killed white versus nonwhite victims (p -values 0.09 and 0.02, respectively) and an insignificant difference on the opposite direction for white defendants.

The remainder of panel A reports results for other crime characteristics which are available only for the HC sample. A possible aggravating factor is the fact that the defendant killed a policeman, fireman, guard, or other public official. One could conjecture that crimes involving minority defendants and white victims are more represented in this category and that this generates the higher error rates we find. When we repeat the analysis considering cases in which none of the victims were one of these public officials (indicated as “no police victim” in the table), we find no significant difference in error rates for white defendants, and a difference of 13 percentage points for minority defendants, with p -value 0.03, in the full sample. In the South the difference is even larger (18 percentage points) and significant at the 1 percent level. So our results are not driven by this type of murders.

A commonly held view is that cases in which an outsider who does not know the victim commits a murder are perceived as particularly threatening and sanctioned more severely. Perhaps cases involving minority defendants and white victims fall disproportionately in this category. In the fourth section of panel A of Table 5 we examine the subset of cases where the defendant was not connected to the community where the crime occurred, according to the information recorded in FL. These cases should be relatively comparable along this dimension. Our results show that in the full sample the likelihood of error is 15 percentage points higher for minority defendants who killed white victims compared to minority defendants whose victims were not white (p -value 0.03). In the South the difference is 21 percentage points, significant at the 1 percent level. The corresponding difference in error rates for white defendants has the opposite sign and is not statistically significant. In the fifth panel we consider the subset of cases where the victims were not “high status,” as classified by FL. We find a difference of 9 percentage points (p -value 0.12) in the full sample and 16 percentage points (p -value 0.02) in the South for the combination of minority defendants and white victim, and no difference for the opposite combination.

Another way to gauge the role of potentially omitted crime characteristics is to confine our attention to murders that occurred in *similar* environmental conditions. In particular, in the sixth section of panel A of Table 5 we consider murders committed during a robbery. The likelihood of judicial error is 18 percentage points higher for minority defendants who killed at least one white victim during a robbery compared to minority defendants whose victims were all nonwhite, and is significant

TABLE 5—ERROR RATES CONDITIONAL ON CRIME CHARACTERISTICS

Defendant's race	Victim's race							
	Full sample				South			
	White	Nonwhite	<i>p</i> -values	Observations	White	Nonwhite	<i>p</i> -values	Observations
<i>Panel A. Habeas corpus</i>								
No female victim								
White	0.394 (0.047)	0.500 (0.151)	0.748	121	0.368 (0.052)	0.500 (0.224)	0.718	93
Nonwhite	0.374 (0.049)	0.235 (0.074)	0.059	133	0.378 (0.051)	0.185 (0.076)	0.018	117
Observations	208	46			177	33		
Single victim								
White	0.359 (0.031)	0.471 (0.125)	0.808	254	0.354 (0.034)	0.455 (0.157)	0.734	206
Nonwhite	0.384 (0.037)	0.292 (0.057)	0.087	242	0.390 (0.039)	0.241 (0.059)	0.017	213
Observations	414	82			354	65		
No police victim								
White	0.367 (0.032)	0.438 (0.128)	0.703	242	0.365 (0.036)	0.400 (0.163)	0.584	191
Nonwhite	0.390 (0.039)	0.262 (0.057)	0.032	220	0.397 (0.041)	0.216 (0.058)	0.005	192
Observations	385	77			322	61		
Defendant not connected to community where crime occurred								
White	0.384 (0.037)	0.500 (0.139)	0.790	191	0.362 (0.040)	0.500 (0.167)	0.789	159
Nonwhite	0.409 (0.042)	0.265 (0.064)	0.030	186	0.427 (0.045)	0.220 (0.065)	0.004	165
Observations	314	63			273	51		
No high status victim								
White	0.398 (0.035)	0.385 (0.140)	0.463	209	0.386 (0.039)	0.375 (0.183)	0.476	166
Nonwhite	0.398 (0.043)	0.309 (0.063)	0.121	183	0.404 (0.046)	0.244 (0.065)	0.023	159
Observations	324	68			272	53		
Robbery								
White	0.342 (0.055)	0.600 (0.245)	0.848	81	0.292 (0.057)	0.600 (0.245)	0.889	70
Nonwhite	0.436 (0.057)	0.250 (0.090)	0.040	102	0.431 (0.059)	0.105 (0.072)	0.000	91
Observations	154	29			137	24		
Felony								
White	0.344 (0.043)	0.429 (0.202)	0.658	129	0.305 (0.045)	0.429 (0.202)	0.725	112
Nonwhite	0.391 (0.046)	0.257 (0.075)	0.063	150	0.394 (0.048)	0.167 (0.069)	0.003	134
Observations	237	42			209	37		
<i>Panel B. Direct appeal</i>								
No female victim								
White	0.358 (0.017)	0.403 (0.058)	0.768	831	0.393 (0.021)	0.463 (0.068)	0.836	576
Nonwhite	0.413 (0.020)	0.333 (0.026)	0.008	926	0.444 (0.024)	0.360 (0.034)	0.022	638
Observations	1,343	414			963	251		
Single victim								
White	0.389 (0.013)	0.374 (0.051)	0.389	1,419	0.414 (0.016)	0.418 (0.061)	0.526	1,012
Nonwhite	0.388 (0.016)	0.361 (0.022)	0.163	1,382	0.412 (0.019)	0.391 (0.028)	0.264	1,001
Observations	2,220	581			1,639	374		

Note: Standard errors of the means in parentheses.

at the 5 percent level. Results are even stronger in the South, with a difference of 32 percentage points, significant at the 1 percent level. The difference for white defendants is in the opposite direction and not statistically significant. Finally, when we restrict the sample to cases that are similar in the sense of being classified as “felony murders,” we find again a higher error rate for nonwhite defendants who killed white victims (13 percentage points in the full sample, 22 in the South, with p -values 0.07 and 0.004, respectively), and no corresponding difference for white defendants.

We have less information on crime characteristics for the DA compared to the HC sample. In panel B of Table 5 we begin by testing whether the gender of the victim is a significant factor in our results. In the first panel we restrict the attention to cases in which none of the victims was female. We find 8 percentage points higher error on minority defendants who killed white men than on those who killed nonwhite men (p -values 0.01 and 0.02 in the full sample and in the South, respectively). The corresponding difference for white defendants is in the opposite direction and not significant. In the second panel we restrict the analysis to homicides with only one victim we find a difference of 3 percentage points for nonwhite defendants who killed white versus nonwhite victims, both in the full sample and in the South, but the p -values increase to 0.16 and 0.26, respectively.

Robustness: Legal Assistance.—We now analyze whether differences in error rates are due to unequal quality of legal assistance of the defendant. A possible interpretation of our finding is that minority defendants who killed a white victim receive systematically worse legal assistance compared to minority defendants who killed a minority victim. This would actually be another source of racial bias, which we discussed in Section IIC and which would distort the distribution of the signal for given characteristics of the crime. Note that if a minority defendant received a worse defense regardless of the race of the victim, this would not invalidate our test nor change the interpretation of the results.

In Table 6 we repeat our tests restricting the sample to cases that are relatively similar in terms of some trial characteristics. We can only do this for HC cases because no trial characteristic is available in the DA dataset. As a proxy for the quality of legal assistance at the trial stage we use the fact that “ineffective assistance of counsel” in the guilt and sentencing phase was included among the claims for relief. We start by restricting the sample to 219 HC cases in which ineffective assistance of counsel was not raised among the claims in the appeal. In this subset of cases the difference in error rates for minority defendants who killed a white versus a nonwhite victim is 18 percentage points (p -value 0.04) in the full sample, and 28 percentage points (p -value 0.004) in the South. Comparing these results to those in Tables 2 and 3 suggests that variation in the quality of legal assistance across racial combinations of defendants and victims may actually lead us to *underestimate* the extent of bias.

In the remaining parts of Table 6, we consider the subset of cases in which “prosecutor’s suppression or withholding of evidence or other prosecutorial misconduct” was not raised among the claims, nor was “improper interrogation.” That is, there was no involuntary confession or guilty plea or request for attorney denied. In both subsamples the order of magnitude of the differences in error rates and the significance level remain comparable to those of Tables 2 and 3, and the

TABLE 6—ERROR RATES CONDITIONAL ON TRIAL CHARACTERISTICS, HABEAS CORPUS

	Victim's race							
	Full sample				South			
	White	Nonwhite	<i>p</i> -values	Observations	White	Nonwhite	<i>p</i> -values	Observations
<i>Panel A. Ineffective assistance of counsel not in any claim</i>								
White	0.458 (0.048)	0.571 (0.202)	0.708	114	0.471 (0.054)	0.500 (0.289)	0.539	91
Nonwhite	0.410 (0.054)	0.227 (0.091)	0.043	105	0.437 (0.059)	0.158 (0.086)	0.004	90
Observations	190	29			158	23		
<i>Panel B. Prosecutorial suppression/withholding of evidence not in any claim</i>								
White	0.383 (0.038)	0.600 (0.163)	0.902	177	0.366 (0.042)	0.667 (0.211)	0.919	137
Nonwhite	0.385 (0.043)	0.300 (0.073)	0.160	170	0.388 (0.045)	0.235 (0.074)	0.039	150
Observations	297	50			247	40		
<i>Panel C. Improper interrogation not in any claim</i>								
White	0.362 (0.033)	0.500 (0.129)	0.849	234	0.348 (0.036)	0.500 (0.167)	0.813	188
Nonwhite	0.385 (0.038)	0.283 (0.059)	0.073	221	0.394 (0.041)	0.220 (0.059)	0.008	192
Observations	379	76			320	60		

Note: Standard errors of the means in parentheses.

rank order test rejects the null of absence of racial bias according to our model in all cases except for the second panel in the full sample, where the *p*-value for nonwhite defendants increases to 0.16.

Note that although the above variables seem reasonably good proxies for the quality of legal assistance, some of them reflect discretionary choices on behalf of the defense in the appeal process (e.g., which claims to present) and in this sense they may not be fully objective. Nonetheless, we take the evidence in Table 6 as suggestive that differences in the quality of legal assistance are not entirely responsible for our results.

Robustness: Reverse Discrimination of Appeal Courts.—So far we have assumed that the appeal courts are unbiased, with or without full information, thus they make no errors or they make unbiased errors due to imperfect information. If the appeal court is biased in the same direction of the trial court, our test will underestimate the extent of racial bias because the (biased) appeal court will reverse the trial court decision less often than an unbiased court would do. The challenge for us would arise from a bias in the opposite direction, namely if the appeal court were inclined to give relief more often than an unbiased court would do. Note that a simple bias of the appeal courts in favor of black defendants (for example on the ground that they are on average poorer and may not be able to afford good legal assistance) would not invalidate our tests of racial bias. What would be problematic for us is a situation where the bias is linked to a particular combination of defendant/victim race, e.g., if the appeal court rules systematically more in favor of nonwhite defendants who killed white victims. Although we cannot rule this out a priori, we test

the plausibility of this scenario by exploiting information on the political orientation of appeal judges. We conjecture that, if a bias in favor of minorities who killed white victims existed, this would more likely be found among liberal judges than among conservative ones, characteristics which we assume to be correlated with party affiliation. Thus we repeat our analysis conditioning on party affiliation of the appeal judges.

Let's begin with the HC sample. For each sentence, we collected the names of the judges who served on the appeal court that decided on that sentence, and recovered information on these judges from the Biographical Directory of Federal Judges available from the Federal Judicial center. This directory contains biographical information on all judges that served on US District Courts, the US Courts of Appeals, the US Supreme Court, and the US Circuit Courts since 1789. We recorded the year in which each judge was appointed to the relevant court and classified the political orientation of the judge as "Republican" if he or she was appointed under a Republican president and "Democratic" if he or she was appointed under a Democratic president. If our results were driven by reverse discrimination on behalf of appeal judges, we should not find discrimination (or find it to a lesser extent) when we look at courts that are predominantly composed of Republican judges.

Panel A of Table 7 reports the results for the subset of HC cases where the majority (first panel) or the totality (second panel) of the judges were appointed under a Republican president. The left-most part of the table employs the full sample, while the right most part restricts the analysis to the South. Both sets of results are consistent with our earlier findings, and indicate a higher likelihood of relief for non-white defendants who killed white victims. When we consider appeal courts where a majority of the judges are Republican (first panel), the magnitude of the difference in error rates is 7 percentage points in the full sample and 12 percentage points in the South (p -values 0.20 and 0.08, respectively). This differences increase to 13 and 16 percentage points when we restrict our test to courts that are entirely composed of Republican-appointed judges (second panel, p -values 0.09 and 0.06).

In the third section of panel A we consider the possibility that political climate may affect relief rates, and restrict the sample to death sentences that occurred under a Republican administration. We find a difference of 15 percentage points in the full sample and 23 percentage points in the South for nonwhite defendants who killed white victims, significant at the 5 and 1 percent level, respectively. In all three panels the corresponding difference for white defendants is in the opposite direction and not significant.³³

In panel B of Table 7 we conduct a similar exercise for the DA dataset. In this case we have available both the party affiliation of the direct appeal judges and the measure of judges' ideology proposed by Brace, Langer, and Hall (2000), which they label PAIID.³⁴ The first panel of panel B shows that when we restrict the sample to first stage appeals decided by courts in which at least 50 percent of the judges were Republican, error rates on nonwhite defendants who killed white victims are higher

³³ Results are similar if we consider habeas corpus appeal sentences that occurred under a Republican administration.

³⁴ Essentially PAIID measures judges' ideology on a scale from conservative to liberal based upon party affiliation modified by a set of criteria allowing for differences across states. We match this measure to reflect the composition of the state appeal court the year in which the appeal sentence was issued.

TABLE 7—POSSIBLE BIAS OF APPEAL COURTS

Defendant's race	Victim's race							
	Full sample				South			
	White	Nonwhite	<i>p</i> -values	Observations	White	Nonwhite	<i>p</i> -values	Observations
<i>Panel A. Habeas corpus</i>								
Majority of final federal panel Republican								
White	0.355 (0.041)	0.455 (0.157)	0.730	149	0.361 (0.046)	0.500 (0.224)	0.728	114
Nonwhite	0.284 (0.043)	0.216 (0.069)	0.200	146	0.313 (0.048)	0.194 (0.072)	0.084	127
Observations	247	48			204	37		
All judges appointed under Republican								
White	0.250 (0.073)	0.500 (0.289)	0.799	40	0.212 (0.072)	0.500 (0.500)	0.716	35
Nonwhite	0.216 (0.058)	0.083 (0.083)	0.096	63	0.244 (0.065)	0.083 (0.083)	0.063	57
Observations	87	16			78	14		
Sentence under Republican administration								
White	0.343 (0.037)	0.438 (0.128)	0.760	182	0.328 (0.040)	0.400 (0.163)	0.665	147
Nonwhite	0.402 (0.043)	0.250 (0.063)	0.024	180	0.412 (0.045)	0.179 (0.062)	0.001	158
Observations	298	64			256	49		
<i>Panel B. Direct appeal</i>								
Majority of state supreme court Republican								
White	0.285 (0.023)	0.280 (0.092)	0.477	428	0.365 (0.061)	0.333 (0.333)	0.463	66
Nonwhite	0.292 (0.031)	0.217 (0.034)	0.051	368	0.419 (0.090)	0.176 (0.095)	0.032	48
Observations	619	177			94	20		
Median ideology of state supreme court conservative								
White	0.364 (0.015)	0.406 (0.062)	0.745	1,151	0.377 (0.016)	0.421 (0.066)	0.740	974
Nonwhite	0.380 (0.018)	0.354 (0.025)	0.199	1,050	0.386 (0.019)	0.367 (0.027)	0.286	957
Observations	1,781	420			1,555	376		
Chief justice of state supreme court conservative								
White	0.363 (0.015)	0.426 (0.060)	0.845	1,155	0.380 (0.016)	0.443 (0.064)	0.829	993
Nonwhite	0.385 (0.018)	0.352 (0.025)	0.142	1,072	0.391 (0.019)	0.366 (0.026)	0.221	988
Observations	1,790	437			1,584	397		

Note: Standard errors of the means in parentheses.

than on those who killed nonwhite victims (differences of 7.5 and 24 percentage points in the full sample and in the South, respectively, with *p*-values 0.05 and 0.03). For white defendants, error rates are virtually the same across victim races.

In the remaining panels we rely on the continuous measure of ideology proposed by Brace, Langer, and Hall (2000) and define as “conservative” judges whose ideology score falls in the top 50 percent of the distribution of PAJID. The second panel restricts the sample to courts whose median member (in terms of ideology) is

“conservative,” while the third does the same but with reference to the chief justice. In both cases we find that the direction and the magnitude of the differences in error rates are comparable to our main results in Tables 2 and 3, though we lose statistical significance. Furthermore, this result does not depend on the particular cutoff for the definition of conservative. Figure 6 shows that the positive difference in error rates for minority defendants who killed white versus nonwhite victims holds for each and every quartile of the distribution of PAJID, indicating that our main result is not driven by the ideological orientation of the court. The corresponding differences for white defendants are instead sometimes positive, sometimes negative, and vary by quartile.

To sum up, we find no evidence that left-leaning judges are those who correct more mistakes in pairs involving minority defendants and white victims. In fact, we find that our results hold strong when we restrict the sample to relatively conservative appeal courts. Thus, we find no obvious evidence of reverse discrimination by higher courts.

V. Conclusions

This paper proposes a test for racial bias in capital sentencing in the United States over the period 1973–1995. We use the share of judicial errors in first degree sentencing as an indicator of racial bias of such courts. Using an originally collected dataset, we uncover a bias against minority defendants killing white victims. The bias is present, according to our test, only in Southern states. More precisely, according to our interpretation first degree courts tend to place less weight on the possibility of condemning an innocent in cases of minority defendants with one or more white victims relative to minority defendants who did not kill whites.³⁵ This result is not explained by differences in observable characteristics of the crime or of the trial, nor by the ideological orientation of appeal courts.

APPENDIX: PROOF THAT THE ERROR RATE IS DECREASING IN x^*

To simplify the notation, here we omit subscripts and superscripts related to race. That is, we write x instead of x_{rR} , π instead of π^r , and f_g, f_n instead of f_g^r, f_n^r . The error rate (4) is

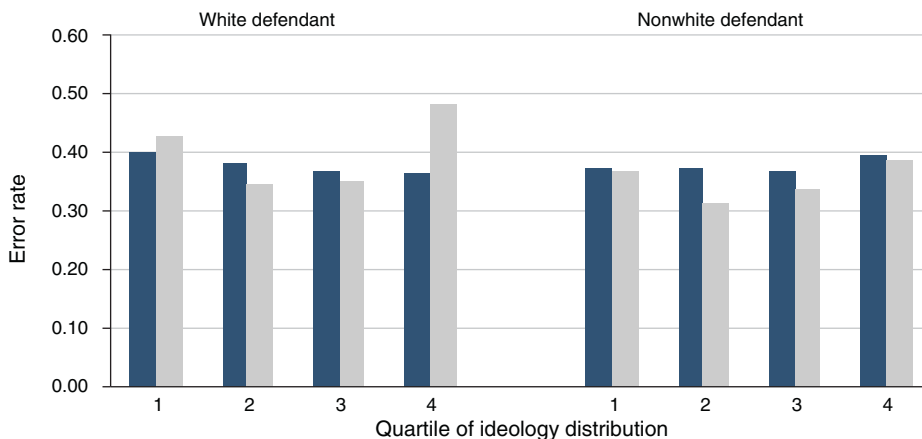
$$(A1) \quad E(r, R) = 1 - \frac{1}{1 + \frac{1-\pi}{\pi} \frac{1-F_n(x^*)}{1-F_g(x^*)}}.$$

Expression (A1) is decreasing in x^* if and only if $\frac{1-\pi}{\pi} \frac{1-F_n(x^*)}{1-F_g(x^*)}$ is decreasing in x^* . Taking the first derivative of this product with respect to x^* , its sign is the same as the sign of:

$$(A2) \quad \frac{1}{[1 - F_g(x^*)]^2} \left[\int_{x^*}^1 f_g(x^*) f_n(x) dx - \int_{x^*}^1 f_g(x) f_n(x^*) dx \right].$$

³⁵ A difference in the opposite direction (bias against whites killing minorities compared to whites killing whites) also emerges, but is never statistically significant.

Panel A. Conservativeness of median member of state supreme court



Panel B. Conservativeness of chief justice of state supreme court

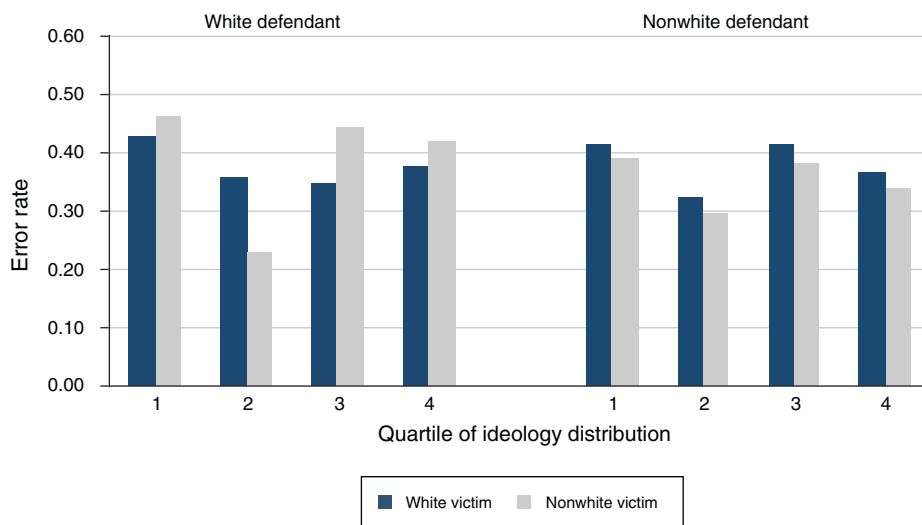


FIGURE 6. ERROR RATE AND IDEOLOGY OF APPEAL COURTS

To see that the term in square brackets is negative, recall that from MLRP we know that $\frac{f_g(x)}{f_n(x)} > \frac{f_g(x^*)}{f_n(x^*)}$ for any $x > x^*$. Because the integrals in (A2) are calculated for $x \in (x^*, 1]$, then in this range $f_g(x^*)f_n(x) < f_g(x)f_n(x^*)$, hence (A2) is negative.

REFERENCES

- Abrams, David S., Marianne Bertrand, and Sendhil Mullainathan. 2012. "Do Judges Vary in Their Treatment of Race?" *Journal of Legal Studies* 41 (2): 347–84.
- Alesina, Alberto, and Eliana La Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing: Dataset." *American Economic Review*. <http://dx.doi.org/10.1257/aer.104.11.3397>.
- Antonovics, Kate, and Brian G. Knight. 2009. "A New Look at Racial Profiling: Evidence from the Boston Police Department." *Review of Economics and Statistics* 91 (1): 163–77.

- Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson. 2012. "The Impact of Jury Race in Criminal Trials." *Quarterly Journal of Economics* 127 (2): 1017–55.
- Anwar, Shamena, and Hanming Fang. 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 96 (1): 127–51.
- Argys, Laura M., and H. Naci Mocan. 2004. "Who Shall Live and Who Shall Die? An Analysis of Prisoners on Death Row in the United States." *Journal of Legal Studies* 33 (2): 255–82.
- Baldus, David C., Charles Pulanski, and George Woodworth. 1983. "Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience." *Journal of Criminal Law and Criminology* 74 (3): 661–753.
- Barnes, Katherine, David Sloss, and Stephen Thaman. 2009. "Place Matters (Most): An Empirical Study of Prosecutorial Decision Making in Death-Eligible Cases." *Arizona Law Review* 51 (2): 305–79.
- Becker, Gary S. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Blume, John, Theodore Eisenberg, and Martin T. Wells. 2004. "Explaining Death Row's Population and Racial Composition." *Journal of Empirical Legal Studies* 1 (1): 165–207.
- Brace, Paul, Laura Langer, and Melinda Gann Hall. 2000. "Measuring the Preferences of Supreme Court Judges." *Journal of Politics* 62 (2): 387–413.
- Coyne, Randall, and Lyn Entzeroth. 2006. *Capital Punishment and the Judicial Process*. Durham: Carolina Academic Press.
- Donohue, John J., III, and Justin Wolfers. 2005. "Uses and Abuses of Empirical Evidence in the Death Penalty Debate." *Stanford Law Review* 58 (3): 791–846.
- Ehrlich, Isaac. 1975. "The Deterrent Effect of Capital Punishment: A Question of Life and Death." *American Economic Review* 65 (3): 397–417.
- Fagan, Jeffrey, and James Liebman. 2002. "Processing and Outcome of Death Penalty Appeals After Furman v. Georgia, 1973–95: [United States] [Computer file] [ICPSR version]." New York: Columbia School of Law and Mailman School of Public Health [producers] (2002). Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor] (2002). doi:10.3886/ICPSR03468.v1.
- Gennaioli, Nicola, and Andrei Shleifer. 2007. "The Evolution of Common Law." *Journal of Political Economy* 115 (1): 43–68.
- Gross, Samuel R., and Robert Mauro. 1984. "Patterns of Death: An Analysis of Racial Disparities in Capital Sentencing and Homicide Victimization." *Stanford Law Review* 37 (1): 27–153.
- Grossman, Gene M., and Michael L. Katz. 1983. "Plea Bargaining and Social Welfare." *American Economic Review* 73 (4): 749–57.
- Humphrey, John A., and Timothy J. Fogarty. 1987. "Race and Plea Bargained Outcomes: A Research Note." *Social Forces* 66 (1): 176–82.
- Iyengar, Radha. 2007. "An Analysis of Attorney Performance in the Federal Indigent Defense System." National Bureau of Economic Research Working Paper 13187.
- Katz, Lawrence, Steven D. Levitt, and Ellen Shustorovich. 2003. "Prison Conditions, Capital Punishment, and Deterrence." *American Law and Economics Review* 5 (2): 318–43.
- Knowles, John, Nicola Persico, and Petra Todd. 2000. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109 (1): 203–29.
- Liebman, James S., Jeffrey Fagan, and Valerie West. 2000. "A Broken System: Error Rates in Capital Cases, 1973–1995." Columbia Law School, Public Law Research Paper 15.
- Radelet, Michael R., and Glenn R. Pierce. 1985. "Race and Prosecutorial Discretion in Homicide Cases." *Law and Society Review* 19 (4): 587–622.
- Reinganum, Jennifer F. 1988. "Plea Bargaining and Prosecutorial Discretion." *American Economic Review* 78 (4): 713–28.
- US Department of Justice, Bureau of Justice Statistics. 1996. "Murder Cases in 33 Large Urban Counties in the United States, 1988: [ICPSR version]." Washington, DC: US Department of Justice, Bureau of Justice Statistics [producer] (1988). Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor] (1996). doi:10.3886/ICPSR09907.v1