NBER WORKING PAPER SERIES

A TEST OF RACIAL BIAS IN CAPITAL SENTENCING

Alberto F. Alesina
Eliana La Ferrara

Working Paper 16981
http://www.nber.org/papers/w16981

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2011

We thank Shamena Anwar, Abhijit Banerjee, Katherine Barnes, Francesco Corielli, Richard Dieter, John Donohue, Jeffry Fagan, James MacKinnon, Nicola Persico, Andrei Shleifer and seminar participants at CIAR, Duke, Warwick, Royal Holloway, IIES Stockholm University, Queen's University, 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, Mariagiovanna Di Feo, Damiano Briguglio, Giovanni Rizzo, Matteo Fiorini and Marta Barazzetta provided excellent research assistance. This paper was written while Alesina was visiting IGIER 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

A Test of Racial Bias in Capital Sentencing
Alberto F. Alesina and Eliana La Ferrara
NBER Working Paper No. 16981
April 2011
JEL No. K42

ABSTRACT

This paper proposes a test of racial bias in capital sentencing based upon patterns of judicial errors in lower courts. We model the behavior of the trial court as minimizing a weighted sum of the probability of sentencing an innocent and that of letting a guilty defendant free. We define racial bias as a situation where the relative weight on the two types of errors is a function of defendant and/or victim race. The key prediction of the model is that if the court is unbiased, ex post the error rate should be independent of the combination of defendant and victim race. We test this prediction using an original dataset that contains the race of the defendant and of the victim(s) for all capital appeals that became final between 1973 and 1995. We find robust evidence of bias against minority defendants who killed white victims: In Direct Appeal and Habeas Corpus the probability of error in these cases is 3 and 9 percentage points higher, respectively, than for minority defendants who killed minority victims.

Alberto F. Alesina
Department of Economics
Harvard University
Littauer Center 210
Cambridge, MA 02138
and IGIER
and also NBER
aalesina@harvard.edu

Eliana La Ferrara
Universita' Bocconi
Dept of Economics
via Roentgen 1
20136 Milano
Italy
eliana.laferrara@unibocconi.it

1 Introduction

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 per cent of the national population, they account for more than 40 per cent 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 that may be correlated with race, is crucial if one wants to rigorously test 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 observable and unobservable dimensions.

We develop a model where courts minimize the probability of making judicial errors and we derive a simple test for racial bias. Our test builds on the following insight. Even if we do not observe in the data all the elements that trial courts consider when imposing a death verdict, if these courts are unbiased, ex post they should not end up making 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 US, namely that all first degree capital sentences are automatically appealed, 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. On the other hand, if we assume 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] Our 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 non-white victims, they should also commit more errors on *white* defendants who killed white victims than on those who killed non-white 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.

In order to implement this test we embarked on a challenging data collection exercise. We started from the data on capital appeals assembled by Liebman, Fagan and West (2000) for

[1] Amnesty International, *USA Death by Discrimination - The Continuiung Role of Race in Capital Cases*, April 2003, p.1.

[2] Remember that the death penalty applies almost exlusively to homicides so there is always at least one well identified victim.

the period 1973-1995 and we supplemented it with information 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 et al.'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 US 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. This result holds strong both for the cases that reach the final stage of revisions and appeal, the Habeas Corpus stage in Federal Courts, and for the full sample of cases in the first appeal stage, called Direct Appeal. For Habeas Corpus cases involving a minority defendant, the error rate was 37.5 percent if the victim was white, and 28.4 percent if it was not white, 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 non-white, 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 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 by Southern States. The difference in error rates in these States is large: in Habeas Corpus cases, the error rate is 15.5 percentage points higher for minority defendants with white victims as compared to minority defendants with non-white victims (p-value .01). For the Direct Appeal sample the corresponding difference is 3.3 percentage points (p-value .13).

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 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: higher reversal for minority defendants who killed white victims compared to those who killed non-

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

white victims, but not so for white defendants who killed white victims compared to white defendants who killed non-white victims.

Another assumption upon which our test rests is that possible unobservable characteristics, 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 investigate whether our results are driven by *observable* characteristics of the crime or of the trial which differ between defendant-victim pairs, such as the type of crime (e.g., robberies versus other crimes), the characteristics of the victim (gender, number, status in the community) and the quality of legal counsel. We do not find any evidence of this.

Our test is related to those developed in the literature on racial bias in motor vehicle searches and in particular to 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. Abrams, Bertrand and Mullainathan (forthcoming) empirically test whether there are systematic differences across judges in the racial gap in sentencing for felony crimes and exploit the random assignment of cases to judges in one county in Illinois. 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.

Our paper is also related to the literature on the death sentence and its usefulness. We do not touch upon the question of the deterrence effect of the death penalty and 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 potentially 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

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

[5]On this point see among others, Erlich (1975), Katz, Levitt and Shustorovich (2003) and Donohue III and Wolfers (2005) for a review.

sentence being higher when the victim was white), but the race of the defendant had no residual explanatory power. 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 2 offers a brief synthesis of the institutional details useful to understand judicial errors in capital cases in the US.[6] Section 3 describes our model of behavior of Courts and defendants and derives our test of racial bias. Section 4 describes the data. Section 5 presents our empirical methodology and results. The last section concludes.

2 History and institutional background

2.1 A brief history of the death penalty in the US

We can identify four periods of the history of capital punishment in the US: from the arrival of the Pilgrim to the Civil War, from the Civil War to the second world war, from 1945 to 1972 and from 1976 to today. In the pre Civil Wars period there were significant differences in the application of the death penalty in the Northern Colonies (and then states) and Southern colonies (and then states). In the North the death penalty was applied to murder, treason a few other violent crimes and several religious transgressions like witchcraft, blasphemy, idolatry and acts of sodomy. In the South in addition to those crimes the death penalty served the purpose of enforcing the slave system and therefore was prescribed for slave stealing and (especially) the organization or instigation of slaves against owners. In addition a Black Code enlisted crimes punishable with death for black defendants but not whites. The most heinous of those was the rape of a white woman by a black man. After the civil war the Black Code was abolished but in the South the application of the death penalty implicitly continued to follow the guidelines of this code. This is quite suggestive for the evidence we present below.

After the Second World war there was a steady decline in executions. The last execution conducted before the Supreme Court declared the death penalty unconstitutional was in Colorado in 1967. In Furman v. Georgia in 1972, the Supreme Court reversed all existing death sentences and capital statutes as unconstitutional. The ruling of the Court held that the administration of capital punishment at the time violated the eighth amendment prohibition of cruel and unusual punishment. All the Justices in the majority motivated their decision relying on the arbitrary and/or discriminatory nature of the death penalty as implemented by existing statutes. Starting the following year, new death penalty statutes were written, and in Gregg v. Georgia (1976), the Court allowed capital punishment to

[6] Section 2 largely draws on Coyne and Enzeroth (2006).

resume.[7] The statutes that were deemed constitutional typically contained explicit lists of aggravating and mitigating factors that should guide the decisions of the juries and reduce the extent of discretion. The death penalty is currently allowed only for treason or murder implying therefore a much narrower definition of capital crimes relative to previous historical periods.

2.2 Procedures

Today, thirty-five states in the US allow capital punishment.[8] 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* in 1976. That statute prescribed: a) an independent trial of guilt or innocence; b) a second hearing solely to determine the sentence; c) a finding of at least one aggravating circumstance; d) an automatic review by the Georgia Supreme Court and e) 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.[9] 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 one per cent of the murders committed in a year ends up in a death sentence.[10]

The appeal process

The most important aspect of the capital punishment procedural rules for the purpose of our study is that all capital sentences, with no exceptions, are automatically appealed in state

[7] In reality since 1972 Utah and Florida had already reinstated death penalty statutes and many states did not stop sentencing defendants. By 1976, 460 individuals had been sentenced to death across different states.

[8] The Federal Government has two death penalty statues on for the military and the other for non military crimes.

[9] Woodson v. North Carolina (1976) and Roberts v. Louisiana (1976).

[10] See Barnes, Sloss and Thaman (2008) for a recent discussion of criteria according to which prosecutors purse the death penalty in about 4 percent of capital crimes in Missouri.

high courts. In all but two states the appeal runs directly from the trial court to the state supreme court, while in Alabama and Ohio it goes through an intermediate court of criminal appeals before reaching the highest court. Sentences that 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 et al., 2000). When all appeals are exhausted, the only hope left for the defendant is an act of clemency from the State Governor.

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 overturned cases 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.

3 The model

3.1 The setup

We consider a defendant whom a court can condemn to the death penalty or to a lesser penalty. 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. Our assumption (to be discussed below) is that while the lower courts can make mistakes, higher courts make no mistakes. Our empirical test holds identically under the more general assumption that even the highest courts can make mistakes but these are uncorrelated with the race of the defendant and of the victim.

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 $R \in \{W, M\}$. The court observes several signals over the characteristics of the defendant and of the crime and summarizes them in a single dimension which we denote with x and that we can think of as the evidence. We normalize the support of x so that $x \epsilon [0, 1]$. 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.[11] While the quality of defense is not explicitly modeled in our framework,

[11] Results 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

it can be incorporated in a different distribution of evidence faced by the courts for minority and white defendants. Thus we allow $F_g^r(x)$ and $F_n^r(x)$ to depend on the race of the defendant. 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:

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

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

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 than an individual or race r killing an individual of race R is sentenced to death is: $p_r(x) = \Pr(x \geqslant x_{rR})$.

3.1.1 The problem of the potential criminal

An individual considering whether to commit a capital crime (crime in short) compares the costs and benefits of it. The benefit of committing the crime and getting away with it is $b > 0$; think of it as the money stolen from a bank (with a killing during the robbery) or the pleasure of killing an enemy. The cost of being sentenced to death having committed the crime is $c_g > 0$. The cost of being sentenced to death not having committed the crime is c_n, with $0 < c_n < c_g$.[12] All of the above b, c_g and c_n are public information.

For an individual the cost of committing a crime, which may include the moral cost, is v and it is drawn from a distribution $\Im^r(v)$ with support in $\mathbb{R}^+$. The court knows the distribution but not the realization of v which is known only to the individual. We allow the distribution of costs to differ across races, thus allowing a higher propensity of minorities to commit crimes.

The individual chooses whether to commit a crime taking into account the likelihood of being convicted and takes x_{rR} as given since it is chosen by the court.

In certain types of crimes the defendant cannot choose the victim and therefore his or her race. One example is a bank robbery with the killing of guards, whose race was unknown to the criminals ex ante. In a second type of crime the defendant wants to kill, say, a relative, in which case he also cannot choose the race of the victim but the race of the victim is known ex ante. In a third type of crime the defendant can choose the race of the victim, say in a rape with murder. For expositional simplicity we present the second case in the text and the first and third in the Appendix.

The expected payoff from the crime for an individual with race r and a certain realization v is given by:

$$\left[1 - F_g^r(x_{rR})\right][-v - c_g] + F_g^r(x_{rR})[b - v]$$

The first term represents the cost of being convicted, the second term the benefit of getting away with the crime. The expected payoff from not committing that crime is:

disproportionately represented by the latter.

[12] Remember that by assumption there are no mistakes in the final ruling of higher courts, therefore no innocent individual is executed. Thus the cost c_n represents the costs of being on death row until the first sentence is reversed.

$$-\left[1-F_n^r\left(x_{rR}\right)\right]c_n.$$

Comparing costs and benefits, an individual commits a crime if and only if:

$$v \leqslant F_g^r\left(x_{rR}\right)b-\left[1-F_g^r\left(x_{rR}\right)\right]c_g+\left[1-F_n^r\left(x_{rR}\right)\right]c_n \equiv v^*(x_{rR}) \tag{1}$$

Thus $v^*\left(x_{rR}\right)$ is the threshold of individual cost v below which an individual of race r chooses to commit a crime against a victim of race R. Define:

$$Prob\left(v \leqslant v^*(x_{rR})\right)=\Im^r\left(v^*(x_{rR})\right) \equiv \pi^r(x_{rR}) \tag{2}$$

as the probability of guilt, i.e. the probability that the realization of v is low enough so that a crime is committed. Note that if the court applied a different standard of proof depending on the race of the victim, e.g. $x_{m,W}<x_{m,M}$, then potential criminals would internalize that and ceteris paribus we would observe fewer crimes involving m,W pairs than m,M pairs.

3.1.2 The problem of the court

The court wants to sentence guilty defendants to the death penalty, but wants to avoid the mistakes of sentencing innocent defendants. Note 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 (c_n) to innocent defendants and the costs of reputation losses of having made mistakes.[13] These considerations can be summarized by assuming that the court minimizes a weighted average of the probability of condemning an 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:

$$\min_{x_{rR}}\left\{\alpha_{rR}\left[1-\pi^r(x_{rR})\right]\left[1-F_n^r\left(x_{rR}\right)\right]+\left(1-\alpha_{rR}\right)\left[\pi^r(x_{rR})\right]\left[F_g^r\left(x_{rR}\right)\right]\right\} \tag{3}$$

$$0<\alpha_{rR}<1$$

for $r\in\{w,m\}$ and $R\in\{W,M\}$. The first and second term in (3) are, respectively, the type I and type II error. 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 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:

$$\frac{f_g^r\left(x_{rR}^*\right)}{f_n^r\left(x_{rR}^*\right)}=\frac{\alpha_{rR}}{1-\alpha_{rR}}\frac{1-\pi^r(x_{rR}^*)}{\pi^r(x_{rR}^*)} \tag{4}$$

The cutoff value x_{rR}^* is thus the "standard of proof" applied by the court. Inspection of (4) 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.

[13] In the more general version of the model in which even superior courts can actually make mistakes, although uncorrelated with race, lower courts also have to worry about the possibility that innocent people sentenced to death are never released, which implies a very high moral cost.

The equilibrium of the model is given by (4) together with:

$$\pi^r\left(x^*_{rR}\right) = \Im^r(v^*(x^*_{rR})) \tag{5}$$

By Brouwer's fixed point theorem an equilibrium exists. The proof of uniqueness is in the Appendix. Note that even if $\alpha_{rR} = \alpha$ for any r, 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 (4) depends on the race of the defendant, r, because f^r_g and f^r_n depend on r; and the right hand side of (4) depend on r because the distribution of the cost of committing a crime $\Im^r$ (and hence π^r) depends on r. For example, if minorities have a lower cost 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 (4) is through the parameter α_{rR}. This is the key insight upon which our test is based.

3.2 A test of racial bias

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

Definition 1 *Bias only on the race of the defendant:* $\alpha_{rR} = \alpha(r)$

In this case 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 2 *Bias only on the race of the victim:* $\alpha_{rR} = \alpha(R)$

In this case 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 3 *Bias on both the race of the victim (R) and the race of the defendant (r):* $\alpha_{rR} = \alpha(r, R)$

This is a situation where 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 lenient on both if the victim is non-white, 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:

$$\gamma(r, R) = \pi^r\left(x^*_{rR}\right)\left[1 - F^r_g\left(x^*_{rR}\right)\right] + \left[1 - \pi^r\left(x^*_{rR}\right)\right]\left[1 - F^r_n\left(x^*_{rR}\right)\right]. \tag{6}$$

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

$$E(r, R) = 1 - \frac{\pi^r\left(x^*_{rR}\right)\left[1 - F^r_g\left(x^*_{rR}\right)\right]}{\gamma(r, R)} \tag{7}$$

with $\gamma(r, R)$ given by (6). 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 (7) through the optimal threshold x^*_{rR} derived from (4), allowing us to build a test. The race of the victim R only affects x^*_{rR} if α_{rR} depends on R. In fact in the equilibrium condition (4) the left hand side does not depend on R, nor does the second term on the right hand side, so $\alpha_{rR}/(1-\alpha_{rR})$ is the only term through which R could affect the equilibrium threshold x^*_{rR} and consequently the error rate. On the other hand, the race of the defendant r can affect x^*_{rR} either because of differences in α or due to differences in the signal distributions f^r or due to differences in the cost of committing crimes that translates in different proportions of guilty $\pi^r(\cdot)$ in expression (4). 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, this is evidence of racial bias in favor of white victims. In 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.

[Figures 1 and 2 about here]

Figure 1 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^r_n and f^r_g 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^r_n and f^r_g 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^*.[14]

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 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

[14] The fact that the threshold x^* is the same for both defendant races is neither a necessary nor a sufficient condition for unbiasedeness of the court. It is used in figure 2 purely for illustrative purposes. Also, the fact that we allow the distributions of idiosyncratic costs v 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.

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:[15]

$$E(m,W) > E(m,M) \iff E(w,W) > E(w,M) \tag{8}$$

Expression (8) 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.

3.3 Discussion and extensions

Bias in the collection of evidence

A crucial assumption underlying our test is that the functions $F^r_n(x)$ and $F^r_g(x)$ do not depend jointly on the race of the defendant and on the race of the victim. A similar restriction is common to other outcome based tests of prejudice, such as Knowles et al. (2001) and Anwar and Fang (2006). Nonetheless, even if we were to allow dependence of $F^r_n(x)$ and $F^r_g(x)$ on R as well as r, it is unclear in which direction the bias would go. Suppose for instance that such dependence on race pairs were due to police work. On the one hand, the police may work harder to get accurate evidence in cases involving minority defendants and white victims. This would make it easier for the courts to "separate" guilt and innocence and the proportion of errors would be lower for these cases. On the other hand, a biased police may "fabricate" evidence against minorities when the victim is white. In this case the bias would go in the opposite direction, i.e. we would find more errors for minority defendants/white victims pairs.[16]

[15] Expression (8) refers to the case of bias in favor of white victims. Obviously for bias in favor of minority victims the inequalities should be reversed. 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.

[16] Empirically, Radelet and Pierce (1985) analyzed a sample of 1017 homicides in Florida in the period 1973-77 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 than by the police.

Bias in the decisions of superior courts

Another hypothesis in our model is that superior courts never make mistakes and are racially unbiased. If superior courts made errors which were uncorrelated with the combination of defendant/victim race, this would not invalidate our test of racial bias. We can also relax this assumption in one direction. Suppose for example that superior courts were racially biased in the same direction of lower courts. This would go against finding higher errors 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.

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. 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 US.[17] This dataset includes information on race of the defendant and of the victim, as well

[17] U.S. Dept. of Justice, Bureau of Justice Statistics. *Murder Cases in 33 Large Urban Counties in the United States, 1988.* Distributed by ICPSR 9907, 1996.

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.[18]

4 The data

To implement our test of racial bias we could not rely on any readily available dataset. In fact all existing data sets 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 or not in appeal. The only comprehensive dataset containing information on judicial errors in capital cases, that is the one use in Liebman et al's (2000) study and compiled by Fagan and Liebman (2002, from now on 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 et al. (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 US. We use two datasets originally compiled by FL:[19]

- DIRECT APPEAL dataset (DA from now on): 4,546 state capital cases whose direct appeal decisions became final between January 1, 1973 and December 31, 1995.[20] This is the universe of all capital sentences that were reviewed on direct appeal by a state high court.
- HABEAS CORPUS dataset (HC from now on): 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.

[18]Specifically, our estimated linear probability model is

$$Plea = \underset{(.052)}{.398} - \underset{(.066)}{.014}\, ND + \underset{(.042)}{.033}\, WV + \underset{(.065)}{.037}\, (ND * WV)$$

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

[19]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.

[20]"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: (1) affirmed the capital judgment or (2) overturned the capital judgement (either the conviction or the sentence) on one or more grounds" (Liebman et al. (2000), p. 126).

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.[21]

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, a serious error is reversible if, in addition to satisfying the three conditions required for DA, it violates the federal Constitution.[22]

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 Lexis Nexis 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 web sites specialized in death penalty issues, plus communications with police officers and defense lawyers.[23] After over two years of attempts we managed to construct an almost complete data set 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 US that contains information on race of defendant and victim.

[Table 1 about here]

Table 1 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). Appendix Tables A1 and A2 contain summary statistics on the share of missing observations by state and by year.

[21] 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 Lexis-Nexis, 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).

[22] Some additional technical rules for reversibility at the HC stage are listed in Liebman et al. (2000), p. 130.

[23] A detailed description of the search procedure and of the sources is available from the authors upon request.

[Table 2 about here]

In Table 2 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 sub sample 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.

Panel A refers to HC cases. For demographic characteristics of the defendant like race and gender we do not find any significant difference across the two sub samples. Age is higher for defendants whose victims' race is missing (p-value .11). Characteristics like prior felony convictions, history of drug abuse, and deprived family background are also balanced across the two sub samples. History of alcohol abuse is more frequent among cases for which victim's race is available (p-value .09). Turning to victims' characteristics: the number of victims, the presence of policemen or public officials among the victims, and whether the victim had high status in the community all have similar means when victims' race is missing and when it is not. The only characteristic for which the difference is statistically significant is the gender of the victim, but we show below that our results are robust to excluding female victims.

Turning to crime characteristics, we report the means of two variables that may be correlated with unobservables specific to pairs of defendant/victim races. The first is whether the defendant knew the victim: crimes against unknown victims may be considered as relatively more threatening by the jury and be sanctioned more harshly. The second is a dummy for whether the crime involved heinous, atrocious, cruel circumstances. We do not find a statistically significant difference in means across sub samples for either of these measures. These findings increase 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.

In Table 2B we repeat a similar analysis for the DA dataset, only for a smaller number of variables because the information coded by FL in the DA dataset is much more limited. Again, characteristics like gender and race of the defendant are perfectly balanced, while defendants' age is higher by 1.5 years in the sample for which victims' race is missing. The number and gender of the victims, as well as the likelihood that the victim is the partner, are not statistically different across subsamples. On the other hand, the likelihood that the victim is a policeman or a public official is significantly higher when victim race is available, likely due to media coverage. One of the robustness tests we conduct below is to exclude cases where the victim is a policeman or public official, and our results hold.

4.1 Descriptive statistics

Table 3 reports summary statistics on the main variables of interest in the HC (Panel A) and DA (Panel B) datasets.

[Table 3 about here]

In the HC data the error rate, measured by the variable "Relief" as the fact that relief is granted at some stage of the review process, is .36. Regarding the race of the defendant, 51 percent of the cases involve white defendants, 44 percent African Americans, with the remaining fraction being 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 non-white 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 non-white victim. The remaining cases are split between non-whites who killed non-whites (13 percent) and whites who killed whites (48 percent). The proportions are fairly similar for the DA sample (first appeal): 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 non-white defendants who killed a white victim is .30.

5 Results

The test for racial bias we derived in section 3 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 $\widehat{E(r,R)}$, the average error rate for cases in which a defendant of race r killed a victim of race R. We test the null $\widehat{E(r,W)} = \widehat{E(r,M)}$ (absence of racial bias) against the alternative $\widehat{E(r,W)} > \widehat{E(r,M)}$ (racial bias in favor of white victims) using the Z-statistic:

$$Z = \frac{\widehat{E(r,W)} - \widehat{E(r,M)}}{\sqrt{\frac{SVar_{rW}}{n_{rW}} + \frac{SVar_{rM}}{n_{rM}}}} \tag{9}$$

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 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 (9) 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 (8).

5.1 Main results

[Table 4 about here]

Table 4 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, $\widehat{E(r,R)}$, and the associated standard error (in parenthesis). The p-values reported at the end of each row are those associated with test statistic (9). 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 Table 4, Panel A 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 non-white, with a difference of -11 percentage points.[24] On the other hand, in cases involving a minority defendant, the error rate is 37.5 if the victim is white, and 28.4 percent if it is non-white, with a difference of $+9$ percentage points (or a 32 percent increase over the the non-white/non-white 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 with a p-value of .08; for cases involving white defendants we fail to reject the null against the alternative (p-value .80). 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. 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 non-white victim, though not statistically significant. In the case of minority defendants the difference is $+3$ percentage points (a 9 percent increase over the the non-white/non-white error rate of .35) and is significant at the10 percent level. Again, we reject the null of no racial bias on defendant/victim racial pairs.

Note that our 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 differences 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.

[Table 5 about here]

In Table 5 we examine whether the pattern of results differs across regions. We find that the pattern of racial bias we uncovered is driven by Southern states. 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 non-whites: the difference is striking at 15.4 percentage points (a 66 percent increase over the non-white/non-white error rate of .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 (Table 5B), 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.

When we conduct analogous tests for all other regions we fail to reject the null based on our rank order test for both HC and DA. In HC the error rate is higher with non-white than with white victims both if the defendant is white and if he is not. In DA the sign

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

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.

5.2 Potential confounding factors

So far we have interpreted the results of our rank order tests as indicative of potential racial bias on behalf of the trial court. An alternative interpretation would be that the pattern of inequalities in error rates is generated by unobserved characteristics that are systematically correlated with different combinations of defendant and victim races. In the notation of our model, this would amount to allowing the distribution of the evidence to depend on *both* races, e.g., $F_n^{r,R}(x)$ 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 by conditioning on a set of available characteristics that might be correlated with such factors.

Crime characteristics

One may conjecture that cases involving minority defendants and white victims may differ in the type of crime involved and that the difference in error rates reflects characteristics of the crime rather than racial bias. The short answer is that we do not find any evidence in support of this interpretation. Although this does not eliminate the possibility that differences in unobservables exist -and indeed this is one of the motivating factors of our analysis- if we find that our results are not affected by conditioning on observables that proxy for the type of crime.

[Table 6 about here]

In table 6 we present our evidence. In Panel A we start from the HC sample, for which relatively detailed information on the crime was collected by FL. First we test whether the gender of the victim is a significant factor in our results. In the first panel of Table 6A we restrict the attention to cases in which none of the victims was female. We find higher error on minority defendants who killed white men than on those who killed non-white men (the difference is 10.6 percentage points, p-value .12). The corresponding difference for white defendants is -11.6 (p-value .77).

An aggravating factor that may be responsible for the results we find is the fact that the defendant killed a police, or fireman, or 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 was 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 .03. So our results are not driven by this types of murders.

Another aggravating factor might be the presence of multiple victims. Restricting the analysis to homicides with only one victim shows a difference of 9 percentage points for non-white defendants who killed white versus nonwhite victims (p-value .09) and an insignificant difference on the opposite direction for white defendants.

The remaining of Table 6A reports results for other crime characteristics which are available only for the HC sample. 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 with more severe punishments. One could conjecture that cases involving minority defendants and white victims fall disproportionately in this category. In the fourth panel of Table 6A 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 this subsample 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 .03). 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) 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 panel of Table 6A 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 non-white, and is significant at the 5 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 a 13 percentage points higher error rate for nonwhite defendants who killed white victims (p-value .07), and no corresponding difference for white defendants.

We have less information on crime characteristics for the DA compared to the HC sample. In Table 6B we show what we have for the DA sample. First we test 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-value .01). The corresponding difference for white defendants is in the opposite direction and not significant. In the second panel we only consider murders that did not involve public officials and we find a difference in error rates of 3 percentage points for minority defendants (with a borderline p-value .105), and no significant difference for white defendants. Finally, when we restrict the analysis to homicides with only one victim we find again a difference of 3 percentage points for nonwhite defendants who killed white vs. nonwhite victims, but the p-value increases to .16.

Legal assistance

Differences in error rates could be due to unequal quality of legal assistance of the defendant. Thus a possible interpretation of our main finding is that minority defendants who killed a white victim receive systematically worse legal assistance compared to minority defendants who killed a minority victim. Incidentally this may be another source of racial bias, but it would be of a different nature than the one modelled in this paper and would not be a bias of the courts.

[Table 7 about here]

In Table 7 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.[25] We start by restricting the sample to 381 HC cases in which ineffective assistance of counsel was *not* raised as the *first* claim in the appeal. In this subset of cases the difference in error rates for minority defendants who killed a white vs. a non-white victim is 14.5 percentage points (p-value .02). If we further restrict the sample to the 220 cases in which ineffective assistance of counsel was not raised at all among the claims, the difference increases to 19 percentage points (p-value .04). Comparing these results to those in table 4 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 7, 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 sub samples the order of magnitude of the differences in error rates and the significance level remain comparable to those of table 4, and the rank order test rejects the null of absence of racial bias according to our model (except for the third panel where the p-value for nonwhite defendants increases to .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 claim to present first, etc.) and in this sense they may not be fully objective. Nonetheless, we take the evidence in table 7 as suggestive that differences in the quality of legal assistance are not entirely responsible for our results.

5.3 Possible bias of appeal courts

So far we have assumed that the appeal courts are unbiased. As we mentioned above, errors uncorrelated with pairs of defendant/victim races are irrelevant for our empirical test. 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 non-white defendants who killed white victims. Although we cannot rule this out a priori, we question 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

[25] Fagan and Liebman's (2002) dataset contains the list of claims raised, as well as the order in which the claims were raised.

conservative ones. Normally, the ideology of judges is correlated with party affiliation, hence 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 U.S. District Courts, the U.S. Courts of Appeals, the Supreme Court and the U.S. 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.

[Table 8 about here]

Table 8A 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. Both sets of results are consistent with our earlier findings, and indicate a higher likelihood of relief for nonwhite defendants who killed white victims. The magnitude of the difference in error rates is 7 percentage points when we consider appeal courts where a majority of the judges are Republican (first panel, p-value .20), and 13 percentage points when we restrict our test to courts that are entirely composed of republican-appointed judges (second panel, p-value .09). In the third panel we consider the possibility that the political climate in a given year may affect relief rates, and restrict the sample to Habeas Corpus appeal sentences that occurred under a Republican administration. We find a difference of 17 percentage points for non-white defendants who killed white victims (p-value .01). In all three panels the corresponding difference for white defendants is in the opposite direction and not significant.

In Table 8B 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 PAJID.[26] The first panel of Table 8B 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 7.5 percentage points higher than on those who killed nonwhite victims (p-value .05). 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 et al. (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 Table 4, though we lose statistical significance. Furthermore, this result does not depend on the particular cutoff for the definition of "conservative". Figure 3 shows that the positive difference in error rates for minority defendants who killed white versus nonwhite

[26] Essentially PAIJD 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.

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.

[Figure 3 about here]

To sum up, we find no evidence that left liberal 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.

6 Conclusions

This paper proposes an outcome based test for racial bias in capital sentencing in the US. We use the share of judicial errors in first degree sentencing as an indicator of racial bias of such courts. Our maintained assumption is that superior courts (especially the federal courts dealing with the Habeas Corpus final appeal stage) have less racial prejudice or no prejudice at all. Note that, if they had, this would bias our results against finding bias in the first sentencing. Using an originally collected dataset, we uncover a bias against minority defendants killing white victims. 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. The same does not hold for white defendants. 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

To simplify the notation, in sections A.1 and A.2 we omit subscripts and superscripts related to race, i.e. we write x instead of x_{rR} and f_g, f_n instead of f_g^r, f_n^r.

A.1 Proof of uniqueness of the equilibrium

Claim 1. There exists an $\widehat{x} \in [0,1)$ such that $\frac{\partial v^*(x)}{\partial x} > 0$ for all $x > \widehat{x}$.

Proof. From *(1)* we can calculate the derivative of $v^*(x)$ with respect to x as

$$\begin{aligned} \frac{\partial v^*(x)}{\partial x} &= f_g(x)b + f_g(x)c_g - f_n(x)c_n \\ &= f_n(x)c_n\left[\frac{f_g(x)}{f_n(x)}\frac{(b+c_g)}{c_n} - 1\right]. \end{aligned} \tag{A1}$$

From MLRP, $\frac{f_g(x)}{f_n(x)}$ is strictly increasing in x. By assumption $\frac{f_g(x)}{f_n(x)} \to +\infty$ as $x \to 1$. Therefore there exists a value $\widehat{x} \in [0,1)$ such that the expression in square brackets in (A1) is positive for all $x > \widehat{x}$.

Claim 2. If x^* is an equilibrium, then $x^* > \widehat{x}$.

Proof. From (1) we have $v^*(0) = c_n - c_g < 0$. From (5) we have $\Im(v^*(0)) = 0$ because $\Im$ has support in $\mathbb{R}^+$. Furthermore, $\Im(v^*(x)) = 0$ for all $x \leq \widehat{x}$. Suppose that in equilibrium $x^* < \widehat{x}$. Then we would have $\pi(x^*) = \Im(v^*(x^*)) = 0$. But in this case the optimal response of the court would be to set $x^* = 1 > \widehat{x}$, a contradiction.

Claim 3. The equilibrium x^* is unique.

Suppose there were two equilibria, x_0^* and x_1^*, with $\widehat{x} < x_0^* < x_1^*$. From Claim 1 this would imply $0 < v^*(x_0^*) < v^*(x_1^*)$, and in turn $\pi(x_0^*) < \pi(x_1^*)$. But then the optimal response of the court would involve setting $x_0^* > x_1^*$, a contradiction.

A.2 Proof that the equilibrium error rate is decreasing in x^*

The error rate (7) can be rewritten as

$$E(r, R) = 1 - \frac{1}{1 + \frac{1-\pi(x^*)}{\pi(x^*)} \frac{1-F_n(x^*)}{1-F_g(x^*)}}. \tag{A2}$$

Expression (A2) is decreasing in x^* if and only if $\frac{1-\pi(x^*)}{\pi(x^*)} \frac{1-F_n(x^*)}{1-F_g(x^*)}$ is decreasing in x^*. Taking the first derivative of this product with respect to x^* we obtain:

$$-\frac{1}{[\pi(x^*)]^2} \frac{\partial \pi(x^*)}{\partial x^*} + \frac{1}{[1 - F_g(x^*)]^2} \left[\int_{x^*}^{1} f_g(x^*) f_n(x) dx - \int_{x^*}^{1} f_g(x) f_n(x^*) dx \right]. \tag{A3}$$

The first addendum in (A3) is negative because the equilibrium $\pi(x^*)$ is increasing in x^*, following Claims 1 and 2 above. To see that the second addendum is also 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 (A3) are calculated for $x \in (x^*, 1]$, then in this range $f_g(x^*) f_n(x) < f_g(x) f_n(x^*)$, hence the expression in square brackets is negative.

A.3 Case with random race of the victim

Consider the case in which the defendant cannot choose the race of the victim and the latter is unknown ex ante. Define $\beta \in (0, 1)$ as the exogenous probability that the victim of the crime is white. The expected payoff to an individual of race r from committing the crime is:

$$\begin{aligned} &\beta \left\{ \left[1 - F_g^r(x_{rW})\right](-v - c_g) + F_g^r(x_{rW})(b - v) \right\} \\ &+(1 - \beta) \left\{ \left[1 - F_g^r(x_{rM})\right](-v - c_g) + F_g^r(x_{rM})(b - v) \right\}. \end{aligned}$$

The payoff from not committing a crime is:

$$-\beta \left[1 - F_n^r(x_{rW})\right] c_n - (1 - \beta) \left[1 - F_n^r(x_{rM})\right] c_n.$$

Let us define

$$\begin{aligned} \Gamma_g(x_{rW}, x_{rM}) &\equiv \beta F_g^r(x_{rW}) + (1 - \beta) F_g^r(x_{rM}) \\ \Gamma_n(x_{rW}, x_{rM}) &\equiv \beta F_n^r(x_{rW}) + (1 - \beta) F_n^r(x_{rM}) \end{aligned}$$

Following the same procedure as in the text, we obtain the threshold level of v below which a crime is committed.

$$v \leqslant \Gamma_g(x_{rW}, x_{rM})b - [1 - \Gamma_g(x_{rW}, x_{rM})]\, c_g + [1 - \Gamma_n(x_{rW}, x_{rM})]\, c_n \equiv v^*(\beta, x_{rW}, x_{rM}).$$

Obviously, $v^*(\cdot)$ depends on all the other parameters, namely β, b, c_g and c_n, but the latter do not depend upon the races neither of the defendant nor of the victim and are common knowledge. Relative to the case developed in the text, now the choice of each potential criminal depends on both cutoff points relative to the race of the victim. Repeating the same steps of the proof in the text one reaches the same implications for our test of racial bias.

A.4 Case where the race of the victim can be chosen

Consider now the case in which the criminal can choose the race of the victim. Under the assumptions of our model if the court were biased and this led to setting a lower threshold of evidence x^* for, say, white victims, all potential criminals would choose minority victims. If instead the court were unbiased potential criminals would be indifferent on the race of the victim and would randomize. This implies that under the assumptions of our model in the presence of bias we should not observe in equilibrium a condition (killing white victims) that allows us to test for bias. To be able to derive a test for bias in cases where the race of the victim is a choice variable one should adopt a different theoretical framework, e.g. one in which there are differential benefits to killing victims of different races or the potential criminal was uncertain about the bias of the court or the distribution of the signal. This goes beyond the scope of the current analysis.

References

[1] Abrams D. M. Bertrand and S. Mullainathan (forthcoming), "Do Judges Vary in their Treatment of Race?", *Journal of Legal Studies*.

[2] Antonovics, K. and B. Knight (2009), "A New Look at Racial Profiling: Evidence from the Boston Police Department", *Review of Economics and Statistics*, 91(1), 163-175.

[3] Anwar, S. and H. Fang (2006), "An Alternative Test of Racial Profiling in Motor Vehicle Searches: Theory and Evidence", *American Economic Review*, 96(1), 127-151.

[4] Argys, L. and N. 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), pp. 255-92.

[5] Barnes, K., D. Sloss, and S. Thaman (2008), "Life and Death Decisions: Prosecutorial Discretion and Capital Punishment in Missouri", unpublished.

[6] Becker G. (1957) *The Economics of Discrimination*, Chicago: University of Chicago Press

[7] Blume, J., T. Eisenberg and M.T. Wells (2004), "Explaining Death Row's Population and Racial Composition", *Journal of Empirical Legal Studies*, 1(1), 165-207.

[8] Brace, P., L. Langer and M. Gann Hall (2000) " Measuring the Preferences of Supreme Court Judges" *Journal of Politics,May 387-413*

[9] Coyne, R. and L. Entzeroth (2006) *Capital Punishment and the Judicial Process,* Carolina Academic Press.

[10] Donohue III, J. and J. Wolfers (2005), "Uses and Abuses of Empirical Evidence in the Death Penalty Debate", *Stanford Law Review*, 58(3), 791-846.

[11] Erlich, I. (1975), "The Deterrent Effect of Capital Punishment: A Matter of Life and Death", *American Economic Review*, 65(3), 397-417.

[12] Fagan, J. and J. Liebman (2002), PROCESSING AND OUTCOME OF DEATH PENALTY APPEALS AFTER FURMAN V. GEORGIA, 1973-1995. ICPSR version 3468. New York City, NY: Columbia School of Law and Mailman School of Public Health [producers]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

[13] Gennaioli N. and A Shleifer (2007) "The Evolution of Common Law" *Journal of Political Economy, 115 (1) 43-67.*

[14] Gross, S.R. and R. Mauro (1984), "Patterns of Death: An Analysis of Racial Disparities in Capital Sentencing and Homicide Victimization", *Stanford Law Review*, 37, 27-153

[15] Grossman G. and M. Katz (1983) "Plea Bargain and Social Welfare" *American Economic Review, 73 (4) 749-57*

[16] Humphrey, J. and T. J. Fogarty (1987), "Race and Plea Bargained Outcomes: A Research Note," *Social Forces*, 66, 176-85.

[17] Katz, L., S. Levitt and E. Shustorovich (2003), "Prison Conditions, Capital Punishment, and Deterrence", *American Law and Economics Review*, 5(2), 318-343.

[18] Knowles J., N. Persico and P. Todd (2001), "Racial Bias in motor vehicles searches:Theory and Evidence," *Journal of Political Economy*, 109, 203-29.

[19] Iyengar, R. (2007), "An Analysis of Attorney Performance in the Federal Indigent Defense System", NBER Working Paper, n. 13187.

[20] Liebman, J., J. Fagan and V. West (2000), *A Broken System: Error Rates in Capital Cases, 1973-1995*; Columbia University.

[21] Radelet, M. and G. Pierce (1985), "Race and Prosecutorial Discretion in Homicide Cases",19 *Law and Society Review*, 587, 601–15.

[22] Reinganum, J. (1988), "Plea Bargain and Prosecutorial Discretion", *American Economic Review, 78 (4), 713-27.*

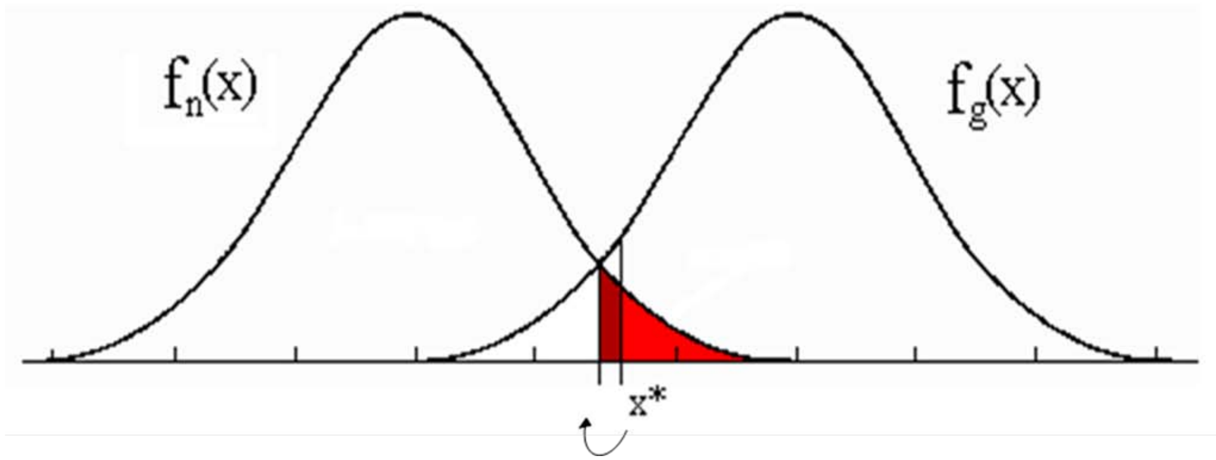

Figure 1:
Error rate for given defendant's race

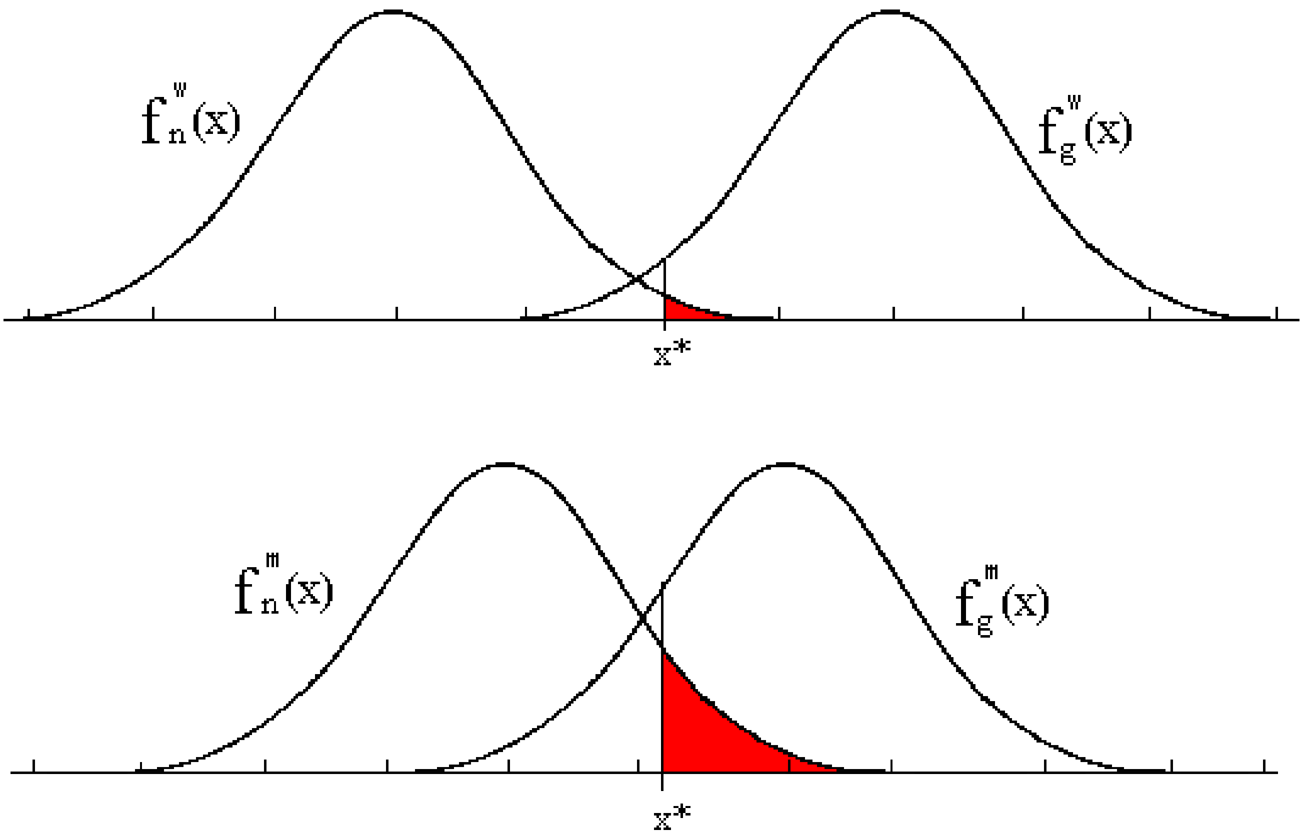

Figure 2:
Error rate for given victim's race

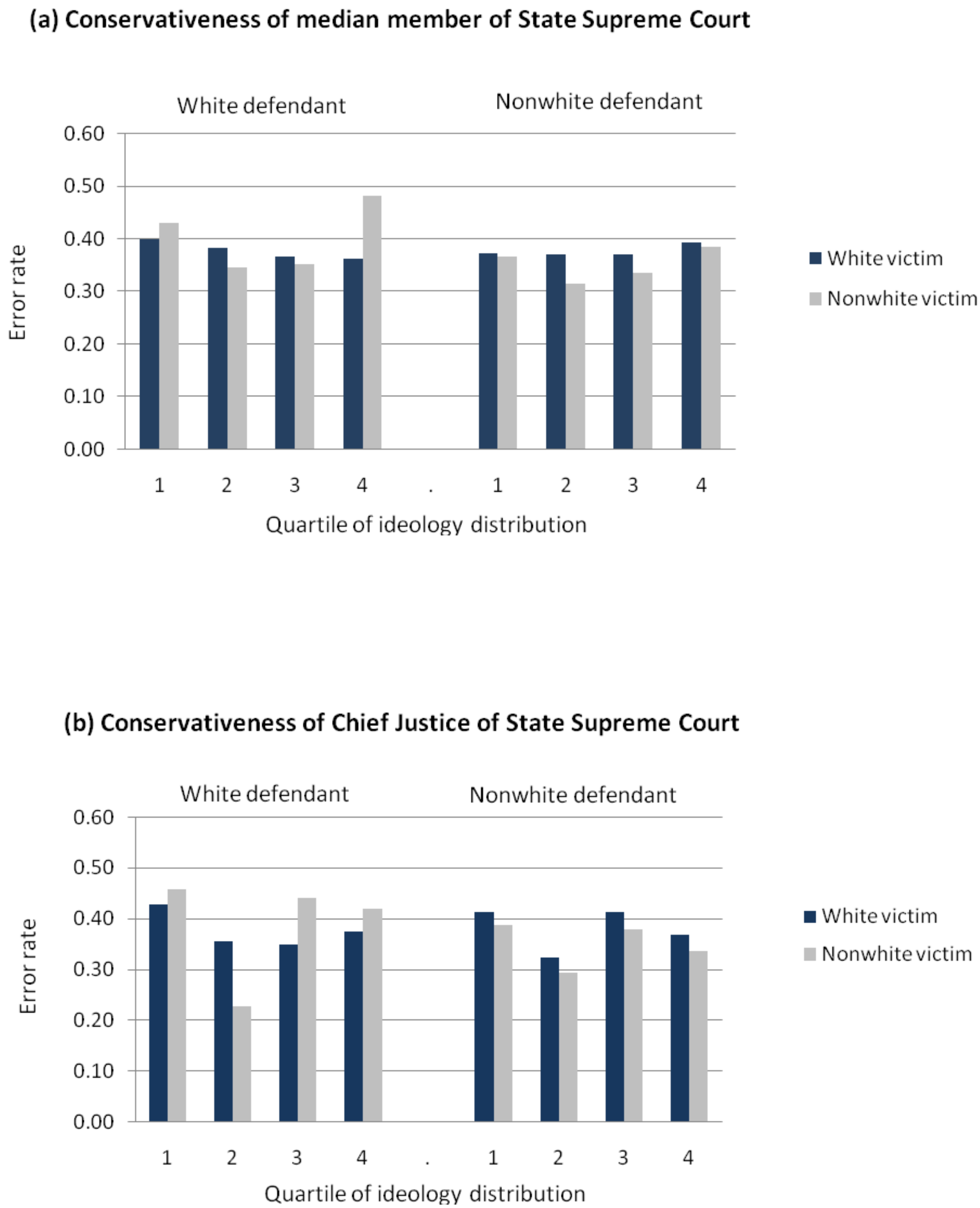

Figure 3:
Error rates and ideology of Appeal Courts

Table 1: Missingness of race

Panel A: Habeas Corpus

		Missing victim's race		
		No	*Yes*	*Total*
Missing defendant's race	*No*	508	20	528
	Yes	3	0	3
	Total	511	20	531

Panel B: Direct Appeal

		Missing victim's race		
		No	*Yes*	*Total*
Missing defendant's race	*No*	3,717	447	4,146
	Yes	130	122	252
	Total	3,847	569	4416

Table 2: Selection in missingness of victim's race

Variable	*Nonmissing victim's race*	*Missing victim's race*	*Diff=0 (p-val)*
Panel A: Habeas Corpus			
Defendant characteristics			
Defendant is White	0.51	0.45	0.59
Defendant is African American	0.44	0.50	0.60
Male defendant	0.99	1.00	0.66
Age of defendant	28.2	41.5	0.11
Prior felony	0.22	0.25	0.79
History of alcohol abuse	0.13	0.00	0.09
History of drug abuse	0.17	0.10	0.44
Deprived/Abused background	0.02	0.00	0.57
Victim characteristics			
Number of victims	1.41	1.22	0.65
Female victim	0.48	0.24	0.05
High status victim	0.23	0.30	0.46
Police victim	0.09	0.00	0.16
Crime characteristics			
Defendant knew victim	0.26	0.25	0.94
Heinous crime	0.40	0.30	0.37
Panel B: Direct Appeal			
Defendant characteristics			
Defendant is White	0.51	0.52	0.68
Defendant is African American	0.41	0.40	0.69
Male defendant	0.98	0.98	0.50
Age of defendant	31.3	32.9	0.00
Victim characteristics			
Number of victims	1.37	1.44	0.22
Female victim	0.52	0.50	0.41
Police victim	0.05	0.02	0.00
Partner victim	0.05	0.04	0.25

Table 3: Summary statistics

Panel A: Habeas Corpus

	Sample with non-missing victim's race		
Variable	*No. Obs*	*Mean*	*Std. Dev.*
Relief	508	0.36	0.48
African American defendant	508	0.44	0.50
White defendant	508	0.51	0.50
African American victim	508	0.13	0.34
White victim	508	0.83	0.37
White def., Non-white vict.	508	0.03	0.18
Non-white def., White vict.	508	0.36	0.48
White def., White vict.	508	0.48	0.50
Non-white def., Non-white vict.	508	0.13	0.34

Panel B: Direct Appeal

	Sample with non-missing victim's race		
Variable	*No. Obs*	*Mean*	*Std. Dev.*
Relief	3717	0.37	0.48
African American defendant	3717	0.41	0.49
White defendant	3717	0.51	0.50
African American victim	3717	0.17	0.38
White victim	3717	0.78	0.41
White def., Non-white vict.	3717	0.03	0.17
Non-white def., White vict.	3717	0.30	0.46
White def., White vict.	3717	0.48	0.50
Non-white def., Non-white vict.	3717	0.19	0.39

Table 4: Error rates by Defendant and Victim's race

Panel A: Habeas Corpus

Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
	Victim's race			
White	0.358	0.471	0.809	260
	(0.031)	(0.125)		
Non-white	0.375	0.284	0.083	251
	(0.036)	(0.055)		
N.obs	427	84		

Panel B: Direct Appeal

Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
	Victim's race			
White	0.373	0.395	0.673	1908
	(0.011)	(0.046)		
Non-white	0.377	0.347	0.097	1809
	(0.015)	(0.018)		
N.obs	2911	806		

Note: Standard errors of the means in parenthesis

Table 5: Error rates by race and region

Panel A: Habeas Corpus

	South				**Other regions**			
	Victim's race				Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*	*White*	*Non-white*	*p-values*	*N.obs*
White	0.350	0.455	0.741	208	0.391	0.500	0.678	52
	(0.034)	(0.157)			(0.073)	(0.224)		
Non-white	0.387	0.232	0.012	219	0.286	0.545	0.917	32
	(0.038)	(0.057)			(0.101)	(0.157)		
N.obs	360	67			67	17		

Panel B: Direct Appeal

	South				**Other regions**			
	Victim's race				Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*	*White*	*Non-white*	*p-values*	*N.obs*
White	0.397	0.443	0.785	1305	0.322	0.286	0.324	603
	(0.014)	(0.056)			(0.020)	(0.077)		
Non-white	0.409	0.376	0.134	1234	0.288	0.304	0.657	575
	(0.017)	(0.024)			(0.026)	(0.028)		
N.obs	2048	491			863	315		

Note: Standard errors of the means in parenthesis

Table 6: Error rates conditional on crime characteristics

Panel A: Habeas Corpus

No female victim				
	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.384	0.500	0.769	124
	(0.046)	(0.151)		
Non-white	0.363	0.257	0.117	137
	(0.048)	(0.075)		
N.obs	214	47		
No police victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.367	0.438	0.703	242
	(0.032)	(0.128)		
Non-white	0.391	0.262	0.030	222
	(0.039)	(0.057)		
N.obs	387	77		
Single victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.359	0.471	0.808	254
	(0.031)	(0.125)		
Non-white	0.383	0.292	0.089	245
	(0.036)	(0.057)		
N.obs	417	82		
Defendant not connected to community where crime occurred				
	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.384	0.500	0.790	191
	(0.037)	(0.139)		
Non-white	0.413	0.265	0.027	187
	(0.042)	(0.064)		
N.obs	315	63		
No high status victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.398	0.385	0.463	209
	(0.035)	(0.140)		
Non-white	0.397	0.309	0.124	186
	(0.043)	(0.063)		
N.obs	327	68		
Robbery				
	White	*Non-white*	*p-values*	*N.obs*
White	0.342	0.600	0.848	81
	(0.055)	(0.245)		
Non-white	0.430	0.250	0.045	103
	(0.056)	(0.090)		
N.obs	155	29		
Felony				
	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.344	0.429	0.658	129
	(0.043)	(0.202)		
Non-white	0.385	0.257	0.073	152
	(0.045)	(0.075)		
N.obs	239	42		

Table 6 (cont'd): Error rates conditional on crime characteristics

Panel B: Direct Appeal

No female victim				
	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.358	0.403	0.768	831
	(0.017)	(0.058)		
Non-white	0.413	0.333	0.008	926
	(0.020)	(0.026)		
N.obs	1343	414		
No police victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.371	0.394	0.678	1825
	(0.012)	(0.048)		
Non-white	0.376	0.346	0.105	1688
	(0.015)	(0.018)		
N.obs	2738	775		
Single victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.389	0.374	0.389	1419
	(0.013)	(0.051)		
Non-white	0.388	0.361	0.163	1382
	(0.016)	(0.022)		
N.obs	2220	581		

Note: Standard errors of the means in parenthesis

Table 7: Error rates conditional on trial characteristics

Habeas Corpus

Ineffective assistance of counsel not 1st claim

	White	*Non-white*	*p-values*	*N.obs*
White	0.354	0.385	0.585	194
	(0.036)	(0.140)		
Non-white	0.374	0.229	0.025	187
	(0.041)	(0.061)		
N.obs	320	61		

Ineffective assistance of counsel not in any claim

	White	*Non-white*	*p-values*	*N.obs*
White	0.458	0.571	0.708	114
	(0.048)	(0.202)		
Non-white	0.417	0.227	0.037	106
	(0.054)	(0.091)		
N.obs	191	29		

Prosecutorial suppression/witholding of evidence not in any claim

	White	*Non-white*	*p-values*	*N.obs*
White	0.383	0.600	0.902	177
	(0.038)	(0.163)		
Non-white	0.383	0.300	0.162	173
	(0.042)	(0.073)		
N.obs	300	50		

Improper interrogation not in any claim

	White	*Non-white*	*p-values*	*N.obs*
White	0.362	0.500	0.849	234
	(0.033)	(0.129)		
Non-white	0.387	0.283	0.070	223
	(0.038)	(0.059)		
N.obs	381	76		

Note: Standard errors of the means in parenthesis

Table 8: Possible bias of Appeal Courts

Panel A: Habeas Corpus

Majority of final federal panel Republican

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.355	0.455	0.730	149
	(0.041)	(0.157)		
Non-white	0.284	0.216	0.200	146
	(0.043)	(0.069)		
N.obs	247	48		

All judges appointed under Republican

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.250	0.500	0.799	40
	(0.073)	(0.289)		
Non-white	0.216	0.083	0.096	63
	(0.058)	(0.083)		
N.obs	87	16		

Sentence under Republican administration

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.347	0.438	0.751	183
	(0.037)	(0.128)		
Non-white	0.402	0.234	0.013	179
	(0.043)	(0.062)		
N.obs	299	63		

Panel B: Direct Appeal

Majority of State Supreme Court Republican

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.285	0.280	0.477	428
	(0.023)	(0.092)		
Non-white	0.292	0.217	0.051	368
	(0.031)	(0.034)		
N.obs	619	177		

Median ideology of State Supr Court conservative

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.364	0.406	0.745	1151
	(0.015)	(0.062)		
Non-white	0.380	0.354	0.199	1050
	(0.018)	(0.025)		
N.obs	1781	420		

Chief justice of State Supr Court conservative

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.363	0.426	0.845	1155
	(0.015)	(0.060)		
Non-white	0.385	0.352	0.142	1072
	(0.018)	(0.025)		
N.obs	1790	437		

Note: Standard errors of the means in parenthesis

Tab A1: Missingness of Victim's race by year and state, Habeas Corpus

Years of 1st sentence	No. Total obs	No. Missing obs	Share missing
1973	4	0	0.00
1974	25	4	0.16
1975	30	5	0.17
1976	25	2	0.08
1977	45	3	0.07
1978	48	1	0.02
1979	51	2	0.04
1980	53	1	0.02
1981	66	0	0.00
1982	68	1	0.01
1983	41	0	0.00
1984	26	1	0.04
1985	25	0	0.00
1986	15	0	0.00
1987	6	0	0.00
1988	2	0	0.00
1989	1	0	0.00
			.

State	No. Total obs	No. Missing obs	Share missing
AL	19	1	0.05
AR	24	0	0.00
AZ	14	1	0.07
CA	4	0	0.00
DE	2	0	0.00
FL	95	4	0.04
GA	84	2	0.02
ID	3	0	0.00
IL	10	0	0.00
IN	4	0	0.00
KY	1	0	0.00
LA	34	0	0.00
MD	1	0	0.00
MO	26	1	0.04
MS	21	0	0.00
MT	4	0	0.00
NC	10	0	0.00
NE	6	0	0.00
NV	4	0	0.00
OK	11	2	0.18
PA	3	0	0.00
SC	7	0	0.00
TN	1	0	0.00
TX	108	9	0.08
UT	3	0	0.00
VA	27	0	0.00
WA	3	0	0.00
WY	2	0	0.00

Tab A2: Missingness of Victim's race by year and state, Direct Appeal

Years od 1st sentence	No. Total obs	No. Missing obs	Share missing
1974	10	1	0.10
1975	27	9	0.33
1976	45	9	0.20
1977	61	22	0.36
1978	77	15	0.19
1979	118	23	0.19
1980	128	25	0.20
1981	161	20	0.12
1982	159	16	0.10
1983	219	17	0.08
1984	238	18	0.08
1985	271	27	0.10
1986	215	17	0.08
1987	246	20	0.08
1988	320	30	0.09
1989	253	30	0.12
1990	238	22	0.09
1991	280	28	0.10
1992	300	27	0.09
1993	258	24	0.09
1994	306	26	0.08
1995	234	21	0.09

State	No. Total obs	No. Missing obs	Share missing
AL	256	50	0.20
AR	75	9	0.12
AZ	189	31	0.16
CA	229	26	0.11
CO	2	0	0.00
CT	2	0	0.00
DE	22	0	0.00
FL	709	80	0.11
GA	269	11	0.04
ID	31	11	0.35
IL	218	0	0.00
IN	67	0	0.00
KY	44	13	0.30
LA	88	4	0.05
MD	42	6	0.14
MO	81	6	0.07
MS	110	3	0.03
MT	13	2	0.15
NC	226	6	0.03
NE	22	5	0.23
NJ	36	3	0.08
NM	8	0	0.00
NV	91	17	0.19
OH	104	0	0.00
OK	186	32	0.17
OR	29	4	0.14
PA	178	42	0.24
SC	115	11	0.10
TN	102	13	0.13
TX	495	51	0.10
UT	13	3	0.23
VA	93	4	0.04
WA	15	4	0.27
WY	4	0	0.00