A Test of Racial Bias in Capital Sentencing*

Alberto Alesina
Harvard University and IGIER

Eliana La Ferrara
Bocconi University and IGIER

First Draft: April 2009
Revised: April 2010

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 the race of the defendant, or of the victim, or of the combination of the two. The key prediction of the model is that if the court is unbiased, ex post the error rate should be independent of the race of the defendant and/or the victim. We test this prediction using an original dataset that contains the outcomes of all capital appeals that became final between 1973 and 1995, together with the race of the defendant and of the victim(s). We do not find evidence of bias when looking at the race of the defendant per se. However, in Habeas Corpus cases we find strong and robust evidence of bias against minority defendants who killed white victims: in the full sample, the probability of error in these cases is 10 percentage points higher than for minority defendants who killed minority victims. If we restrict the analysis to Southern states, this difference becomes 15.4 percentage points.

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 popula-

*We thank Shamena Anwar, Abhijit Banerjee, Katherine Barnes, Richard Dieter, Jeffrey Fagan, Nicola Persico, Andrei Shleifer and seminar participants at Duke, Warwick and IIES Stockholm University 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.

tion, 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 those can hardly be interpreted as evidence of racial bias because violent crime rates are higher amongst minorities than amongst whites. Accounting for differences in patterns of offending is crucial if one wants to rigorously test for the presence or absence of discrimination.

Therefore, in the present paper we propose a test of racial bias in capital sentencing that allows for the possibility that members of one racial group are more likely to commit capital crimes than members of another. We develop a model where courts minimize the probability of making judicial errors and we derive a simple test for racial bias. Our test refers only to the first sentencing, namely the capital sentencing of first degree courts and are based upon the assumption that superior courts can only improve upon the accuracy of first sentencing and therefore remove part or all racial bias. That is, we focus upon errors of first degree courts reversed by higher courts.

If we allow for the possibility that racial minorities commit more crimes (and we do not investigate why in this paper), then a simple test concerning errors in judgements against minority defendant is inconclusive and cannot prove racial bias in the court. We then consider the race of the victim.[2] by focusing on pairs of victim/defendant. We develop a test of racial bias based upon the principle 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 had to (painfully) construct a new dataset. We started from the data assembled by Liebman, Fagan and West (2000) but we had to supplement it with a detailed research conducted on a case by case basis. An especially difficult variable to reconstruct was the race of the victim of each case, a variable not available in any usable data set.

Using this data set we find evidence of racial prejudice: first degree courts are more severe (i.e., they tend to give more death sentences which are then reversed) to cases involving a minority defendant killing one or more white victims. This result holds strong for the cases that reach the final stage of revisions and appeal, the Habeas Corpus stage in Federal Courts. For Habeas Corpus cases involving a minority defendant, the error rate was 37.5 if the victim was white, and 27.3 percent if it was not white. The difference, equal to 10.2 percentage points, is statistically significant with a p value of 0.06. For cases involving a white defendant the difference indicates higher reversal rates when the victim is non-white, but it is not statistically 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, about 16 percentage points, 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 only to homicides so the is always at least one well identified victim.

minority defendants with white victims as compared to minority defendants with non-white victims, and it is statistically significant at the 1 per cent level. We do not find evidence of racial bias in other States and we also find evidence of some reduction of racial bias in the South in the eighties relative to the seventies.

Our preliminary results on a subset of Direct Appeal cases, namely the first appeal stage after first degree sentencing, indicate a slightly higher reversal rate on sentences given to minority defendants who killed white victims compared to minority defendants who killed non-white victims (.38 versus .36), although the difference is not statistically significant. This difference between results on Habeas Corpus and Direct appeal may be driven by the non-representativeness of the Direct Appeal sample, or be a "true" feature of the review process. In particular, such difference would arise if higher and higher courts kept being less and less prejudiced against minorities, therefore more and more racially induced mistakes are corrected as the process of revision moves forward. That is, a State court may confirm a racially biased sentence by the first degree court, but the Federal Court may catch the mistake.

Our test is closely 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). The first two papers do not find evidence of racial bias in car searches. Our model is different from these papers in several ways. In particular in those 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 and there is a review of the first decision. The objective function of our courts is therefore different from that of the troopers stopping cars.[3]

We share with Gennaioli and Shleifer (2007) an interest in the effect of bias in judges' decisions. These authors however are interested in a different issue from ours, 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.

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 (see, among others, Erlich (1975), Katz, Levitt and Shustorovich (2003) and Donohue III and Wolfers (2006) for a review). We focus only on the question of whether or not the death penalty is applied with racial biases.

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. Among others, in an important study Gross and Mauro (1984) matched data from the Supplementary Homicide Reports filed by local police with the FBI and from the census of inmates on the death row compiled by the NAACP Legal Defense and Education Fund, for eight states and four

[3]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.

years (1976-80). They constructed an index of aggravating factors and found that, after controlling for them, the race of the victim was still a strong predictor of capital sentencing (the likelihood of a death sentence being higher when the victim was white), but the race of the defendant had no residual explanatory power.[4]

Our study differs from this body of legal studies because it builds a test which is based on the assumption of rational behavior of the court and of the defendant, and in particular on the result that in equilibrium unobserved characteristics of the defendant or of the case should not systematically be correlated with the probability of error. Notice that our test refers to the stage of capital sentencing and not to the decision of carrying out a capital sentence -once affirmed- so that it results in an execution. It is possible that officers vested with the power of clemency (e.g., governors) use this power in a discretionary way and, ceteris paribus, treat minority and white death row inmates differently. This is the test of Argys and Mocan (2004). The two aspects may be related, in that if discrimination is known to occur at the sentencing stage, it is possible that clemency proceedings exercise "reverse discrimination" in response.

The paper is organized as follows. Section 2 offers a very minimal synthesis of the institutional details useful to understand judicial errors in capital cases in the US. Section 3 derives our formal 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.

In this section we briefly review a few historical and institutional points which are relevant for our discussion of the death penalty.[5]

2.1 A Brief History

We can think of 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) which are quite relevant for the issues studied in the present paper. 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.[6] 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 followed the guideline of such code. The number of execution

[4] Abrams, Bertrand and Mullainhatan (20..) taking advantage of the random assignment of cases to judges show that differetn judges are affected differently by the race of the defdnat in Cook County Illinois.

[5] This section largely draws on Coyne and Enzeroth (2006).

[6] Even within the North there were great differences in the aggressiveness with which the death penalty was applied.

nationwide declined partly as a result of political opposition to the death penalty: however executions reached an all timer high during the Great Depression. In the nineteen thirties for several years execution were around 130 per year. The decline in opposition due to the "distraction " caused by the economic difficulties was probably the cause of this pattern. After the Second World war there was a steady decline in executions. The last one 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 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.[8]

2.2 Procedures

Today, thirty six states in the US allow capital punishment.[9] Each state has its own statute but a fair amount of similarity exists among them. Most statutes are in fact modelled around the Georgia one approved by the Supreme Court in the *Gregg vs. Georgia* sentence 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 very similar, the actual application of the death penalty varies greatly across states, ranging from those who use it routinely like in several Southern states to those that almost never use it.

2.2.1 First trial and sentencing

The Supreme court has ruled that no state statute can prescribe mandatory capital punishment, that is no one found guilty of a capital crime can be automatically sentenced to death.[10] 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 an aggravating circumstance and the absence of at least one mitigating factor. 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

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

[8] The Supreme Court however left the door slightly ajar for other horrible crimes like child rape with physical harm to the victim.

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

[10] Woodson ves North Carolina (1976) and Roberst verus Luisiana (1976).

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.[11]

2.2.2 The appeal process

The most important aspect of the capital punishment procedural rules for our study is that all capital sentences, with no exceptions, are automatically appealed in state 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 (Liebman et al. (2000)). 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.[12] 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". When all the appeal processes including the habeas corpus process is over, 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. There is widespread consensus among legal scholars that most (if not all) errors in the first stage trial and sentencing are corrected by the appeal process.

3 The model

3.1 Basic setup

We consider a defendant whom a court can condemn to the death penalty or to a lesser penalty or set free. If the court decides for the death penalty, there is an appeal. In case of a lesser sentence or a no guilt verdict there is no appeal and the decision stands. In appeal, the superior court can either confirm the death penalty or reverse the decision of the lower court because of errors. An error can either 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 (either by charging innocent individuals or giving the death penalty rather than prison terms) higher courts make no mistakes. In fact our empirical tests hold identically under a more general assumption namely that even the highest courts can make mistakes but the latter are uncorrelated with

[11] 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.

[12] Normally the first appeal is to a normal state appellate court, in some state a first appeal would go directly to the State Supreme court. Oklahoma and Texas states which use the death penalty frequently have specialized courts for this first appeal.

the race of the defendant and of the victim. More discussion of this is below but we derive the model under the expositionally simpler;assumption of no mistakes of higher courts.

Defendants are characterized by their race, either White or Minorities and by a set of characteristics of the person or the crime or the relationship between the person and crime. The courts observe several signals over the characteristics of the defendant and of the crime and they summarize them in a single dimension which we label x and we can think of it as the evidence. We normalize the support of $x\epsilon\left[0,1\right]$. The distribution of x is given by $F_g\left(x\right)$ if the defendant is guilty and $F_n\left(x\right)$ if he is not guilty. The signal is therefore informative for the court. The densities $f_g(x)$ and $f_n(x)$ satisfy the strict monotone likelihood ratio property (MLRP), that is, $f_g(x)/f_n(x)$ is strictly increasing in x. Intuitively, this property implies that higher values of the signal x are associated with a relatively higher probability of guilt. We also assume $f_g(x)/f_n(x) \longrightarrow +\infty$ as $x \longrightarrow 1$.

The court chooses to sentence the defendant to death if the evidence is above a certain threshold. Let $r \in \{w, m\}$ be the race of the defendant, where w stands for "white" and m for "minority". Define the threshold x_r^*, which as indicated in the notation could vary with the race of the defendant. Thus the probability than an individual or race r is sentenced to death is: $p_r\left(x\right) = p_r\left(x \geqslant x_r^*\right)$.

3.2 A naive test

3.2.1 The problem of the potential criminal

An individual considering whether to commit a crime or not compares costs and benefits of it. The benefit of the crime and getting away with it is b; think of it as the money stolen from a bank.[13] The cost of being sentenced to death having committed the crime is c_g; the cost of being sentenced to death not having committed the crime is c_n. 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 (sizeable) costs of staying on death row until the first sentence is reversed. All of the above b, c_g and c_n are public information.

For an individual the costs of committing a crime, which may include the moral costs of it, are v and they are drawn from a certain distribution $\Im(v)$ the support of which is $\mathbb{R}^+$. The court knows the distribution but not the realization of v which is known only to the individual. Notice that, for the moment, we are not assuming any difference in this distribution of costs across races. The individual takes x_r^* as given since it is chosen by the court. The expected payoff from the crime for an individual with a certain realization v and a race r is given by:

$$\left[1 - F_g\left(x_r^*\right)\right]\left[-v - c_g\right] + F_g\left(x_r^*\right)\left[b - v\right] \tag{1}$$

The first term represents the cost of being caught, the second term the benefit of getting away with the crime. The expected cost from not committing a crime is

$$\left[1 - F_n\left(x_r^*\right)\right] c_n \tag{2}$$

[13] The death penalty is given only for homicides, so imagine that the robbery includes one or more homicides.

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

$$v_r \leqslant F_g\left(x_r^*\right) b - \left[1 - F_g\left(x_r^*\right)\right] c_g + \left[1 - F_n\left(x_r^*\right)\right] c_n \tag{3}$$

Define the right hand side of (3) as $v^*\left(x_r^*\right)$.This is the threshold of v, which of course depends on x_r^* below which an individual of race r chooses to commit a crime. Define :

$$Prob\left(v \leqslant v_r^*\right) = \Im\left(v_r^*\right) \equiv \pi(x_r^*). \tag{4}$$

This the probability that the realization of v is low enough so that a crime is committed.

3.2.2 The problem of the court

The court wants to deter crime but sentencing 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 in high profile cases, like those involving capital punishment. 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 of sentencing innocent people to death, which might imply a very high moral cost.

These considerations can be summarized by assuming that the court minimizes a weighted average of type I and type II error, that is the probability of condemning an innocent (type I) and the probability of letting a guilty person free (type II). Therefore the court chooses x_r^* as the value of x which minimizes:

$$\min_{x_r} \left\{\alpha_r \left[1 - \pi(x_r)\right]\left[1 - F_n\left(x\right)\right] + \left(1 - \alpha_r\right) F_g\left(x\right)\right\} \tag{5}$$

$$0 \leqslant \alpha_r \leqslant 1$$

The first term represents the type I error, the probability of sentencing an innocent, the second term the type II error, the probability of letting a guilty person free. The parameter α_r, $r = \alpha_m, \alpha_w$ represents the weight given by the court to type I error. We allow this parameter to vary by the race of the defendant and it defines our test of racial bias for the court:

Definition 1 *A court is racially unbiased if and only if* $\alpha_m \neq \alpha_w$.

Note that in case the bias goes against the minorities, then $\alpha_m < \alpha_w$ namely the court is less "worried" about a type I error in the case of a minority defendant.

The first order conditions of the court are:

$$\frac{f_g(x_r^*)}{f_n(x_r^*)} = \frac{\alpha_r}{1-\alpha_r}\frac{1-\pi(x_r^*)}{\pi(x_r^*)} \tag{6}$$

$$r = w, m$$

where $f_g(.)$ and $f_n(.)$ are the density of $F_g(.)$ and $F_n(.)$. Notice that $\frac{\partial x^*}{\partial \alpha_r} > 0$, that is the "standard of proof" increases if the court weighs more the type I error. Notice also that if $\alpha_w > \alpha_m$ then the standard of proof is higher for whites than for minority. This would be the direction of bias against minorities. The equilibrium is given by:

$$\pi(x_r^*) = \Im(v_r^*) \tag{7}$$

$$\frac{f_g(x_r^*)}{f_n(x_r^*)} = \frac{\alpha_r}{1-\alpha_r}\frac{1-\pi(x_r^*)}{\pi(x_r^*)} \tag{8}$$

By Brouwer fixed point theorem the equilibrium exists. It is unique for x_r^* and $\pi(x_r^*)$ and in Appendix we prove uniqueness. Note that if $\alpha_w = \alpha_m$ then $x_w^* = x_m^*$ and $\pi(x_m^*) = \pi(x_w^*)$. That is if there is no racial bias the probability of committing a crime is the same for whites and minorities and the court chooses the same threshold. If, instead, $\alpha_w > \alpha_m$, namely if there is racial bias against minorities the court chooses a higher threshold for whites $x_w^* > x_m^*$. This implies that $\pi(x_w^*) > \pi(x_m^*)$, namely in equilibrium whites will commit more crimes since they internalize the fact that they will be treated more leniently by the courts.

3.2.3 The derivation of a test

Our empirical observations are on the number of reversals of first degree death sentences. We derive a test for $\alpha_w \neq (or =)\ \alpha_m$ based upon only this information.

Define the capital sentencing rate as:

$$\gamma(r) = \pi(x_r^*)[1 - F_g(x_r^*)] + [1 - \pi(x_r^*)][1 - F_n(x_r^*)] \tag{9}$$

The equilibrium error rate $E(r)$ is given by:

$$E(r) = 1 - \frac{\pi(x_r^*)[1 - F_n(x_r^*)]}{\gamma(r)} = 1 - \frac{\pi(x_r^*)[1 - F_g(x_r^*)]}{\pi(x_r^*)[1 - F_g(x_r^*)] + [1 - \pi(x_r^*)][1 - F_n(x_r^*)]} \tag{10}$$

A test for racial bias in this model is therefore as follows:
Test: If $\alpha_w = \alpha_m$ then $x_m^* = x_w^*$ and $E(m) = E(w)$. If $\alpha_w < (>)\alpha_m$ then $x_w^* < (>)x_m^*$ and $E(w) > (<)E(m)$.

Thus in this model we have racial bias in the court if the observed error rate in sentencing of minorities is greater that the error rate of sentencing whites.

3.2.4 Discussion

The critical assumption of this version of the model is that whites and minorities have the same propensity to commit crimes, i.e. they have the same $\Im^m(v) = \Im^w(v)$. If there were no racial bias in the courts the proportion of whites and minorities committing crimes would be the same. With courts biased against minorities the latter would commit fewer crimes because they would internalize the bias of the court. The above observations are counterfactual, since it is well known that crime rates are higher amongst minorities, because of different levels of income, education and other reasons which we do not investigate in this paper.

In particular suppose that the distribution of the costs of committing a crime is different across races, in particular assume that for every $v : \Im^m(v) > \Im^w(v)$.[14] Thus we maintain the assumption that these three parameters, b, c_g and c_n are the same for everyone. Then it follows that $\pi^m(x^*) > \pi^w(x^*)$, that is for any given threshold chosen by the courts more minorities would commit a crime than whites. Thus, even racially unbiased courts would choose different thresholds, namely $x^*_m < x^*_w$. This would amount to statistical discrimination, not racial bias. As a result our test defined above would not capture racial bias.

3.3 A superior test: the race of the victim

In order to derive a more stringent test for racial bias we introduce the race of the victim. Remember that the death sentence applies only to homicides, therefore there is always at least one victim.[15] We allow the cost of committing a crime to differ across the race if the defendant. We indicate with $R \in \{W, M\}$ the race of the victim. The parameters b, c_g, c_n are, as above, independent of the race (of the victim and of the defendant) and public information. The courts now can choose four cutoffs x_{rR} which correspond to the four possible combinations defendant/victim.

Suppose that the distribution of evidence depends upon the race of the defendant, for instance because of the quality of legal assistance. If minority defendants (on average poorer) have a lower quality defense they may face more errors against them in the first trails. While the quality of defense is not explicitly modeled in or framework it could be incorporated in a different distribution of evidence faced by the courts for minority and white defendants. This is because a bad lawyer may not be able to present a good case for the accused. [16]

[14] Note that there could also be differences across races in b, c_g or c_n, but their effect would be isomorphic to differences in $\Im(.)$.We briefly return to this issue in Appendix.

[15] In reality there could be multiple victims of different races. We assume this away in the theory and we address it empirically below. Cases of homicides with multiple victims of different races are a small portion in the data.

[16] Results by Iyengar (2007) indeed suggest that this may be the case. This author study the effectiveness of two types of defense lawyers, provided for indigent defendants, namely public defenders or court private lawyers compensated by the hours. The former perform better, and minority defendants are disproportionately represented by court appointed private lawyers. Let us then model the difference in quality of legal assistance through the use of distribution functions $F^r_g(x)$ and $F^r_n(x)$ which differ according to the race of the defendant: $r = m, w$.

3.3.1 The problem of the potential criminal

The cost realization depends on the race of the individual, but not on that of the victim: $v \sim \Im^r(v)$. There are different types of crimes. In certain types of crimes the defendant cannot choose the victim, and therefore his or her race. Imagine a bank robbery with the killing of guards, who were unknown ex ante to the criminals. 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 which is given. In a third type of crime the defendant can choose the race of the victim, say in a rape with murder. We present the second case (the most common) in the text, and the first and third in the Appendix. We show that our test applies in all these three cases.

Consider then a crime like an homicide of a specific person, say a relative. In this case the criminal cannot choose the race of the victim but he can internalize that the treatment of the court might be different depending on his race and that of the victim. Lets continue to assume that b, c_g and c_n are the same for every crime and known to everybody. Also the distribution of costs cost of committing a crime may depend on the race of the defendant but not on the race of the victim, that is not on the pair criminal/victim. This assumption is critical. Following the same steps as a above from the problem of the potential criminal one can derive a threshold of costs below which a crime of a person or race r kills a person or race R as

$$v(r,R) \leq v^*(x^*_{r,R}) \tag{11}$$

Now we have different threshold of cost of committing crimes for every pair of criminal and victim since the court may choose different threshold of evidence for every pair. Note that now in the decision about whether to commit a crime or not only the threshold of evidence for that crime (i.e. that pair of victim defendant) matter since the pair is known in advance before deciding whether to commit a crime or not. The court maximizes

$$\min_{x_{rR}} \left\{ \alpha_{rR} \left[1 - \pi(x^*_{rR})\right] \left[1 - F^r_n(x_{rR})\right] + (1 - \alpha_{rR}) \left[\pi(x^*_{rR})\right] \left[F^r_g(x_{rR})\right] \right\} \tag{12}$$

Following the same step as above one can derive the equilibrium value for the four threshold of evidence and the probability of committing the four type of crimes. Note that if the court are biased that is $\alpha_{m,W} < \alpha_{m,M}$ then $x^*_{m,W} < x^*_{m,M}$ then potential criminals will internalize that and there will be fewer m, W pairs that m, M pairs. This is consistent with the evidence even though many additional factors explain this fact, such as proximity, family relationships etc. Following the same steps as above one can show that our test based on cross ranking of pairs or criminal and victim remain valid.

3.3.2 The problem of the court

Consider a court that faces a defendant of race r, who has killed a victim of race R. The problem is:

$$\min_{x_{rR}} \left\{ \alpha_{rR} \left[1 - \pi(x^*_{rW}, x^*_{rM})\right] \left[1 - F_n(x_{rR})\right] + (1 - \alpha_{rR}) \left[\pi(x^*_{rW}, x^*_{rM})\right] \left[F_g(x_{rR})\right] \right\} \tag{13}$$

Note that $\pi(.)$ the probability that an individual of race r has committed a crime depends on both the cutoff points chosen by the courts x^*_{rW}, x^*_{rM}, namely the two different cutoff

points on the evidence chosen as a function of the race of the victim. We assume that a court facing, say, a minority defendant takes the cutoff point on a white defendant as given, since it would be chosen in a different trial. Similarly, a court facing a white defendant takes the cutoff point on a minority defendant as given. The two cutoff points are then determined in equilibrium.

Definition 2 *A court is racially unbiased if and only if* $\alpha_{rR} = \alpha$, *for all* r *and* R.

The first order condition (6) would be modified as follows:[17]

$$\frac{f_g^r(x_{rR})}{f_n^r(x_{rR})} = \frac{\alpha_{rR}}{1-\alpha_{rR}} \frac{1-\pi(x_{rW,}x_{rM})}{\pi(x_{rW,}x_{rM})} \tag{14}$$

The equilibrium is given by:

$$\pi\left(x^*_{rW,}x^*_{rM}\right) = \Im^r(v^*(x^*_{rW,}x^*_{rM})) \tag{15}$$

$$\frac{f_g^r(x^*_{rW})}{f_n^r(x^*_{rW})} = \frac{\alpha_{rW}}{1-\alpha_{rW}} \frac{1-\pi(x^*_{rW,}x^*_{rM})}{\pi(x^*_{rW,}x^*_{rM})} \tag{16}$$

$$\frac{f_g^r(x^*_{rM})}{f_n^r(x^*_{rM})} = \frac{\alpha_{rM}}{1-\alpha_{rM}} \frac{1-\pi(x^*_{rW,}x^*_{rM})}{\pi(x^*_{rW,}x^*_{rM})} \tag{17}$$

By Brouwer fixed point theorem an equilibrium exists. The proof of uniqueness follows the same logic as the one given in Appendix for the previous version of the model and is omitted. Note that even if $\alpha_{rR} = \alpha$ for any r, R the cutoff point chosen by the court could be, and in general will be different for W and M, because the distribution of costs of committing a crime is different.

The court will choose $x^*_M < x^*_W$ since the likelihood that a minority defendant is guilty is higher, ceteris paribus, as the cost of committing crimes is lower for minorities. And, in fact, because of that reason, in equilibrium minorities commit more crimes.

Let us now derive a test for racial bias. To save on notation we indicate with π_{rR} the probability of guilt for every pair. We can then define the capital sentencing rate for r, R pairs as:

$$\gamma(r,R) = \pi_{rR}\left[1-F_g^r(x^*_{rR})\right] + (1-\pi_{rR})\left[1-F_n^r(x^*_{rR})\right] \tag{18}$$

The equilibrium error rate $E(r,R)$ is given by:

$$E(r,R) = 1 - \frac{\pi_{rR}\left[1-F_g(x^*_{rR})\right]}{\gamma(r,R)} \tag{19}$$

Result: If $\alpha_{rM} = \alpha_{rW}$ then $\gamma(r,B) = \gamma(r,W)$ and $E(r,M) = E(r,W)$.

[17] The same considerations apply if the distribution function on the evidence depend upon the race of the victim but not on that of the defendant. In that case, the left hand side of (14) would be $f_g^R(x_{rR})/f_n^R(x_{rR})$, but the format of our tests would not change.

This result shows that if the court does not discriminate over the race of the victim, the error rate should be independent of the race of the victim, even though it could depend upon the race of the defendant. We now provide several possible tests of bias against minorities. Obviously for a bias against whites the inequalities should be reversed.

Proposition 1 *If a court does not exhibit racial bias, then the ranking of average error rates* $E(r, W)$ *and* $E(r, M)$ *should not depend on* r *the race of the defendant.*

Proof. By inspection of equation (19). ■

In other words we posit that if the court is racially unbiased the following condition must hold:

$$E(m, W) > E(m, M) \iff E(w, W) > E(w, M) \tag{20}$$

Intuitively, test (20) says that if we find a higher error rate on minority defendants who killed white victims, compared to minority defendants who killed minority victims, absence of racial bias on behalf of the court is consistent with finding a higher error rate on white defendants who killed white victims than on white defendants who killed minority victims. 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 other words, for each victim's race the ranking of error rates across defendants' race must be the same. Failure to satisfy this condition implies the presence of racial bias in our model. [18]

3.3.3 Discussion and extensions

Bias in the collection of evidence

Our test would not be valid if the functions $F_n^r(x)$ and $F_g^r(x)$ were allowed to depend both on the race of the defendant and the race of the victim. In other words, we cannot allow minority defendants to have a systematically worse legal defense relative to white defendants when the case involves a white victim than when it involves a minority victim. However, if we were to allow dependence of $F_n^r(x)$ and $F_g^r(x)$ on R as well as r, it is unclear in which direction the bias would go. Suppose for instance that the dependence on the evidence function on race pairs is due to police work. This may occur, for instance, if the police work is biased, depending upon race. On the one hand the police may work harder to bring in court proofs against minority defendants paired with white victims, and in that case it would be easier for the courts to "separate" guilt and innocence for minorities than for whites, as particular effort has been put in collecting as much evidence as possible. Note that then the proportion of errors would be lower for minorities, due to this effect. This bias would "work against" finding more errors on minorities. On the other hand a biased police may "fabricate" cases against minorities when the victim is white. In this case the bias would go

[18] This is the test we perform using our data below. Notice that our use of rank order tests are reminiscent of Anwar and Fang (2006), with the difference that in their case one of the two dimensions over which 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.

in the opposite direction of finding more errors in the case of minority defendants and white victims paired. Ex ante it is unclear in which direction the bias would go.[19]

Plea Bargain

Many potentially capital crime 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 persecution to pursue or not a plea bargaining (Ehrard, 2009). Comprehensive empirical studies of the nature and characteristics of plea bargaining agreements are hard to come by due to data limitations. With particular reference to the race of the defendants, the only result we could identify was that minority defendants receive a harsher plea bargained prison term, though this result applies not only (or not especially) to capital cases (Humphrey and Gogarth, 1987). The same study also finds that plea bargained terms are generally higher for those with previous criminal history, which is more likely to occur for minority defendants.

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 pair defendant/victim, our test would be unaffected. If it were not, then this correlation might introduce a bias, but it is quite hard to develop an intuition for the direction of it, as the direction of the bias would depend among other things on the shape of the distribution of the signal.

To assess the potential relevance of this source of bias from an empirical point of view, we used data on a representative sample of murders adjudicated in 1988 in 33 of the largest counties in the US.[20] This dataset includes information on race of the defendant and of the victim, as well as the final disposition outcome of the case (among which guilty plea). When we regress the likelihood of guilty plea on race of the defendant, race of the victim, and the interaction of the two, the coefficient on the interaction is not statistically different from zero.[21]

Bias in decisions of superior courts

One maintained hypothesis in our model is that superior courts never make mistakes and are racially unbiased which may or may not be true. If superior courts made errors which were unrelated to the race of the defendant or the victim this would just be random noise

[19] Radelet and Pierce (1983) analyzed a sample of 1,400 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.

[20] 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.

[21] 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.

in the data. By the law of large numbers these mistakes would not influence our tests or racial bias. Suppose instead that superior courts were racially biased in the same direction of lower courts. This would go against finding racially biased errors. That is, if we did not find a 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 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 stronger in later stages of appeal (e.g., Habeas Corpus) than in earlier ones (e.g., Direct Appeal). This is what we find below.

What we cannot allow 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).

Notice also that if one were to model the behavior of higher courts incorporating the assumption that higher courts do not receive a perfect signal but a noisy signal, as long as it is known that lower courts are biased in a given direction, it may be socially optimal to pick a parameter α biased in the opposite direction for higher courts to counteract the errors made in the first stage. In this case the assumption we would need for our test to remain valid is that the extent of bias of higher courts, even if in the opposite direction, does not exceed that of lower courts.

4 Construction of the dataset

To implement the tests described above we could not rely on any readily available dataset. In fact all existing datasets containing information on the race of the defendant and of the victim in capital cases have limited geographical and temporal coverage and – most importantly for our purposes – do not contain information on whether the capital sentence was reaffirmed or not in appeal.

On the other hand, 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.'s (2000) report. In what follows we start by briefly reviewing the characteristics and scope of FL's data, then discuss our search methodology and then 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. It covers three sets of cases:

- DIRECT APPEAL dataset: 4,546 state capital cases whose direct appeal decisions became final between January 1, 1973 and December 31, 1995.[22] This is the universe of all capital sentences that were reviewed on direct appeal by a state high court.

- POST-CONVICTION dataset: 248 state post-conviction reversals, which is a subsample of the state post-conviction appeals that resulted in a reversal; and

- HABEAS CORPUS dataset: 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.

Because the post-conviction dataset is an incomplete and selected set of cases (due to the fact that state post-conviction decisions are often not published), in our analysis we shall only employ the direct appeal (from now on DA) dataset and the habeas corpus (from now on HC) dataset, which comprise instead the universe of available cases at those stages.

For 84 of the 4,546 cases in the DA dataset and 13 of the 557 cases in HC, the sentence indicated in FL's data could not be found in Lexis-Nexis, hence we drop those cases. Also, because some observations in the datasets correspond to multiple sentences for the same crime and we want to record error once for each crime, we use one observation per defendant/crime and attribute an error if it was found in at least one stage of the appeal process.

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.[23]

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, plus a

[22] "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).

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

number of websites specialized in death penalty issues.[24] Unfortunately, despite the large amount of time invested in this task, we could not find information on defendant's and victim's race for every case in our dataset.

[Table 1 about here]

Table 1 reports a tabulation of cases with missing information on the race of the defendant or/and on the race of the 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 23 cases out of 531, that is 4 percent of the sample. In the DA data, we have information on defendants' race for 3,543 cases out of 3,782 (94 percent of the sample), and on victims' race for 3,041 cases out of 3,782 (80 percent). Our ability to recover information on victims' race has been lower in DA cases, due both to the size of the overall data and to the fact that these cases received relatively low media coverage compared to HC ones.

[Table 2 about here]

In table 2A we try to gauge the extent of selection in the pattern of missingness of victims' race (the binding constraint in the HC data) for all cases in which we have information on the defendant's race. We report the means of several variables related to defendant, victim, and crime characteristics for the subsample in which we have information on victims' race (column 1) and the cases in which we don't (column 2). We conduct a t-test for the equality of these means and report the p-values in column 3.

For none of the demographic characteristics of the defendant (race, gender, age) we find a significant difference across the two subsamples. Characteristics like prior felony convictions, history of drug or alcohol abuse, and deprived family background are also balanced across the two subsamples. Similar results obtain for 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 approaches statistical significance is the gender of the victim, but we show below that our results are robust to excluding female victims.

Turning to crime characteristics, we have two pieces of information. 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. Again, we would expect harsher punishment for the latter type of crime. We do not find a statistically significant difference in means across subsamples for either of these measures.

These observations coupled with the small number of missing data points 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. Contrary to our previous results, in this data the pattern of missingness is mixed. Characteristics like gender and race of the defendant are perfectly balanced, and also for the number of victims and the likelihood

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

that the victim is the partner, there is no significant difference across subsamples. On the other hand, the likelihood that the victim is female or is a public official is significantly higher when victim race is available, again likely due to media coverage. In this subsample, there is also a 3 percentage points higher likelihood that the defendant knew the victim.

We cannot know what is the source of the difference in observables that we find in the DA dataset. But given that this pattern may undermine the validity of our results, in most of our analysis we will confine ourselves to the HC dataset. The HC dataset also has the advantage of containing much more information on each case. Nonetheless, we shall also report the main results for DA cases, which however will have to be interpreted with caution.

Appendix tables A1 and A2 contain summary statistics on the share of missing observations by state and by year in the HC and DA datasets, respectively.

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 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, 84 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).

In Panel B of table 3, similar figures are reported for the DA dataset, i.e. the subsample for which victim's race is not missing. The error rate (relief) in the first stage of appeal is 37 percent. The shares of white and African American defendants are similar to the HC dataset, at .51 and .41, respectively. The shares of white and African American victims are .78 and .18, respectively. Finally, in this sample the share of non-white defendants who killed a white victim is .29.

[Figure 1 about here]

In figure 1 we examine the time pattern of error rates, plotting the mean of the "relief" variable by year, both for the HC and the DA dataset.[25] Error rates are declining over time. Under the interpretation we propose below, namely that error rates are correlated with racial bias, this pattern would suggest a decline in the amount of racial bias over time.

[25] Note that the HC data end in 1986 that is nine years before the coverage of our sample because it takes that long for cases to reach this stage of the appeal process. For a recent study on the time needed to reach the HC stage see King, Cheesman and Ostrom (2008).

5 Empirical methodology and results

The test for racial bias we derived in the theoretical section required that 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, $R = W, M$. 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) using the Z-statistic:

$$Z = \frac{\widehat{E(r,W)} - \widehat{E(r,M)}}{\sqrt{\frac{SVar_W}{n_W} + \frac{SVar_M}{n_M}}} \sim N(0,1) \tag{21}$$

where $SVar_R$ is the sample variance of the error dummy in the cases involving a defendant of race r and a victim of race R; and n_R is the number of cases involving a defendant of race r and a victim of race R; and $R = W, M$.

5.1 Main results, Habeas Corpus dataset

[Table 4 about here]

Table 4 contains the outcome of our test (21) using the HC data, and the main result of the paper. Each cell reports the average probability of error 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 (21). 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) is true.

The first row of table 4 shows that in cases involving a white defendant the average error rate is 35.7 percent if the victim is white and 47.1percent if it is non-white, with a difference of -11.4 percentage points.[26] On the other hand, in cases involving a minority defendant, the error rate is 37.5 if the victim is white, and 27.3 percent if it is not white, with a difference of 10.2 percentage points. 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 .06; for cases involving white defendants we fail to reject the null against the alternative (p-value .19). Based on our rank order test, we therefore reject the hypothesis of no racial bias on behalf of trial courts.

[Tables 5 and 6 about here]

In table 5 we examine whether the pattern of results differs depending on whether we consider the earlier cases in the sample (Panel A: sentences imposed between 1973 and 1979)

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

or later on (Panel B: sentences imposed from 1980 to 1990).[27] This table shows a difference for sentences before and after 1980. In the first decade we find results very similar to those in Table 4. In the following decade the pattern is similar (more error for minorities killing whites) but the difference is smaller and not statistically significant.

In table 6 we repeat a similar exercise but look at differences across regions rather than over time. When we consider sentences imposed in the South (Panel A) a striking differences emerges between minority defendants and white ones. This table shows that comparing minorities killing whites the differences in errors is large and statistically significant in the South (15.4 per cent, p-value 0.01)

When we conduct analogous tests for all other regions (Panel B of table 6) we fail to reject the null of absence of racial bias. Conditioning on the race of the defendant, the error rate is higher for non-white than for white victims both if the defendant is white and if he is not. Neither of these differences, though, is statistically different from zero at conventional levels. One caveat about Panel B, however, is that it covers a substantially smaller number of cases compared to Panel A.

Therefore these result viewed together suggest that our evidence of racial bias is strong especially in the South in the first half of the sample. This is consistent of racial issues being more salient in the South and of an increase in awareness against racial bias in more recent decades.

5.2 Interpretation and robustness

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.

Legal assistance

One of these characteristics is the quality of legal assistance. 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. In table 7 we repeat our tests restricting the sample to cases that are "relatively similar" in terms of some trial characteristics.

[Table 7 about here]

For example, being represented by an out of state attorney at the appeal stage is likely to be an indication of having a good defense because out of state attorneys are typically from prominent law firms that take the case pro bono, and have more resources at their disposal in terms of investigators, mitigation specialists and defense team in general. In fact being

[27] The criterion for splitting the years in this way was to try and keep a similar number of observations in both subsamples, although exact balancing was not possible due to the high number of sentences imposed in 1980. Sentences imposed after 1990 do not appear in the HC dataset, as at the time the FL data was collected the appeal process had not reached the Habeas Corpus stage.

represented by an out of state attorney increases the likelihood of obtaining relief in our data. In the first part of the table we restrict the sample to cases that were *not* represented by out of state attorneys. The results are very similar to those obtained in table 4 using the full sample. Minority defendants have about a 10 percentage points higher probability of sentence reversal when their victim was white than when it was not (p-value .11), while no significant difference is found for white defendants.

Similar conclusions can be drawn if we use as a proxy for the quality of legal assistance at the trial stage the fact that "ineffective assistance of counsel" in the guilt and sentencing phase was later included among the claims for relief.[28] We start by restricting the sample to 379 Habeas Corpus cases for which we have information on victim race and 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 rate for minority defendants who killed a white and a non-white victim is 14.5 percentage points (as opposed to 10.2 in the full sample), and the null of no difference is rejected against the alternative of a positive difference with a p-value of .02. If we further restrict the sample to the 218 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 lead us to underestimate the extent of bias.

In the remaining parts of table 7, we consider the subset of cases in which "prosecutorial suppression or withholding of evidence or other prosecutorial misconduct" was not raised among the claims (348 cases), nor was "improper interrogation", that is, there was no involuntary confession or guilty plea or request for attorney denied (454 cases). In both subsamples 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.

Two observations are worth making. First, 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 totally objective. Second, one may argue that even if our results were entirely attributable to the quality of legal assistance -which based on our evidence they are not- this would not prove the absence of bias, but rather give insights on one possible source of bias. Whether the bias results from discriminatory preferences on behalf of the court or from discriminatory access to legal representation is a conceptual difference that our model makes, but it is not obvious that the two can be entirely told apart in the empirics. 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.

Crime characteristics

[Table 8 about here]

Another possible interpretation of our results is that cases involving minority defendants and white victims may differ from other cases for the type of crime involved, and that the difference in error rates reflects characteristics of the crime rather than racial bias. In table

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

8 we address this concern by conditioning our analysis on a set of observable crime characteristics. 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 this should increase our confidence that the findings are not entirely spurious.

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 first panel of table 8 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 Fagan and Liebman's dataset. These cases should be relatively comparable along this dimension. Our results show that among these *376* cases the likelihood of error is *16.3* percentage points higher for minority defendants who killed white victims compared to minority defendants whose victims were not white (p-value .02). The difference in error rates is not statistically significant for white defendants.

An aggravating factor that may be responsible for the difference we find is the fact that the defendant killed a police, or fireman, or guard, or other public official. 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 essentially the same results: no significant difference in error rates for white defendants, and a difference of 14.1 percentage points for minority defendants (p-value .02).

Another aggravating factor might be the presence of multiple victims. Restricting the analysis to homicides with only one victim shows again a significant difference of 10.2 percentage points for minority defendants (p-value .06), and no significant difference for white defendants.

Finally, to test whether the gender of the victim is a significant factor in our results, in the last panel of table 8 we restrict the attention to cases in which none of the victims was female. Again, we find higher error on minority defendants who killed white men than on those who killed non-white men (the difference is about 13 percentage points, and the p-value is .07). The corresponding difference for white defendants is not significant. Notice that "female victim" was the only observable characteristics for which we had found a difference close to significant (p-value .12) in Table 1 between the sample where victim race was known, and the sample where it was missing. The fact that our results hold –and are actually stronger– for the subset of cases where the victim is not a woman is therefore reassuring.

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, we consider murders committed during a robbery (183 cases fall in this category). We find that 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 (p-value .04). The difference for white defendants is not statistically significant.

Finally, when we restrict the sample to cases that are "similar" in the sense of being classified as "felony murders" (last panel of Table 8), again we find a close to 13 percentage points higher error rate for non-white defendants who killed white victims (p-value .07), and no corresponding difference for white defendants.

To sum up, in all cases listed in table 8 our rank order test rejects the null of absence of racial bias according to our model. To the extent that the observable crime characteristics

we control for in this table may be correlated with other unobservable characteristics of the crime, this suggests that our findings are not driven by spurious correlation between pairs of defendant/victim races and severity of the crime.

Reasons for relief

Relief can be given for a variety of reasons from mental illness of the defendant to police misconduct to improper use of testimonies. We have investigated whether there are certain statistically significant patterns of reasons for relief for different combinations of races of victim or defendant. We did not find any particular pattern to report. Results are available.

Possible bias of appeal courts

An important assumption of our model is that the review by the appeal court is unbiased, i.e. that a judicial error is found if and only if there is one. As we mentioned above, 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 our methodology would come 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. This type of bias could result, for example, from an attempt to "correct" for biases in previous stages. The first thing to notice is that, as long as such bias is related only to defendant characteristics (e.g., to compensate for the poor quality of legal assistance available to minorities), our test is still valid. 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 this possibility is something we cannot rule out, we try to shed light on 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 be more likely be found among liberal judges than among conservative ones. For each sentence, we therefore 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. In particular, 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 9 about here]

Table 9 reports the results for the subset of cases where the majority (Panel A) or the totality (Panel B) 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, significant at the 10 percent level in panel B (p-value .13 in panel A). The magnitude of the difference in error rates is 9 percentage points when we consider appeal courts where a majority of the judges are Republican (Panel A of Table 9), and 13 percentage points when we restrict our test to courts that are entirely composed of republican-appointed judges (Panel B).

In Panel C we consider the possibility that the political climate in a given year may affect relief rates, and restrict the sample to final appeal sentences that occurred under a Republican administration. We find a difference of 18.4 percentage points for non-white defendants who killed white victims, and this difference is significant at the 1 percent level.

Overall, we interpret the evidence in table 9 as suggestive that our findings are unlikely to be driven by "reverse bias" of appeal courts. We next try to shed light on the role of trial courts, as opposed to appeal courts.

Role of the jury

A first issue relates to the role of the jury versus the judge in determining the sentence. During our sample period, several capital sentencing states admitted judicial findings of fact at the sentencing stage, implying that the judge could determine the presence of aggravating factors and impose a death sentence, possibly overriding the jury's recommendation. This was later declared unconstitutional by the US Supreme Court (Ring v. Arizona, 536 U.S. 584, 2002). One can thus gather some (limited) insights on the role of the jury versus the judge in making errors on certain combinations of defendant/victim races by repeating our empirical test separately for cases decided under systems where both guilt and sentence were exclusively determined by the jury, and systems where the sentencing procedure provided for judicial fact finding.[29]

[Table 10 about here]

Panel A of Table 10 reports the results of our rank order test for cases decided under judicial fact finding or hybrid systems, while Panel B contains results for cases where the jury was the sole responsible for fact finding at the sentencing stage. In both cases, the qualitative pattern of the results indicates higher error on cases with mixed racial combinations. In particular, in Panel A death sentences for non-whites who killed whites have a 16.5 percentage points higher likelihood of being reversed compared to those for non-whites who killed non-whites. In Panel B the corresponding difference is 8.5 percentage points. Neither of the two differences, however, is significant at conventional levels (p-values .12 and .13, respectively), possibly due to the smaller sample size, especially for Panel A.

Trial judges' incentives

Finally, to understand the incentives of trial court judges and the possibility that reelection concerns may underlie some of the judicial errors we observe, we consider differences across jurisdictions in the way in which judges are reappointed. In a recent paper using data from Kansas, Lim (2008) finds significant differences in sentencing behavior of elected versus appointed judges, with elected judges in conservative districts leaning towards relatively harsher sentences, possibly to increase their reelection probability.

Building on this contribution, we gathered information on the method of reappointment of judges in each jurisdiction of our dataset. Four possible categories emerged: (i) competitive election; (ii) retention election (up or down majority decision by voters); (iii) decision by the state legislature; and (iv) gubernatorial reappointment. We then matched each record in our dataset to the system in place at the time and in the jurisdiction where the death sentence was imposed.

[29] In the latter group we include both states where the judge had the sole authority of establishing the existence of aggravating factors, and hybrid systems where the jury's recommendation regarding the sentence was an advisory one and could be overruled. For the period under consideration, this gives us the following states: Arizona, Colorado, Idaho, Montana, Nebraska, Alabama, Delaware, Indiana and Florida.

For our purposes, the main distinction is between the competitive system (i) and the remaining ones, as the first is the system in which the judge is most sensitive to pleasing his/her constituency. Table 11 therefore reports results for the subsample of cases in which the trial judge would face reelection concerns (Panel A) and the remaining cases (Panel B).

[Table 11 about here]

What emerges from the table is that our evidence of racial bias holds in the sample of cases where judges have reelection concerns but not in the remaining cases. Only in the former we find a significant difference in error rates for nonwhite defendants depending on the race of the victim. Two important caveats apply, though. First, different from Lim (2008) who holds constant most institutional variation by focusing on one state, in our data different reappointment systems overlap with other institutional differences across states and jurisdictions. Second, the inference we can make based on Panel B is very limited, due to the very low number of observations.

5.3 The first stage of appeal, Direct Appeal

Our final body of results relates to the very first stage of appeal, i.e. the one involving State Courts and that is referred to as Direct Appeal cases (see above). We present our results in Table 12, first for the full sample (Panel A) and then separately for Southern states (Panel B).

[Table 12 about here]

In the full sample, the error rate on white defendants is 38 percent regardless of the victim's race; the error rate on nonwhite defendants is 37.8 percent when the victim is white and 35.6 percent when it is not white, though the difference is not significant (p-value .19). When we focus on the South, murders across races seem to display 3.7 percentage points higher error rates, but although the direction of the bias is consistent with our findings for Habeas Corpus cases, the results are not significant at conventional levels (p-value of .13 for non-white defendants).

One possible interpretation of the above is that courts ruling on Direct Appeal maintain some degree of racial bias in the same direction as first degree courts. This would bias our results against finding evidence of racial discrimination in capital sentencing.[30] Another consideration to keep in mind is that the sample for which we have data on victim race may not be representative of all DA cases, as some differences on observable characteristics emerged in Table 2 above.

6 Conclusions

Testing for racial discrimination in capital punishment is quite difficult in general if different groups have different crime rates. Obviously we could not simply look at the ratio of minority

[30] An extreme case is that of Virginia. As noted by Liebman et al. (2000), this state has by far the lowest error detection rate among capital states, and about 1/3 of the national average. When we checked our result leaving Virginia out of the sample we moved closer to statistical significance of difference in error rates consistent with racial bias. (Results are available).

versus white guilty defendant in capital cases given that crime rates are higher for minorities. We used the share of judicial errors in first degree sentencing as an indicator of racial bias of such courts. Our maintained assumption was that superior courts (especially the Federal Court dealing with Habeas Corpus final appeal stage) have less racial prejudice bias or no prejudice at all. Note that, if they had, that would bias our result against finding any bias in he first sentencing. We uncovered 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 other combinations. This racial bias is driven by Southern States and is very strong in the case of Habeas Corpus cases, where the results are statistically significant. We also find that this racial bias was stronger in the earlier part of our sample covering the seventies relative to the eighties and early nineties. While the results on the Habeas Corpus stage show robust evidence of a sizeable racial bias, we do not find strong evidence of bias in Direct Appeal cases. The incompleteness of our data coverage of Direct Appeal cases should caution us about over interpreting these results. The difference between Direct Appeal results and the Habeas Corpus one can also be explained by the fact the racial bias is corrected more and more at higher stages of the review process.

7 Appendix

7.1 Proof of uniqueness of equilibrium given by equation 8.

(The proof follows Anwar and Fang (2006))

Define with $\widehat{x}_R^*$ the level of evidence of which equation 3 holds with equality. Note that $v(0) = c_n - c_g < 0$, and that

$$\frac{\partial v^*(x^*)}{\partial x^*} = f_g(x_r^*)\left[b + c\right] - f_n(x_r^*)c_n \tag{22}$$

By MLRP there exists a $\widehat{x}^* \epsilon\, [0,1]$ such that $v^*(\widehat{x}^*)$ is strictly increasing in $\widehat{x}^*$ when $x^* > \widehat{x}^*$. $v(0) < 0$ and $\pi(x^*) = 0$ for $x^* < \widehat{x}^*$. The equilibrium (or equilibria) must imply $x^* > \widehat{x}^*$. In fact in the region $x^* < \widehat{x}^*$, $\pi(x^*) = 0$ and the court would set $x^* = 1 > \widehat{x}^*$ leading to a contradiction. Suppose now that we have two equilibria, x^* and $\overline{x}^*$ such that $x^* > \overline{x}^* > \widehat{x}^*$. Thus, both equilibria are in the region in which $v(x_R^*)$ is increasing in x_R^*. Since $x^* > \overline{x}^*$ than $v(x^*) > v(\overline{x}^*)$ but then $\pi(x^*) > \pi(\overline{x}^*)$. But then the court would choose $x^* < \overline{x}^*$ which leads to a contradiction.

7.2 The criminal cannot choose the race of the victim.

Consider the case in which the defendant cannot choose the race of the victim because like in a robbery somebody may get killed in ways which were not planned or predicted by the criminal. Define β as the probability that the victim of the crime is white, and $(1-\beta)$ that of killing a black victim. The expected payoff from the crime is:

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

$$+(1-\beta)\left\{\left[1-F_g\left(x_{rM}^*\right)\right]\left[-v(r)-c_g\right]+F_g\left(x_{rM}^*\right)\left[b-v(r)\right]\right\} \tag{23}$$

The cost of not committing a crime is:

$$\beta\left[1-F_n\left(x_{rW}^*\right)\right]c_n+(1-\beta)\left[1-F_n\left(x_{rM}^*\right)\right]c_n \tag{24}$$

Following the same procedure as above, we define a threshold level of $v(r)$ below which a crime is committed:

$$v(r) \leqslant v^*(x_{rW}^*, x_{rM}^*)$$

The derivation and the explicit formula for $v^*(.)$ is in Appendix. Notice that the threshold cost $v(r)$ is different for different races of the defendant and it depends on the (possibly different) cutoff on evidence imposed by the court for minorities or white victims. Obviously, $v(.)$ 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.

7.3 Choosing the race of the victim

Consider now the final possibility namely the case in which the criminal can choose the race of the victim. If the benefits of a white or minority victim where the same, as we have assumed thus far, and if the courts chose a higher threshold of evidence for minority victims, all potential criminals would choose minority victims. With no discrimination of the courts (i.e. with the same threshold) potential criminals would be indifferent on the race of the victim and would randomize. Our test based upon differences in the victim's race would apply to thew first two types of murder in which the race of the victim could vary.

Suppose instead that $b_w > b_m$ where b_r are the benefits of choosing a victim of race r. In order to observe that both whites and minorities are chosen as victims it has to be the case that extra benefit of choosing a white victim more than compensates in equilibrium for the fact that $x_{r,W}^* < x_{r,M}^*$. In this case wed have an equilibrium with both white and minority victims. As above we would have four thresholds of evidence chosen by the court, and the threshold condition would be:

$$v(r,R) \leq v^*(x_{r,R}^*) \tag{25}$$

This case is similar to the one discussed above with one difference. Now a potential criminal has three rather than two choices: committing no crimes, choosing a white or choosing a minority victim. The choice would depend on the realization of v.and the benefits $b_w > b_m$ and of course the thresholds of evidence picked by the court. Consider a court facing a crime with a certain victim/defendant race combination. The court would choose the appropriate threshold $x_{r,R}^*$. This is the threshold that the potential criminal would consider in his decision of whether or not to commit the crime against that victim race, not committing the crime or committing a crime against another victim of a different race. The basic form of our test would then go through.

References

[1] Abrams D. M. Bertrand and S. Mullainathan (20""), "Do Judges Vary in their Treatment of Race?", unpublished.

[2] Antonovics, Kate and Brian 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, Shamena and Hamming Fang (2006), "An Alternative Test of Racial Profiling in Motor Vehicle Searches: Theory and Evidence", *American Economic Review*, 96(1), 127-151.

[4] Argys, Laura and Naci Mocan (2004), "Who Shall Live and Who Shall Die? An Analysis of Prisoners on Death Row in the United States", *Journal of Legal Studies*, 33(2), pp. 255-92.

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

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

[7] Becker Gary (1968), "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76, 129-217.

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

[9] Donohue III, John and Justin Wolfers (2006), "Uses and Abuses of Empirical Evidence in the Death Penalty Debate", *Stanford Law Review*, pp.

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

[11] Erhard, Susan (2008) "Plea Bargain and the Death Penalty:an exploratory study", *Justice System Journal.*

[12] Fagan, Jeffrey, and James 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 Gene and Michael Katz (1983) "Plea Bargain and Soclai Welfare" *American Economic Review, 73 (4) 749-57*

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

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

[18] King N., F Cheesman and B. Ostrom (2008), "Habeas Litigation in US Distrcit Courts", Vanderbilt University Law School.

[19] Knowles John, Nicola Persico and Petra Todd (2001), "Racial Bias in motor vehicles searches:Theory and Evidence," *Journal of Political Economy*, 109, 203-29.

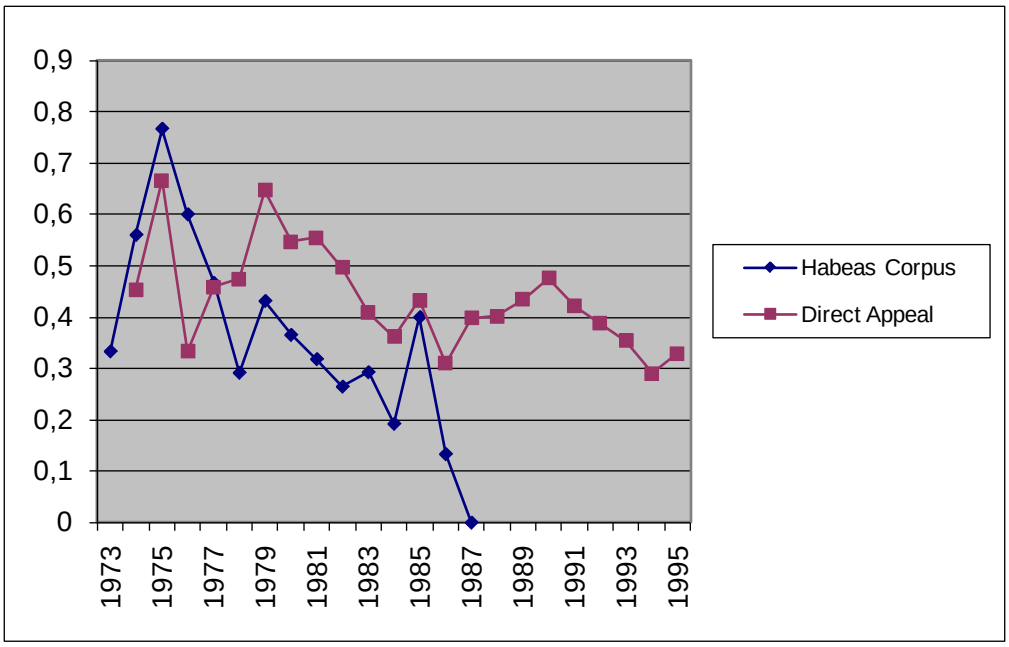
[20] Iyengar, Radha (2007), "An Analysis of the Performance of Fedral Indigent Defense Counsel", NBER Working Paper, n. 13187

[21] Liebman, James, Jeffrey Fagan and Valerie West (2000), *A Broken System: Error Rates in Capital Cases, 1973-1995*; Columbia University.

[22] Lim, Clare (2008), "Turnover and Accountability of Appointed and Elected Judges", unpublished, Stanford University.

[23] Reinganum, Jennifer (1988), "Plea Bargain and Prosecutoral Discretion", *American Economic Review, 78 (4), 713-27.*

Figure 1: Error rates by year of first sentence



Source: Authors' calculations on Fagan and Liebman's (2002) dataset.

Table 1: Missingness of race

Panel A: Habeas Corpus

		Missing victim's race		
		No	*Yes*	*Total*
Missing defendant's race	*No*	505	23	528
	Yes	3	0	3
	Total	508	23	531

Panel B: Direct Appeal

		Missing victim's race		
		No	*Yes*	*Total*
Missing defendant's race	*No*	2,964	579	3,543
	Yes	77	162	239
	Total	3,041	741	3782

Table 2: Selection in missingness of victim's race

Panel A: Habeas Corpus

Variable	*Nonmissing victim's race*	*Missing victim's race*	*Diff=0 (p-val)*
Defendant characteristics			
Defendant is White	0.51	0.48	0.76
Defendant is African American	0.44	0.48	0.73
Male defendant	0.99	1.00	0.63
Age of defendant	28.80	41.50	0.14
Prior felony	0.22	0.26	0.67
History of alcohol abuse	0.12	0.04	0.25
History of drug abuse	0.16	0.13	0.68
Deprived/Abused background	0.02	0.00	0.55
Victim characteristics			
Number of victims	1.40	1.29	0.76
Female victim	0.48	0.30	0.12
High status victim	0.23	0.30	0.39
Police victim	0.09	0.04	0.44
Crime characteristics			
Defendant knew victim	0.26	0.26	0.99
Heinous crime	0.40	0.30	0.36

Panel B: Direct Appeal

Variable	*Nonmissing victim's race*	*Missing victim's race*	*Diff=0 (p-val)*
Defendant characteristics			
Defendant is White	0.51	0.51	0.87
Defendant is African American	0.41	0.41	0.93
Male defendant	0.98	0.98	0.61
Age of defendant	30.99	32.43	0.00
Victim characteristics			
Number of victims	1.38	1.35	0.45
Female victim	0.53	0.43	0.00
Police victim	0.06	0.02	0.00
Partner victim	0.05	0.04	0.22
Crime characteristics			
Defendant knew victim	0.11	0.06	0.00

Table 3: Summary statistics

Panel A: Habeas Corpus

	Sample with non-missing victim's race		
Variable	*No. Obs*	*Mean*	*Std. Dev.*
Any relief	505	0.36	0.48
African American defendant	505	0.44	0.50
White defendant	505	0.51	0.50
African American victim	505	0.13	0.34
White victim	505	0.84	0.37
White def., Non-white vict.	505	0.03	0.18
Non-white def., White vict.	505	0.36	0.48
White def., White vict.	505	0.48	0.50
Non-white def., Non-white vict.	505	0.13	0.34

Panel B: Direct Appeal

	Sample with non-missing victim's race		
Variable	*No. Obs*	*Mean*	*Std. Dev.*
Relief first appeal	2898	0.37	0.48
African American defendant	2898	0.41	0.49
White defendant	2898	0.51	0.50
African American victim	2898	0.18	0.38
White victim	2898	0.78	0.42
White def., Non-white vict.	2898	0.03	0.18
Non-white def., White vict.	2898	0.29	0.46
White def., White vict.	2898	0.48	0.50
Non-white def., Non-white vict.	2898	0.19	0.39

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

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.357	0.471	0.19	258
	(0.031)	(0.125)		
Non-white	0.375	0.273	0.06	250
	(0.036)	(0.055)		
N.obs	425	83		

Note: Standard errors of the means in parenthesis

Table 5: Error rates by race and decade

Panel A: Years 1973-1979

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.46	(a)		100
	(0.050)	(a)		
Non-white	0.477	0.318	0.08	108
	(0.054)	(0.102)		
N.obs	186	22		

Panel B: Years 1980-1990

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.284	0.400	0.20	156
	(0.038)	(0.131)		
Non-white	0.286	0.250	0.33	142
	(0.046)	(0.066)		
N.obs	239	59		

Note: Standard errors of the means in parenthesis

(a) Only 2 observations in this cell

Table 6: Error rates by race and region

Panel A: South

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.349	0.455	0.26	206
	(0.034)	(0.157)		
Non-white	0.387	0.232	0.01	219
	(0.038)	(0.057)		
N.obs	358	67		

Panel A: Other regions

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.391	0.500	0.32	52
	(0.073)	(0.224)		
Non-white	0.286	0.500	0.14	31
	(0.101)	(0.167)		
N.obs	67	16		

Note: Standard errors of the means in parenthesis

Table 7: Error rates conditional on trial characteristics

Defendant not represented by out of state attorney				
		Victim's race		
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.301	0.417	*0.23*	178
	(0.036)	(0.149)		
Non-white	0.333	0.239	*0.11*	169
	(0.043)	(0.064)		
N.obs	289	58		
Ineffective assistance of counsel not 1st claim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.352	0.385	*0.41*	192
	(0.036)	(0.140)		
Non-white	0.374	0.229	*0.02*	187
	(0.041)	(0.061)		
N.obs	318	61		
Ineffective assistance of counsel not in any claim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.457	0.571	0.29	112
	(0.049)	(0.202)		
Non-white	0.417	0.227	0.04	106
	(0.054)	(0.091)		
N.obs	189	29		
Prosecutorial suppression/witholding of evidence not in any claim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.380	0.600	*0.09*	176
	(0.038)	(0.163)		
Non-white	0.383	0.282	*0.11*	172
	(0.042)	(0.073)		
N.obs	299	49		
Improper interrogation not in any claim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.361	0.500	*0.15*	232
	(0.033)	(0.129)		
Non-white	0.387	0.271	*0.05*	222
	(0.038)	(0.058)		
N.obs	379	75		

Note: Standard errors of the means in parenthesis

Table 8: Error rates conditional on crime characteristics

Defendant not connected to community where crime occurred				
	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.381	0.500	*0.20*	190
	(0.037)	(0.139)		
Non-white	0.413	0.250	*0.016*	186
	(0.042)	(0.063)		
N.obs	314	62		
No police victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.364	0.438	*0.29*	241
	(0.032)	(0.128)		
Non-white	0.391	0.250	*0.019*	221
	(0.039)	(0.056)		
N.obs	386	76		
No high status victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.400	0.385	*0.46*	208
	(0.035)	(0.140)		
Non-white	0.397	0.296	*0.093*	185
	(0.043)	(0.063)		
N.obs	326	67		
Single victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.357	0.471	*0.19*	252
	(0.031)	(0.125)		
Non-white	0.383	0.281	*0.06*	244
	(0.036)	(0.057)		
N.obs	415	81		
No female victim				
	White	*Non-white*	*p-values*	*N.obs*
White	0.384	0.500	*0.23*	124
	(0.046)	(0.151)		
Non-white	0.363	0.235	*0.07*	136
	(0.048)	(0.074)		
N.obs	214	46		
Crime occurred during a robbery				
	White	*Non-white*	*p-values*	*N.obs*
White	0.347	0.600	*0.16*	80
	(0.055)	(0.245)		
Non-white	0.430	0.250	*0.04*	103
	(0.056)	(0.090)		
N.obs	154	29		
Felony murder				
	White	Non-white	*p-values*	N.obs
White	0.350	0.429	*0.35*	127
	(0.044)	(0.202)		
Non-white	0.385	0.257	*0.07*	152

	(0.045)	(0.075)
N.obs	237	42

Note: Standard errors of the means in parenthesis

Table 9: Possible bias of Appeal Courts

Panel A: Majority of final federal panel Republican

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.353	0.455	0.27	147
	(0.041)	(0.157)		
Non-white	0.284	0.194	0.13	145
	(0.043)	(0.067)		
N.obs	245	47		

Panel B: All judges appointed under Republican

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.229	0.500	0.18	39
	(0.072)	(0.289)		
Non-white	0.216	0.083	0.09	63
	(0.058)	(0.083)		
N.obs	86	16		

Panel C: Sentence under Republican administration

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.343	0.438	0.24	182
	(0.037)	(0.128)		
Non-white	0.402	0.217	0.01	178
	(0.043)	(0.061)		
N.obs	298	62		

Note: Standard errors of the means in parenthesis

Table 10: Role of Trial Courts

Panel A: Sentence determined by judge or hybrid

Defendant's race	**Victim's race** *White*	*Non-white*	*p-values*	*N.obs*
White	0.432	0.667	0.14	80
	(0.058)	(0.211)		
Non-white	0.396	0.231	0.12	61
	(0.071)	(0.122)		
N.obs	122	19		

Panel B: Sentence determined by jury

Defendant's race	**Victim's race** *White*	*Non-white*	*p-values*	*N.obs*
White	0.323	0.364	0.40	178
	(0.036)	(0.152)		
Non-white	0.368	0.283	0.13	189
	(0.041)	(0.062)		
N.obs	303	64		

Notes: Standard errors of the means in parenthesis

Panel A comprises AZ, CO, ID, MT, NE, AL, DE, IN, FL.

Panel B comprises all other States.

Table 11: Trial judges' incentives

Panel A: Judge reappointed by competitive election

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.380	0.455	0.32	203
	(0.035)	(0.157)		
Non-white	0.407	0.300	0.08	195
	(0.041)	(0.065)		
N.obs	337	61		

Panel B: Judge reappointed by retention election, legislature or gubernatorial appointment

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.306	0.333	0.47	39
	(0.078)	(0.333)		
Non-white	0.138	0.167	0.41	41
	(0.065)	(0.112)		
N.obs	65	15		

Note: Standard errors of the means in parenthesis

Table 12: Error rates by Defendant and Victim's race, Direct Appeal

A. Full sample

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.379	0.380	0.49	1492
	(0.013)	(0.051)		
Non-white	0.378	0.356	0.19	1483
	(0.016)	(0.020)		
N.obs	2312	663		

B. South

	Victim's race			
Defendant's race	*White*	*Non-white*	*p-values*	*N.obs*
White	0.407	0.444	0.28	1018
	(0.016)	(0.063)		
Non-white	0.414	0.377	0.13	1001
	(0.019)	(0.027)		
N.obs	1622	397		

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	3	0.06
1980	53	1	0.02
1981	66	0	0.00
1982	68	2	0.03
1983	41	0	0.00
1984	26	1	0.04
1985	25	1	0.04
1986	15	0	0.00
1987	6	0	0.00
1988	2	0	0.00
1989	1	0	0.00
1990	0		.

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	3	0.27
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	1	0.04
WA	3	1	0.33
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
1973	1	0	0.00
1974	12	1	0.08
1975	30	9	0.30
1976	51	10	0.20
1977	76	20	0.26
1978	78	15	0.19
1979	124	22	0.18
1980	118	24	0.20
1981	148	18	0.12
1982	161	19	0.12
1983	204	23	0.11
1984	211	23	0.11
1985	253	42	0.17
1986	190	24	0.13
1987	220	29	0.13
1988	287	46	0.16
1989	225	39	0.17
1990	212	40	0.19
1991	249	38	0.15
1992	258	35	0.14
1993	214	27	0.13
1994	249	40	0.16
1995	211	35	0.17

State	No. Total obs	No. Missing obs	Share missing
AL	189	52	0.28
AR	65	14	0.22
AZ	156	33	0.21
CA	234	38	0.16
CO	4	0	0.00
CT	2	0	0.00
DE	20	0	0.00
FL	617	76	0.12
GA	244	11	0.05
ID	27	11	0.41
IL	193	0	0.00
IN	64	0	0.00
KY	48	19	0.40
LA	96	8	0.08
MD	39	6	0.15
MO	90	10	0.11
MS	98	5	0.05
MT	11	5	0.45
NC	190	8	0.04
NE	20	7	0.35
NJ	34	8	0.24
NM	8	4	0.50
NV	83	34	0.41
OH	101	0	0.00
OK	169	44	0.26
OR	24	8	0.33
PA	176	62	0.35
SC	93	21	0.23
TN	99	28	0.28
TX	458	52	0.11
UT	20	3	0.15
VA	92	3	0.03
WA	15	9	0.60
WY	3	0	0.00