A Test of Racial Bias in Capital Sentencing

Alberto Alesina   Eliana La Ferrara

Online Appendix

A. Endogenizing the behavior of the criminal

This section of the appendix discusses an extension of our model where the probability of guilt, π^r is endogenously derived from an optimization problem of the criminal. We start by presenting the case where the race of the victim is given and known to the criminal, then discuss cases in which the race of the victim is ex ante unknown or in which it can be chosen by the criminal.

A1. Problem of the criminal

An individual who is considering whether to commit a capital crime trades of the expected benefits and costs of it. Let the benefit of committing the crime and getting away with it be $b > 0$; think of it as the money stolen from a bank (with a killing during the robbery) or the pleasure of killing an enemy. The cost of being sentenced to death having committed the crime is $c_g > 0$. The cost of being sentenced to death not having committed the crime is c_n, with $0 < c_n < c_g$.[1] All of the above b, c_g and c_n are public information. For an individual the cost of committing a crime, which may include the moral cost, is v and it is drawn from a distribution $\Im^r(v)$ with support in $\mathbb{R}^+$. The court knows the distribution but not the realization of v which is known only to the individual. We allow the distribution of costs to differ across races, thus allowing a higher propensity of minorities to commit crimes. The individual chooses whether to commit a crime taking into account the likelihood of being convicted and takes x_{rR} as given since it is chosen by the court.

In certain types of crimes the defendant cannot choose the victim and therefore his or her race. One example is a bank robbery with the killing of guards, whose race was unknown to the criminals ex ante. In a second type of crime the defendant wants to kill, say, a relative, in which case he also cannot choose the race of the victim but the race of the victim is known ex ante. In a third type of crime the defendant can choose the race of the victim, say in a rape with murder. We present the second case here. The others are discussed below.

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

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

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

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

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

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

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

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

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

as the probability of guilt, i.e. the probability that the realization of v is low enough so that a crime is committed. Note that if the court applied a different standard of proof depending on the race of the victim, e.g. $x_{m,W} < x_{m,M}$, then potential criminals would internalize that and ceteris paribus we would observe fewer crimes involving m, W pairs than m, M pairs. With an endogenous probability of committing a crime the potential criminal would incorporate in his calculations the courts' behavior. With exogenous probabilities, of course he would not.

The equilibrium of the model is given by (2) in the text together with:

$$\pi^r\left(x_{rR}^*\right)=\Im^r(v^{r*}(x_{rR}^*)) \tag{3}$$

By Brouwer's fixed point theorem an equilibrium exists. The proof of uniqueness is below.

A.2 Proof of uniqueness of the equilibrium

To simplify the notation, in sections A.2 and A.3 we omit subscripts and superscripts related to race, i.e. we write x instead of x_{rR}, v^* instead of v^{r*}, and f_g, f_n instead of f_g^r, f_n^r.

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

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

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

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

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

Proof. From (1) we have $v^*(0) = c_n - c_g < 0$. From (3) we have $\Im(v^*(0)) = 0$ because $\Im$ has support in $\mathbb{R}^+$. Furthermore, $\Im(v^*(x)) = 0$ for all $x \leq \widehat{x}$. Suppose that in equilibrium $x^* < \widehat{x}$. Then we would have $\pi(x^*) = \Im(v^*(x^*)) = 0$. But in this case the optimal response of the court would be to set $x^* = 1 > \widehat{x}$, a contradiction.

Claim 3. The equilibrium x^* is unique.

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

A.3 Proof that the equilibrium error rate is decreasing in x^*

The error rate (4) can be rewritten as

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

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

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

The first addendum in (A3) is negative because the equilibrium $\pi(x^*)$ is increasing in x^*, following Claims 1 and 2 above. To see that the second addendum is also negative, recall that from MLRP we know that $\frac{f_g(x)}{f_n(x)} > \frac{f_g(x^*)}{f_n(x^*)}$ for any $x > x^*$. Because the integrals in (A3) are calculated for $x \in (x^*, 1]$, then in this range $f_g(x^*) f_n(x) < f_g(x) f_n(x^*)$, hence the expression in square brackets is negative.

A.4 Case with random race of the victim

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

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

The payoff from not committing a crime is:

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

Let us define

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

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

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

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

A.5 Case where the race of the victim can be chosen

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

B. How the error rate varies with π

In this section we discuss how the error rate changes when the proportion of guilty individuals, π, changes. For the sake of compactness, we omit superscripts for race (and occasionally arguments) in the formulas below.[2]

The error rate E is:

$$E(x) = \frac{(1-\pi)\left[1 - F_n\left(x(\pi)\right)\right]}{\pi\left[1 - F_g\left(x(\pi)\right)\right] + \left[1-\pi\right]\left[1 - F_n\left(x(\pi)\right)\right]} \tag{4}$$

where $x(\pi)$ is implicitly defined by.

$$\frac{f_g(x)}{f_n(x)} = \frac{\alpha}{1-\alpha}\frac{1-\pi}{\pi} \tag{5}$$

The numerator of (4) captures the Type I error, as it is the product of (i) the area to the right of the standard of proof x and below the signal distribution for non-guilty defendants, f_n, and (ii) the proportion of non-guilty people in the population, $1-\pi$. The denominator is the sentencing rate, given by the mass of people (guilty and non-guilty) whose realized signal is to the right of x.

We want to assess the sign of the derivative $dE(x)/dx$.

As a preliminary step, it is useful to observe that π has two effects on $E(x)$ that go in opposite directions - we may call them "direct" and "indirect" effect.

To see the **direct effect** of π, suppose that x were independent of π. In that case:

$$\begin{aligned}
\frac{\partial E(\pi)}{\partial \pi} &= \frac{-(1-F_n)[\pi(1-F_g)+(1-\pi)(1-F_n)] - (1-\pi)(1-F_n)[(1-F_g)-(1-F_n)]}{[\pi(1-F_g)+(1-\pi)(1-F_n)]^2} \\
sign &\approx -\pi(1-F_n)(1-F_g) - (1-\pi)(1-F_n)(1-F_g) \\
&= -(1-F_n)(1-F_g) < 0.
\end{aligned}$$

Intuitively, the direct effect captures the fact that, for given standard of proof x, higher values of π imply that fewer of the people condemned will be innocent, hence the error rate is lower.

The **indirect effect** of π on E works through endogenous changes in x. To see how, notice that (i) $\frac{\partial x}{\partial \pi} < 0$ (from (5) and MLRP), and (ii) $\frac{\partial E}{\partial x} < 0$ (proved in the Appendix of the paper). Together, (i) and (ii) push in the direction of $\frac{\partial E}{\partial \pi} > 0$. Intuitively, the indirect effect captures the fact that when π increases the court chooses a lower standard of proof, which –for given π– increases the size of the Type I error.

The combined effect is obtained by taking the total derivative of (4) with respect to π. The sign of this derivative is determined by:

[2] In other words, when we write f_g and F_g we mean $f_g(x)$ and $F_g(x)$, etc.

$$\frac{dE}{dx} \overset{S}{\approx} -\left[(1-F_g) - \pi f_g \frac{\partial x}{\partial \pi}\right] [\pi(1-F_g) + (1-\pi)(1-F_n)] +$$

$$+\pi(1-F_g)\left[(1-F_g) - \pi f_g \frac{\partial x}{\partial \pi} - (1-F_n) - (1-\pi) f_n \frac{\partial x}{\partial \pi}\right]$$

$$= -(1-F_n)(1-F_g) - \pi(1-\pi)\left[f_n(1-F_g) - f_g(1-F_n)\right] \frac{\partial x}{\partial \pi} \tag{6}$$

The first addendum in (6) is negative and captures the direct effect. The second addendum is positive (notice that the term in square brackets is positive by MLRP) and captures the indirect effect. Which of the two dominates depends on the value of π and the functional form of F_g, F_n.

An example

In what follows we provide an illustration assuming a simple Normal parameterization for F_g and F_n: $F_g(x) = N(\mu_g, \sigma_g)$, $F_n(x) = N(\mu_n, \sigma_n)$ with $\mu_n < \mu_g$ and $\sigma_g = k\sigma_n$, $k > 0$.

The equilibrium value of x in this case is given by

$$x = \frac{2(\mu_n \sigma_g^2 - \mu_g \sigma_n^2) + \sqrt{(2\mu_n\sigma_g^2 - 2\mu_g\sigma_n^2)^2 - 4(\sigma_n^2 - \sigma_g^2)\left[\mu_g^2\sigma_n^2 - \mu_n^2\sigma_g^2 + 2\ln\left(\frac{\alpha}{1-\alpha}\frac{1-\pi}{\pi}\frac{\sigma_g}{\sigma_n}\right)\sigma_g^2\sigma_n^2\right]}}{2(\sigma_g^2 - \sigma_n^2)}.$$

Denote with $Erfc(z)$ the complementary error function
$Erfc(z) = 1 - \text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-\frac{t^2}{2}dt}$ and define

$$\Delta \equiv \sigma_g^2\sigma_n^2\left[\mu_g^2 - 2\mu_g\mu_n + \mu_n^2 + 2\ln\left(\frac{\alpha}{1-\alpha}\frac{1-\pi}{\pi}\frac{\sigma_g}{\sigma_n}\right)(\sigma_g^2 - \sigma_n^2)\right].$$

Then the error rate is

$$E(x) = \frac{(1-\pi)\left[2 - Erfc\left(\frac{\mu_n\sigma_n^2 - \mu_g\sigma_n^2 + \sqrt{\Delta}}{\sqrt{2}(\sigma_n^3 - \sigma_g^2\sigma_n)}\right)\right]}{2 - \pi Erfc\left(\frac{\mu_g\sigma_g^2 - \mu_n\sigma_g^2 - \sqrt{\Delta}}{\sqrt{2}(\sigma_g^3 - \sigma_g\sigma_n^2)}\right) + (\pi - 1)Erfc\left(\frac{\mu_n\sigma_n^2 - \mu_g\sigma_n^2 + \sqrt{\Delta}}{\sqrt{2}(\sigma_n^3 - \sigma_g^2\sigma_n)}\right)}.$$

Appendix Figure 1 shows how E varies with π for different degrees of relative noise in the signal distributions.[3] The ratio σ_g/σ_n measured on the rightmost axis is allowed to take values in $(0, 4]$. The remaining parameters are set at $\alpha = 0.5$, $\mu_n = 0.3$, $\mu_g = 6$, $\sigma_n = 2$.

In this example when $\sigma_g < \sigma_n$, i.e. $k < 1$, the error rate is monotonically decreasing in π. For intermediate values of k the relationship is non-monotonic, i.e. the error first decreases and then

[3] The parameterization used in this figure allows x to take values outside $[0, 1]$ because this improves the readability of the graphs. Qualitatively similar patterns can be obtained using the truncated Normal distribution in $[0, 1]$. Also note that the arguments used in this section are based on MLRP which does not require x to be defined on a compact set.

increases with π, and for $\sigma_g >> \sigma_n$ the latter effect dominates. Underlying this pattern is the relative size of the change in the Type I and Type II errors. To understand why, consider figure 2.

Appendix Figure 2 depicts the density functions $f_n(x)$ and $f_g(x)$ parameterized as above and the equilibrium value of x, indicated by a vertical line. In panel (a) $\sigma_g = 0.5\sigma_n$. When π increases from 0.1 (top graph) to 0.8 (bottom graph), the equilibrium value of x decreases from 4.7 to 3.4. This leads to a relative small increase in the area corresponding to the type I error ($1 - F_n(x^*)$ goes from .01 to .06) and a relatively larger decrease in the area corresponding to the type II error ($F_g(x^*)$ goes from .09 to .005). The latter declines at a faster rate because the distribution f_g is less dispersed. The overall effect is a decrease in the error rate from .12 to .01.

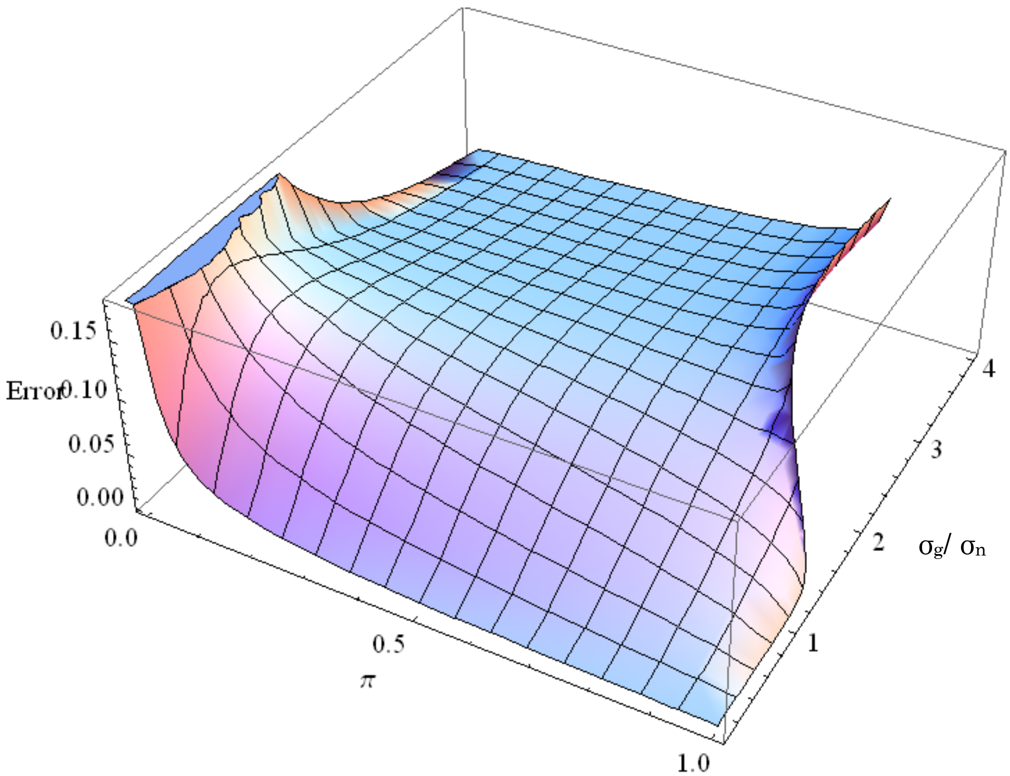
When we consider panel (b), where $\sigma_g = 2.5\sigma_n$, the pattern is reversed. Here the same increase in π from 0.1 to 0.8 induces a much larger reduction in x^*, from 5.3 to 0.9. This in turn leads to a relatively large increase of the area corresponding to the type I error ($1 - F_n(x^*)$ goes from .01 to .06) relative to the decrease in the area $F_g(x^*)$ (from .45 to .16), with a corresponding increase in the error rate from .09 to .10.

We can thus summarize the role played by π in our model and the implications for the interpretation of our results. First, the variable π is *not* the proportion of homicides committed across races (in which case one may conjecture that the proportion of intra-racial homicides would be larger). The variable π is the probability that an individual brought to court for a homicide is guilty of the death penalty. Second, based upon the (limited) empirical evidence that we were able to assemble and that we discuss in the text of the paper there is no clear pattern of π_{rR} which would allow to take a stand on their relative size. In particular, there is no evidence that the (imperfectly) predictable patterns of π_{rR} would systematically go in a direction that inficiates the validity of our test. Notice also that even if we had perfect information about the relative size of π_{rR} we would have to take a stand on the shape of the distributions $F_g(x)$ and $F_n(x)$. In the text we then proceed with the assumption that π does not systematically differ across pairs of defendant/victim races.

C. Additional empirical results

This section of the Appendix contains some additional empirical results.

First, we provide descriptive statistics on the missingness of victim's race. We start by tabulating the number of cases with missing race of the victim (Table A1), then we present a breakdown of the number of observations for which we could not find the race of the victim by year and by state, for both the Habeas Corpus (Table A2) and for the Direct Appeal (Table A3) datasets. We also assess the potential selection in missingness of victim's race through a balance test on observables (Table A4). Finally, We report a balance test for crime characteristics across pairs of defendant and victim's races (Table A5).



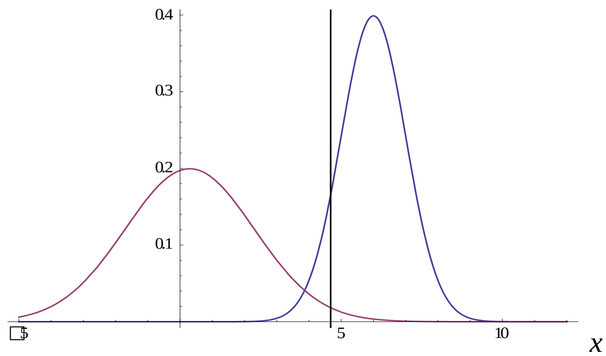
Appendix Figure 1:

Error rate as a function of π and σ_g/σ_n

Panel (a): $\sigma_g < \sigma_n$

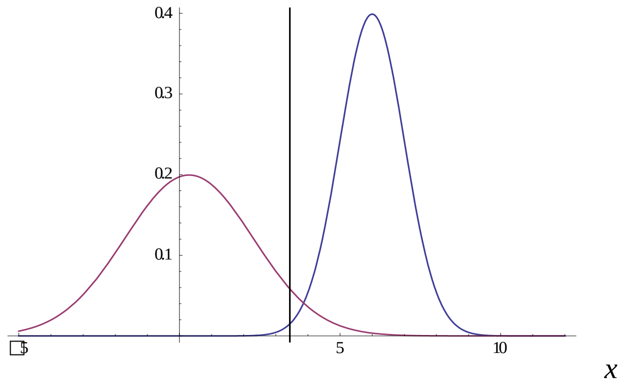
$\pi = 0.1$

$x^* = 4.67$
$1\text{-}F_n(x^*) = .014$
$F_g(x^*) = .09$
Error = .125



$\pi = 0.8$

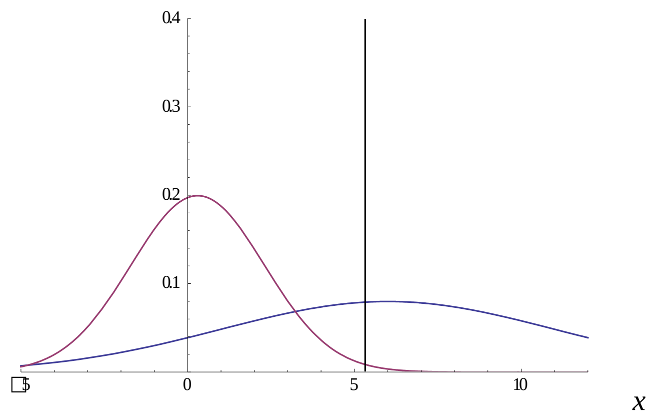
$x^* = 3.43$
$1\text{-}F_n(x^*) = .059$
$F_g(x^*) = .005$
Error = .015



Panel (b): $\sigma_g > \sigma_n$

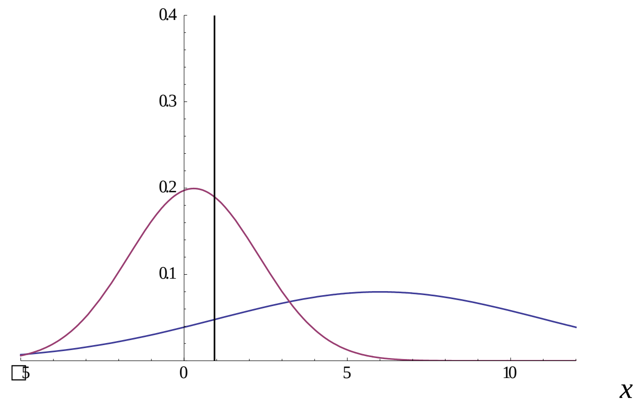
$\pi = 0.1$

$x^* = 5.3$
$1\text{-}F_n(x^*) = .006$
$F_g(x^*) = .45$
Error = .091



$\pi = 0.8$

$x^* = 0.91$
$1\text{-}F_n(x^*) = .379$
$F_g(x^*) = .155$
Error = .10



Appendix Figure 2: Equilibrium *x* for different values of π and σ_g / σ_n

Appendix Table A1: Missingness of race

Panel A: Habeas Corpus

		Missing victim's race		
		No	*Yes*	*Total*
Missing defendant's race	*No*	508	20	528
	Yes	16	0	16
	Total	524	20	544

Panel B: Direct Appeal

		Missing victim's race		
		No	*Yes*	*Total*
Missing defendant's race	*No*	3,717	447	4,164
	Yes	130	122	252
	Total	3,847	569	4416

Appendix Table A2:
Missingness of victim's race by year and State, Habeas Corpus

Year 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	50	2	0.04
1980	51	1	0.02
1981	66	0	0.00
1982	68	1	0.01
1983	41	0	0.00
1984	26	1	0.04
1985	25	0	0.00
1986	15	0	0.00
1987	6	0	0.00
1988	2	0	0.00
1989	1	0	0.00
			.

State	No. Total obs	No. Missing obs	Share missing
AL	19	1	0.05
AR	24	0	0.00
AZ	14	1	0.07
CA	4	0	0.00
DE	2	0	0.00
FL	95	4	0.04
GA	82	2	0.02
ID	3	0	0.00
IL	10	0	0.00
IN	4	0	0.00
KY	1	0	0.00
LA	34	0	0.00
MD	1	0	0.00
MO	26	1	0.04
MS	21	0	0.00
MT	4	0	0.00
NC	10	0	0.00
NE	6	0	0.00
NV	4	0	0.00
OK	11	2	0.18
PA	3	0	0.00
SC	7	0	0.00
TN	1	0	0.00
TX	108	9	0.08
UT	3	0	0.00
VA	26	0	0.00
WA	3	0	0.00
WY	2	0	0.00

Appendix Table A3:
Missingness of victim's race by year and State, Direct Appeal

Year of 1st sentence	No. Total obs	No. Missing obs	Share missing
1973	13	2	0.15
1974	56	15	0.27
1975	72	24	0.33
1976	72	12	0.17
1977	115	24	0.21
1978	184	37	0.20
1979	165	23	0.14
1980	189	15	0.08
1981	226	14	0.06
1982	290	30	0.10
1983	254	18	0.07
1984	275	25	0.09
1985	280	29	0.10
1986	288	31	0.11
1987	251	22	0.09
1988	234	16	0.07
1989	190	20	0.11
1990	165	13	0.08
1991	143	13	0.09
1992	108	10	0.09
1993	66	5	0.08
1994	11	1	0.09
1995	2	0	0.00

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

Appendix Table A4:
Selection in missingness of victim's race

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

Appendix Table A5: Balance test on crime characteristics

	Full sample							
	White defendant			Non-white defendant			Diff-in-diff	
	White victim	*Non-white victim*	*Diff=0 (p-val)*	*White victim*	*Non-white victim*	*Diff=0 (p-val)*	ΔΔ	*(p-val)*
Panel A: Habeas Corpus								
Victim								
Multiple victims	0.02	0.00	0.51	0.02	0.03	0.73	0.03	0.46
Female victim	0.54	0.29	0.05	0.44	0.45	0.92	0.26	0.08
Weapon								
Knife	0.19	0.27	0.48	0.14	0.18	0.50	-0.04	0.77
Handgun	0.37	0.20	0.20	0.50	0.41	0.25	0.07	0.65
Shotgun or rifle	0.16	0.27	0.32	0.20	0.18	0.77	-0.12	0.33
Strangulation	0.09	0.07	0.74	0.04	0.08	0.23	0.07	0.39
Circumstances								
Defendant connected to	0.27	0.18	0.39	0.24	0.27	0.68	0.12	0.34
Burglary or theft	0.12	0.06	0.50	0.11	0.07	0.41	0.02	0.85
Robbery	0.36	0.31	0.69	0.47	0.42	0.53	0.00	1.00
Kidnapping	0.07	0.06	0.95	0.07	0.07	0.92	0.01	0.92
Rape or sex related	0.21	0.13	0.40	0.19	0.16	0.56	0.05	0.65
Institutional killing	0.04	0.06	0.71	0.10	0.09	0.85	-0.03	0.71
Panel B: Direct Appeal								
Victim								
Multiple victims	0.26	0.19	0.12	0.20	0.29	0.00	0.16	0.00
Female victim	0.57	0.37	0.00	0.47	0.50	0.28	0.23	0.00
Weapon[(a)]								
Knife	0.28	0.35	0.47	0.27	0.26	0.87	-0.08	0.46
Handgun	0.34	0.45	0.32	0.45	0.42	0.60	-0.14	0.27
Shotgun or rifle	0.12	0.15	0.66	0.06	0.08	0.52	-0.02	0.85
Strangulation	0.09	0.00	0.17	0.02	0.00	0.10	0.07	0.23

Appendix Table A5 (cont'd): Balance test on crime characteristics

	South							
	White defendant			Non-white defendant			Diff-in-diff	
	White victim	*Non-white victim*	*Diff=0 (p-val)*	*White victim*	*Non-white victim*	*Diff=0 (p-val)*	ΔΔ	*(p-val)*
Panel A: Habeas Corpus								
Victim								
Multiple victims	0.01	0	0.74	0.01	0.04	0.11	0.04	0.29
Female victim	0.55	0.46	0.54	0.43	0.48	0.53	0.14	0.40
Weapon								
Knife	0.18	0.20	0.86	0.13	0.17	0.57	0.01	0.92
Handgun	0.40	0.30	0.53	0.50	0.38	0.17	-0.02	0.91
Shotgun or rifle	0.18	0.30	0.34	0.20	0.21	0.85	-0.11	0.47
Strangulation	0.08	0.00	0.37	0.04	0.07	0.48	0.10	0.25
Circumstances								
Defendant	0.24	0.09	0.25	0.23	0.27	0.52	0.20	0.19
Burglary or theft	0.13	0.00	0.22	0.10	0.08	0.68	0.11	0.32
Robbery	0.38	0.50	0.44	0.49	0.39	0.22	-0.22	0.21
Kidnapping	0.06	0.10	0.66	0.06	0.08	0.62	-0.02	0.86
Rape or sex related	0.23	0.20	0.81	0.18	0.18	1.00	0.03	0.83
Institutional killing	0.05	0.00	0.46	0.09	0.08	0.88	0.05	0.62
Panel B: Direct Appeal								
Victim								
Multiple victims	0.23	0.14	0.05	0.15	0.25	0.00	0.19	0.00
Female victim	0.57	0.32	0.00	0.46	0.52	0.07	0.31	0.00
Weapon[(a)]								
Knife	0.26	0.25	0.90	0.28	0.27	0.95	0.01	0.93
Handgun	0.35	0.38	0.83	0.46	0.42	0.51	-0.06	0.64
Shotgun or rifle	0.10	0.19	0.28	0.06	0.09	0.47	-0.06	0.45
Strangulation	0.09	0.00	0.22	0.02	0.00	0.16	0.07	0.30

Notes: (a) information available for 935 out of 3717 cases