CARESS Working Paper 01-09

Racial Fairness and Effectiveness of Policing

Nicola Persico[1]

University of Pennsylvania

February 2001

[1] Department of Economics, 3718 Locust Walk, Philadelphia, PA 19104. Email persico@ssc.upenn.edu, URL www.ssc.upenn.edu/~persico. This paper develops a line of inquiry stemming from joint work with John Knowles and Petra Todd. I am grateful to both of them, as well as to the NSF for financial support under grants SBR-9905564 and SES-0078870. Thanks also to Aldo Colussi for research assistanship.

Abstract

Citizens of two racial groups choose whether to engage in illegal activities, and police audit citizens. Citizens are heterogeneous according to legal earning opportunities, which are distributed differently across groups. We define fairness of policing as policing groups with the same intensity, and effectiveness of interdiction as reducing the amount of crime. We show that sometimes, forcing the police to behave more fairly can increase effectiveness of interdiction, and give exact conditions under which this is so. These conditions are based on the distributions of legal earning opportunities in the two groups, and are expressed as contraints on the QQ plot of these distributions.

Legal earning opportunities are not observable for those citizens who become criminals. However, we give conditions under which the QQ plot of legal earning opportunities equals the QQ plot based on reported income distributions (which are observable). We also discuss whether our notion of fairness is meaningful when the cost of being searched reflects the shame of being singled out by the police.

1 Introduction

Law enforcement practices often have disparate impact on different ethnic and racial groups. For instance, the current debate on racial profiling has shown that motorists on highways are much more likely to be searched by police looking for illegal drugs if the motorists are African-American. Similar allegations are made about searches by customs officials at airports.[1] Being searched imposes time costs on innocent citizens and, perhaps more importantly, can be a debasing experience and a humiliating one when conducted in public.[2] Such racial disparities are an undeniable reality of policing, and are widely perceived to be discriminatory. Public attention to this issue has resulted in policy changes aimed at correcting the disparity. The proposed remedies would reduce the extent to which racial or ethnic characteristics can be used in policing.[3]

However, forbidding the police from using some characteristics may reduce the effectiveness of policing. Often, those who engage in certain criminal activities predominantly belong to specific socio-economic backgrounds and tend to share certain characteristics such as low legal income, little educational attainment, and few outside economic opportunities. Observing whether a particular individual possesses these characteristics may help the police uncover criminal behavior, and conversely, ignoring

[1] See Knowles, Persico, and Todd (2001) for a case that is representative of the disparity in highway searches. Concerning airport searches, Gibeaut (1999) reports that, of all passengers who were strip-searched by customs officials in 1997, 60 percent were black or Hispanic. In the case of airport searches, however, the racial disparity may be due in part to the provenance of "risk" flights.

[2] For a description of the debasement connected with being bodily searched by airport customs officials, see Adedeji v. United States, 782 F. Supp. 688 (D. Mass.). See Anderson (1990) for a more general view of the relationship between the police and African-Americans.

[3] In response to well-publicized charges of racial profiling against New Jersey police, the Attorney General of New Jersey has released a report which calls for a ruling by the New Jersey Supreme Court holding that "race may play no part in an officer's determination of whether a particular person is reasonably likely to be engaged in criminal activity." (See Verniero and Zoubek (1999), pp. 52-3). Part V of the report lists a number of remedial actions aimed at avoiding that race be a factor in the decision to search a motorist.

some characteristics may lead to less efficient policing and to increased crime.[4]

These conflicting considerations reflect a fundamental tension of principle between the practical demands of law enforcement and equal treatment under the law. It is important to establish whether this tension exists in practice, not least because the trade-off between racial fairness and effectiveness of policing is the *sine qua non* premise of the existing judicial line on racial disparities in search and seizure. The current Supreme Court position is expressed in *U.S. v Whren*,[5] which finds that the Fourth Amendment permits pretextual stops based on objective justification, and does not admit second-guessing of the police officer's subjective motivation. This line explicitly upholds the use of racial and other characteristics as a factor in determining the likelihood that a person is engaging in a crime, as long as this use is reasonably related to law enforcement and is not a pretext for racial harassment.[6] This position gives precedence to public policy concerns for effectiveness of interdiction over the plight of innocent citizens who bear a disproportionate amount of the policing effort. Only if these two goals are in contrast is the position defensible.

This paper studies the trade-off between racial fairness and effectiveness of interdiction in the context of a simple model in which citizens of two racial groups choose whether to engage in illegal activities, and police audit (or search) citizens. The first point of this paper is that, in principle, there need not be a trade-off. Constraining police to behave in a racially fair manner—in our context, searching both racial groups with the same intensity—need not *per se* result in a greater crime rate. This is because, under a plausible incentive scheme for police, the equilibrium allocation does not coincide with the efficient (crime-minimizing) allocation, and so placing constraints on the behavior of the police may actually move the system towards efficiency.[7]

[4] See Beck and Daly (1999) for a discusssion of the public policy argument in favor of discretional searches of motorists.

[5] 517 U.S. 806, 116 S.Ct. 1769 (1996)

[6] See Kennedy (1997) for an exposition of this view, and more recently Thompson (1999) for a critical analysis.

[7] In this paper the term "efficiency" captures the effectiveness of interdiction and is used as a

To understand this statement it is necessary to understand why the equilibrium is inefficient in our model. Our assumption is that, when choosing whom to search, the objective of a police officer is to maximize the probability of uncovering criminal behavior. Suppose then that the population is composed of two groups, A and W. In equilibrium it cannot be that one group has a lower fraction of criminals, for otherwise no police officer would search members of the group with the lower fraction of criminals, and members of this group would respond by stepping up their illegal activity. Thus, in equilibrium the two groups must have the same fraction of criminals. The key point is that this equilibrium condition is independent of the elasticity of crime to policing. In contrast, this elasticity determines the conditions for efficiency. For instance, take an equilibrium allocation and suppose an increase of 1 percent in the search probability deters members of group W more than members of group A; then, efficiency requires that 1 percent of resources be transferred from searching group A to searching group W, to take advantage of the higher deterrence effect of policing on group W. This shows that there is a wedge between the equilibrium condition and the efficient allocation.[8]

Now, let us see how constraining police to be more fair can increase effectiveness of interdiction. Keep the assumption that group W is more responsive to policing, and suppose in the unconstrained equilibrium group W is policed with less intensity. Then, forcing the police towards a more fair behavior means increasing the intensity with which group W is policed; this moves the system towards efficiency. In general, if the elasticity to policing is lowest for the group which is policed more intensely, moving the system towards a more fair allocation *increases* the effectiveness of policing: there is no trade-off between fairness and efficiency.

Having formulated the efficiency/fairness trade-off in terms of elasticity to policing, we seek theoretical conditions to compare the elasticity in the two groups. Measuring the elasticity directly is a difficult exercise, which at a minimum requires some sort of exogenous variation in policing (see Levitt (1996), who also reviews the literature). We

synonimous of "crime minimization." This is discussed further in Section 2.1.

[8]How this conclusion depends on the police incentives is discussed in Section 9.

outline an alternative approach based on measuring the primitives of our simple model of policing, specifically, legal earning opportunities. The advantage of this approach is that it does not require observing any variation in policing. The reason why earning opportunities matter is that, given a certain intensity of policing, citizens whose legal earning opportunities are above a threshold will not engage in criminal activities, and those below a threshold will become criminals. Thus the amount of criminal activity—and hence also the elasticity of crime to policing—depends on the distribution of legal earning opportunities. We provide a necessary and sufficient condition to determine whether the group which in equilibrium is policed more intensely is also the group in which the elasticity of crime to policing is higher. Mathematically, this condition is expressed as a constraint on the QQ plot of the distribution of legal earning opportunities in the two groups. The QQ (quantile-quantile) plot $h(x)$ is easily constructed; given the c.d.f.'s F_A and F_W of earning opportunities in the two groups, it is $h(x) = F_A^{-1}(F_W(x))$. We show that—roughly speaking—if the function h has slope smaller than 1, then in equilibrium group A has a higher elasticity of crime to policing *and* is policed more intensely. As discussed above, this implies that moving slightly towards fairness reduces the effectiveness of policing. A similar condition allows to evaluate the consequences of large changes in fairness, i.e., of implementing the completely fair outcome in which the two groups are policed with the same intensity.

Although an agent's legal earning opportunity is unobservable if he/she decides to be a criminal, we can observe the reported earnings of those who work legally. We give conditions under which in equilibrium the QQ plot of reported earnings coincides with the QQ plot of potential legal earnings. As an illustration of this approach, we construct the QQ plot based on the empirical distributions of reported earnings of the African-American and white male residents of statistical metropolitan areas in the US. Intriguingly, we find that this QQ plot indeed seems to have slope smaller than 1 (although the slope seems quite close to 1 for plausible values of the parameters). In the context of our model, this would imply that the elasticity to policing at equilibrium is higher in the A group than in the W group. The interpretation is that *there is* a trade-off between fairness and effectiveness of policing. To the extent that our stylized model is a reasonable approximation of the real-world problem, this suggests that if one were to move towards fairness by constraining police to shift resources from the A group to the W group, then the total amount of crime would rise.

Finally, we discuss our notion of racial fairness. Our notion is based on equating across races the expected costs of being searched. This notion is redistributional in nature because it is based on the idea that police can (and should) redistribute these costs across racial groups. We inquire whether the stigmatization associated with being singled out for search—as distinct from physical costs of being searched—can be understood as a component of the cost to be redistributed. To this end, we present a notion of stigmatization—which we call disrepute—and suggest that this notion cannot justify our redistributional notion of fairness. This is because in our definition the total amount of stigmatization within a race is independent of the search behavior of police. Thus, changes in the interdiction strategy do not affect the amount of stigmatization suffered by any racial group.

1.1 Related literature

The formal economic literature on policing is vast, and originates with Becker (1968) and Ehrlich (1973). We refer to Ehrlich (1996) for a survey of the literature. In this context, the issue of fairness is generally discussed in terms of relationship between crime and sanction. For instance, in Shavell and Polinsky (1998) a sanction is "fair" when it is proportional to the gravity of the act committed. In contrast, our notion of fairness applies to the strategy of policing, and is a statement about the intensity with which two different groups are policed. In a recent paper (Knowles, Persico, and Todd (2000)) a model similar to the one used here was employed in an empirical analysis of highways searches for drugs. That paper asks whether the police force is racially prejudiced, and it does not address issues of fairness.

We discuss the effects of placing constraints on police behavior to achieve outcomes that are more fair. This is a comparative statics exercise, and in this sense is related to the literature on affirmative action (Coate and Loury (1993), Mailath, Samuelson, and Shaked (2000), Moro and Norman (2000), and Norman (2000)). That literature looks at the effects of constraining the hiring behavior of firms. While our analysis is different in many respects, one fundamental difference is worth highlighting. The affirmative action literature builds on Arrow's (1972) model of statistical discrimination, which analyzes a game with multiple equilibria. While that literature has been very successful in highlighting a realistic force leading to discrimination, the comparative statics conclusions about the effect of given policies are often qualified by the fact that they apply to a specific equilibrium. In contrast, while we take as given the heterogeneity between groups and therefore do not explain how identical groups can be treated differently in equilibrium, our model does have a a unique equilibrium, and so our comparative statics results are not subject to the multiple-equilibria criticism.

2 Model

There is a continuum of police officers with measure 1. Police officers search citizens, and each police officer makes exactly $\overline{S}$ searches. When a citizen is searched who is engaged in criminal activity, he/she is uncovered with probability 1. A police officers maximizes the number of successful searches, i.e., the probability of uncovering criminals.

There is a continuum of citizens divided into two racial groups, A and W with measures N_A and N_W respectively. If a citizen does not engage in crime, he/she is legally employed and receives earnings $x > 0$ independent of being searched. If a citizen engages in crime he/she receives utility (in monetary terms) of $H > 0$ if not searched and $J < 0$ if searched.Every citizen knows his own legal earning opportunity. We model x as the realization from a random variable X_R, where $R \in \{A, W\}$ denotes the race of the citizen. Denote with F_R and f_R the relative c.d.f.'s and densities. We assume that X_A and X_W are nonnegative, and that their supports are bounded intervals including zero. Notice that the distribution of legal earning opportunities is allowed to differ across races.

A citizen's legal earning opportunity x is unobservable to the police when they choose whether to search him/her. However, the police can observe the race of the citizen.

In the aggregate there is a measure $\overline{S}$ of searches that is allocated between the two groups. Denote with S_R the measure of searches of citizens of race R. The feasibility constraint requires $S_A + S_W = \overline{S}$. Denote with $\sigma_R = S_R/N_R$ the probability that a random individual of race R is searched. The budget constraint can be expressed as

$$N_A\sigma_A + N_W\sigma_W = \overline{S}. \tag{1}$$

Given σ_R, a citizen of race R with legal earning opportunity x commits a crime if and only if

$$x < \sigma_R J + (1 - \sigma_R) H = \sigma_R (J - H) + H \overset{\text{def}}{=} q(\sigma_R).$$

Thus, the fraction of citizens of race R that engage in criminal activity is $F_R(q(\sigma_R))$. The function $q(\sigma)$ represents the expected reward of being a criminal when the deterrence level is σ.

2.1 Definitions: evaluating outcomes

We now give our definition of efficiency. Consistent with the objective of this paper, we use a definition of efficiency that captures *effectiveness of interdiction*.[9] When defining efficiency of interdiction we should account for costs and benefits of the interdiction

[9] In particular, our definition ignores the utility of citizens who commit a crime, and so is not meant to capture Pareto efficiency. While it may be interesting to discuss the welfare effects of criminal activity and the merits of punishing citizens for committing crimes, this is not the focus of the debate. Rather, the public policy concern is for efficiency of interdiction.

activity. In our model the cost of interdiction is $\overline{S}$ and is taken as given. Therefore, it is appropriate for the next definition to focus only on the benefits of interdiction, i.e., the reduction in crime.

Definition 1 *The outcome* (σ_A, σ_W) *is* ***more efficient*** *than the outcome* (σ'_A, σ'_W) *if the total number of citizens who commit crimes is lower at* (σ_A, σ_W)*.*

We now give our definition of fairness, which should be thought of as *fairness of policing*. An outcome is completely fair if the two racial groups are policed with the same intensity. By extension, making an outcome more fair means reducing the difference between the intensity with which the two groups are policed.

Definition 2 *The outcome* (σ_A, σ_W) *is* ***more fair than*** *the outcome* (σ'_A, σ'_W) *if* $|\sigma_A - \sigma_W| < |\sigma'_A - \sigma'_W|$*. The outcome* (σ_A, σ_W) *is* ***completely fair*** *if* $\sigma_A = \sigma_W$*.*

This definition is tailored to capture the concerns of those who criticize existing policing practices. It is consistent with the idea that there is some (unmodeled) cost imposed on citizens who are searched and that it is desirable to equalize the expected cost across races.[10] This definition is essentially comparative, since it is concerned with the predicament of two citizens who belong to different racial groups; as such, this notion of fairness cannot be discussed in a model with only one group of citizens, such as Polinsky and Shavell (1998).

2.2 Definitions: ordering income distributions

We first introduce the notion of QQ plot and then discuss its interpretation. The QQ plot is a construction of statistics that finds applications in a number of fields: descriptive statistics, stochastic majorization and theory of income inequality, statistical inference and hypothesis testing.[11] A related notion, the Ratio-at-Quantiles plot, was

[10] More on this in Section 8.

[11] See, respectively, Wilk and Gnanadesikan (1968), Arnold (1987) p. 47 and ff., and Lehmann (1988).

employed by Wohlstetter and Coleman (1970) to describe income disparities between whites and nonwhites.

The QQ plot is an increasing function with domain and range being the supports of the random variables X_W and X_A, respectively.

Definition 3 *$h(x) = F_A^{-1}(F_W(x))$ is the QQ (quantile-quantile) plot of F_A against F_W.*

Given any income level q with a given percentile in the white income distribution, the number $h(q)$ indicates the income level with the same percentile in the African-American population. In other words, the function $h(q)$ matches quantiles in the white and African-American populations. An equivalent, perhaps more insightful definition is the following. The QQ plot is the unique function h such that $h(X_W)$ is distributed as X_A.[12] This equivalence is useful for the interpretation; the QQ plot can be thought of as the way in which the random variable X_W must be "compressed" or "stretched" to obtain the random variable X_A.

Relations of stochastic dominance can be expressed in terms of the QQ plot. For our purposes, it is useful to define first-order stochastic dominance in terms of the QQ plot.

Definition 4 *F_W first-order stochastically dominates F_A if $h(x) \leq x$ for all x.*

The previous definition is obviously equivalent to the more familiar definition of dominance requiring that $F_W(x) \leq F_A(x)$ for all x. Finally, we introduce the notion of "stretch" which is novel in this paper.

Definition 5 *F_W is a **stretch** of F_A if $h'(x) \leq 1$ for all x.*

[12] To verify the equivalence of the two definitions, suppose the function h is such that $h(X_W)$ is distributed as X_A or, equivalently, $h^{-1}(X_A)$ is distributed as X_W. Then for all x we have $F_W(x) = \Pr(X_W \leq x) = \Pr\left(h^{-1}(X_A) \leq x\right) = F_A(h(x))$, and inverting the function F_A we reach the expression for h given by Definition 3.

It is easy to give examples of stretches. A Uniform (or Normal) distribution X_W is a stretch of any other Uniform (or Normal) distribution X_A with smaller variance.[13]

While all the definitions in this section are applicable to generic c.d.f.'s F_A and F_W, in our model the infima of the supports of F_A and F_W are assumed to coincide. Under this assumption, if F_W is a stretch of F_A then *a fortiori* F_W first-order stochastically dominates F_A. Therefore, in our environment the property of stretch implies (is stronger than) first-order stochastic dominance.

3 Equilibrium and Efficient Interdiction

3.1 Equilibrium conditions

We use a * superscript to denote equilibrium quantities. A police officer who maximizes the number of successful searches will search only members of the racial group with the highest fraction of criminals. If the equilibrium is interior, that is $\sigma_A^*, \sigma_W^* > 0$, a police officer must be willing to search either group, so both groups must have the same fraction of criminals. Thus, at an interior equilibrium the following condition must be verified

$$F_A\left(q\left(\sigma_A^*\right)\right) = F_W\left(q\left(\sigma_W^*\right)\right). \tag{2}$$

If the equilibrium is not interior then an inequality will generally hold. Whether the equilibrium is interior or not, it is easy to see that the value of the left- and right-hand sides of equation (2) is uniquely determined in equilibrium. Leaving aside uninteresting constellations of primitives,[14] this uniquely pins down the equilibrium

[13] Indeed, when X_W is Uniform (or Normal) the random variable $a + bX_W$ is distributed as Uniform (or Normal). This means that when X_A and X_W are both Uniformly (or Normally) distributed r.v.'s, the QQ plot has a linear form $h(x) = a + bx$. In this case $h' < 1$ iff $b < 1$ iff $Var(X_A) = b^2 Var(X_W) < Var(X_W)$.

[14] If $\overline{S}$ is very large or very small then in equilibrium all citizens (of both races) commit crimes, or no citizen commits a crime. In this case a trivial multiplicity of equilibria arises in terms of the police search intensity.

in tems of σ's. For instance, if the equilibrium is interior then obviously there is a unique pair σ_A^*, σ_W^* which solves the equilibrium condition (2) subject to the feasibility constraint (1). In this paper we restrict attention to constellations of parameters for which the equilibrium is interior, that is, both racial groups are searched with positive probability. From the empirical viewpoint, this is a reasonable restriction. To guarantee interiority of equilibrium in the forthcoming analysis, we will henceforth maintain the following assumption.

Assumption 1 $F_A\left(\frac{\overline{S}}{N_A}(J-H)+H\right) < F_W(H)$.

Assumption 1 is used in the proof of Lemma 1, and is more likely to be verified when $\overline{S}$ is large.

3.2 Efficiency conditions

We denote efficient outcomes by the superscript ^{EFF}. A social planner who minimizes the total amount of crime solves

$$\begin{aligned}&\min_{\sigma_A,\sigma_W} N_A F_A(\sigma_A(J-H)+H)+N_W F_W(\sigma_W(J-H)+H)\\ &\text{subject to } N_A\sigma_A+N_W\sigma_W=\overline{S}\end{aligned}$$

or, using the constraint to eliminate σ_W

$$\min_{\sigma_A} N_A F_A(\sigma_A(J-H)+H)+N_W F_W\left(\frac{\overline{S}-N_A\sigma_A}{N_W}(J-H)+H\right).$$

The derivative with respect to σ_A is

$$N_A(J-H)\left[f_A(\sigma_A(J-H)+H)-f_W(\sigma_W(J-H)+H)\right]. \tag{3}$$

Equating the derivative to zero yields the first order condition for an interior optimum

$$f_A\left(q\left(\sigma_A^{EFF}\right)\right)=f_W\left(q\left(\sigma_W^{EFF}\right)\right). \tag{4}$$

This condition is necessary for full efficiency. The densities f_A and f_W represent elasticities of crime to policing. At an interior efficient allocation, the elasticities are equal. Condition (4) is different from condition (2) and thus will generally not hold in equilibrium; therefore, the equilibrium allocation is generally inefficient.

4 Benchmark: The Symmetric Case

In this section we discuss the benchmark case where $F_A = F_W$, that is, the distribution of legal earning opportunities is the same in the two groups. We call this the symmetric case since here $\sigma_A^* = \sigma_W^*$, i.e., the equilibrium is symmetric and completely fair. Even in the symmetric case, in some circumstances the efficient allocation is very asymmetric and the symmetric (and hence completely fair) allocation *minimizes* efficiency. This shows that the efficient allocation can be highly unfair and this does not depend on the heterogeneity of the income distributions, i.e. on the difference between races.[15] We interpret this result as indication that there is little logical relationship between the concept of efficiency and properties of fairness.

Proposition 1 *Suppose $F_A = F_W$. Then the equilibrium allocation is completely fair. If in addition F_A (and hence F_W) is a concave function on its support then the equilibrium allocation minimizes efficiency, and at the efficient allocation one of the two groups is not policed.*

Proof: Since $F_A = F_W$, the equilibrium condition (2) implies $\sigma_A^* = \sigma_W^*$, so the equilibrium is symmetric and completely fair.

To show the second statement, observe that since f_A and f_W are decreasing, the only pair σ_A, σ_W that solves the efficiency condition (4) together with the feasibility

[15] A similar conclusion that the efficient allocation can be unfair in a symmetric world was reached by Norman (2000) in the context of an Arrow model of statistical discrimination in the workplace. There, race-dependent (asymmetric) investment helps alleviate the inefficiency caused by workers who test badly and end up assigned to low-skilled jobs in spite of having invested a lot in human capital.

condition (1) is $\sigma_A^* = \sigma_W^*$. Now, observe that the second order conditions associated with efficiency maximization (crime minimization) are—twice differentiating the planner's objective function of Section 3.2

$$0 < N_A (J-H)^2 \left[f_A' \left(\sigma_A (J-H) + H\right) + \frac{N_A}{N_W} f_W' \left(\frac{\overline{S} - N_A \sigma_A}{N_W} (J-H) + H \right) \right]$$
$$= \left[N_A (J-H)\right]^2 \left[\frac{1}{N_A} f_A' \left(q\left(\sigma_A\right)\right) + \frac{1}{N_W} f_W' \left(q\left(\sigma_W\right)\right) \right]$$

But by assumption f_A' and f_W' are negative, so the second order conditions for efficiency maximization are not met. Indeed σ_A^*, σ_W^* meets the conditions for efficiency minimization. Finally, to show that either $\sigma_A^{EFF} = 0$ or $\sigma_W^{EFF} = 0$, observe that σ_A^*, σ_W^* is the unique solution to the necessary conditions (4) for an interior efficient allocation. Since σ_A^*, σ_W^* minimizes efficiency, the allocation that maximizes efficiency must be a corner solution. ∎

The reason why symmetric allocations may be so inefficient is apparent: the optimization problem that characterizes the efficient allocation need not be convex in the σ's. Convexity of the objective function fails when the c.d.f.'s are concave, and so the problem of maximizing efficiency may have asymmetric solutions even when groups are identical. Intuitively, there is no reason why searching two identical groups at the same rate should reduce aggregate crime more than focussing all the resources on one group.

The above result highlights the fact that fairness and efficiency may be antithetical, and this for reasons that have nothing to do with differences across groups. Our main focus, however, is on the efficiency consequence of perturbing the equilibrium allocation in the direction of fairness. This question cannot be asked in the symmetric case since, as we have seen, the equilibrium outcome is already completely fair. This is the subject of the next sections.

5 Small Adjustments Towards Fairness and Effectiveness of Interdiction

In this section we present conditions under which a small change from the equilibrium towards fairness, i.e., towards equalizing the search intensities across races, decreases efficiency relative to the equilibrium level.

Lemma 1 *Suppose F_W first-order stochastically dominates F_A. Then making the equilibrium allocation marginally more fair increases the total amount of crime if and only if $h'(q(\sigma_W^*)) < 1$.*

Proof: First, observe that the equilibrium is interior. Indeed, it is not an equilibrium to have $\sigma_A^* = \overline{S}/N_A$ and $\sigma_W^* = 0$ because then we would have

$$F_A(q(\sigma_A^*)) = F_A\left(\frac{\overline{S}}{N_A}(J-H)+H\right) < F_W(H) = F_W(q(0)).$$

where the inequality follows from Assumption 1. This cannot be an equilibrium since in this case a police officer's best response is only to search group W, and this is inconsistent with $\sigma_W^* = 0$. A similar argument shows that it is not an equilibrium to have $\sigma_W^* = \overline{S}/N_W$ and $\sigma_A^* = 0$, since by stochastic dominance $F_A(H) \geq F_W(H) > F_W\left(\frac{\overline{S}}{N_W}(J-H)+H\right)$.

The (interior) equilibrium must satisfy condition (2). Leaving aside trivial cases where in equilibrium every citizen or no citizen is criminal, and so marginal changes in policing intensity do not affect the outcome, condition (2) can be written as

$$q(\sigma_A^*) = h(q(\sigma_W^*)). \tag{5}$$

Now, given any x, by definition of h we have $F_W(x) = F_A(h(x))$, and differentiating with respect to x

$$f_W(x) = h'(x)\cdot f_A(h(x)).$$

Substituting $q(\sigma_W^*)$ for x and using (5) yields

$$f_W(q(\sigma_W^*)) = h'(q(\sigma_W^*)) \cdot f_A(q(\sigma_A^*)). \tag{6}$$

Now, let us evaluate expression (3) at the equilibrium. Using (6) we can rewrite it as

$$N_A(J - H) f_A(q(\sigma_A^*)) [1 - h'(q(\sigma_W^*))]. \tag{7}$$

This formula represents the derivative of the total amount of crime with respect to σ_A. If $h'(q(\sigma_W^*)) < 1$ this expression is negative (remember than $J - H$ is negative), and so a small decrease in σ_A from σ_A^* increases the total amount of crime. To conclude the proof we need to show that moving from the equilibrium towards fairness means decreasing σ_A^*.

To this end, recall that by assumption F_W first-order stochastically dominates F_A which in view of Definition 4 means that for all x we have $h(x) \leq x$. But then equation (5) implies $q(\sigma_A^*) \leq q(\sigma_W^*)$ and so $\sigma_A^* \geq \sigma_W^*$ (remember that $q(\cdot)$ is a decreasing function). Thus, moving from the equilibrium towards fairness requires decreasing σ_A^*. ∎

As is clear from expression (7), the magnitude of h' measures the inefficiency of marginally shifting the focus of interdiction toward group A. When h' is small it pays to increase σ_A because the elasticity to policing at equilibrium is higher in the A group.

The previous result yields the following proposition.

Proposition 2 *Suppose that F_W is a stretch of F_A. Then a marginal change from the equilibrium towards fairness decreases efficiency.*

Proof: Since F_W is a stretch of F_A we have $h'(x) \leq 1$ for all x, and so the second condition of Lemma 1 is verified. As for the first condition, since by assumption F_W and F_A have the same infimum of their supports, the stretch property guarantees that F_W first-order stochastically dominates F_A. Thus all the conditions of Lemma 1 are met for any value $q(\sigma_W^*)$ in the support of h. ∎

6 Implementing Complete Fairness

Under the conditions of Proposition 2, small changes towards fairness unambiguosly decrease efficiency; we now turn to large changes. We give sufficient conditions under which implementing the completely fair outcome decreases efficiency relative to the equilibrium point.

Proposition 3 *Assume F_W first-order stochastically dominates F_A, and suppose there are two numbers $\underline{q}$ and $\overline{q}$ such that*

$$h'(x) \leq \frac{\min_{y \in [h(x),x]} f_A(y)}{f_A(h(x))} \text{ for all } x \text{ in } \left[\underline{q}, \overline{q}\right]. \tag{8}$$

Assume further that the equilibrium outcome (σ_A^*, σ_W^*) is such that $\underline{q} \leq q(\sigma_W^*) \leq q(\sigma_A^*) \leq \overline{q}$. Then, the completely fair outcome is less efficient that the equilibrium outcome.

Proof: See the appendix. ∎

Since $\min_{y \in [h(x),x]} f_A(y) \leq f_A(h(x))$, condition (8) is more stringent than simply requiring that $h'(x) \leq 1$. It is sometimes easy to check whether condition (8) holds.[16] However, condition (8) can only hold on a limited interval $\left[\underline{q}, \overline{q}\right]$. This interval cannot encompass the entire support of the two ditributions, and in general there will exist values of $\overline{S}$ such that $q(\sigma_A^*)$ or $q(\sigma_W^*)$ lie outside the interval $\left[\underline{q}, \overline{q}\right]$. More importantly, there generally exist values of $\overline{S}$ for which the conclusion of Proposition 3 is overturned and there is no trade-off between complete fairness and efficiency. This is shown in the next proposition.

Proposition 4 *Assume the suprema of the supports of F_A and F_W do not coincide. Then there are values of $\overline{S}$ such that the completely fair outcome is more efficient than the equilibrium outome.*

[16] For instance, suppose F_A is a uniform distribution between 0 and K. In this case, any function h with $h(0) = 0$ and derivative less than one satisfies the conditions in Proposition 3 for all x in $[0, K]$. Indeed, in this case $f_A(y)/f_A(h(x)) = 1$ for all y in $[h(x), x]$ so long as $x \in [0, K]$.

Proof: We can restate the assumption on the supports as follows. Denote with $\widehat{q}$ the supremum of the suppport of F_A. Then, $F_W(\widehat{q}) > 1$ (this is without loss of generality, as we can relabel A and W).

Denote with $\overline{\sigma} = \overline{S}/(N_A + N_W)$ the search intensity at the completely fair outcome. The variation in crime due to the shift to complete fairness is

$$\begin{aligned} TV = & N_A [F_A(\overline{\sigma}(J-H)+H) - F_A(\sigma_A^*(J-H)+H)] \\ & + N_W [F_W(\overline{\sigma}(J-H)+H) - F_W(\sigma_W^*(J-H)+H)]. \end{aligned} \tag{9}$$

Case $F_W(H) < 1$. For $\overline{S}$ sufficiently small we get a corner equilibrium in which $\sigma_W^* = 0$ and $\sigma_A^* = \overline{S}/N_A$. Therefore, $\sigma_W^* < \overline{\sigma} < \sigma_A^*$. Also, when $\overline{S}$ sufficiently small σ_A^* approaches zero, so $\sigma_A^*(J-H)+H$ lies above $\widehat{q}$. Summing up we have

$$\widehat{q} < q(\sigma_A^*) < q(\overline{\sigma}) < H.$$

Using these inequalities and recalling that $F_A(\widehat{q}) = 1$ we can rewrite expression (9) as

$$TV = N_A[1-1] + N_W[F_W(q(\overline{\sigma})) - F_W(H)] < 0.$$

Case $F_W(H) = 1$. In this case, pick the maximum value of $\overline{S}$ such that in equilibrium all citizens engage in crime. Formally, this is the value of $\overline{S}$ giving rise to σ_W^*, σ_A^* with the property that $F_W(q(\sigma_W^*)) = F_A(q(\sigma_A^*)) = 1$ and $F_R(q(\sigma_R^* + \varepsilon)) < 1$ for all $\varepsilon > 0$ and $R = A, W$. By definition of $\widehat{q}$ we are guaranteed that $\widehat{q} = q(\sigma_A^*) < q(\overline{\sigma}) < q(\sigma_W^*)$. Expression (9) becomes

$$TV = N_A[1-1] + N_W[F_W(q(\overline{\sigma})) - 1] < 0.$$

∎

Proposition 4 highlights the importance of knowing whether the equilibrium values $q(\sigma_R^*)$ lie inside the interval $[\underline{q}, \overline{q}]$; this is additional information beyond the knowledge of F_A and F_W which was not necessary for Proposition 2. Proposition 4 should be contrasted with Proposition 2. One resulting insight should be that it is easier to predict the impact on efficiency of small changes in fairness, rather than the effects of a shift to complete fairness.[17]

7 Recovering the Distribution of Legal Earning Opportunities

What is relevant for the citizen's decision of whether to engage in crime is his/her *potential* legal earning opportunities. To estimate the QQ plot it would be helpful to know the distributions of potential legal earning opportunities. However, potential legal earnings are sometimes unobservable. This is becaue, according to our model, citizens with potential legal income below a threshold choose to engage in crime. These peple do not take advantage of their legal earning opportunities. Some of these people will be caught and put in jail, and are thus missing from our data, which raises the issue of truncation. The remaining fraction of those people who choose to become criminals escapes detection. When polled, these people are unlikely to report as their income the potential legal income—the income that they would have earned had they not engaged in crime. If we assume that these "lucky" criminals report zero income when polled, then we have a problem of censoring of our data.

In this section we present a simple extension of the model that circumvents these selection problems and is more realistic. Within the assumptions of this augmented model, the QQ plot of reported earnings can easily be adjusted to coincide with the QQ plot of potential legal earnings. Thus, measurement of the QQ plot of potential legal earnings is unaffected by the selection problem even though the c.d.f.'s themselves are affected.

[17] Although in the proof of Proposition 4 we have chosen the value of $\overline{S}$ such that at the completely fair outcome all citizens of one group engage in crime, this feature is not necessary for the conclusion. In fact, it is possible to construct examples in which, even as $h' < 1$, (a) the completely fair allocation is more efficient than the equilibrium one, and (b) at the completely fair allocation both groups have a fraction of criminals between zero and one.

Consider a world in which a fraction α of citizens of both races is *decent*, i.e., will never engage in crime no matter how low their legal earning opportunity. The remaining fraction $(1-\alpha)$ of citizens are *street*, i.e., will engage in crime if the expected return from crime exceeds their legal income opportunity; this is the type of agent we have discussed until now. The police cannot distinguish whether a citizen is decent or street. The base model in the previous sections corresponds to the case where $\alpha = 0$.

Given a search intensity σ_R for group R, the fraction of citizens of race R who engage in crime is $(1-\alpha) F_R(q(\sigma_R))$. The analysis of Section 3 goes through almost unchanged, and for any $\alpha > 0$ the equilibrium and efficiency conditions coincide with conditions (2) and (4). Consequently, the equilibrium and efficient allocations are independent of α, and the analysis of Sections 5 and 6 goes through verbatim. This means that also in this augmented model, our predictions about the relationship between efficiency and fairness depend on the shape of the QQ plot of (unobservable) legal earning opportunities.

Proposition 5 *Assume all citizens who engage in crime report earnings of zero while all citizens who do not engage in crime realize their potential legal earnings and report them truthfully. Then, given any fraction* $\alpha \in (0,1)$ *of decent citizens, the QQ plot of the reported earnings equals, in equilibrium, the QQ plot of potential legal earnings.*

Proof: Let us compute $\widehat{F}_R(x)$, the distribution of reported earnings. All citizens of race R with potential legal earnings of $x > q(\sigma_R^*)$ do not engage in crime regardless of whether they are decent or street. They realize their potential legal earnings, and report them truthfully. This means that $\widehat{F}_R(x) = F_R(x)$ when $x > q(\sigma_R^*)$.

In contrast, citizens of race R with an $x < q(\sigma_R^*)$ engage in crime when they are street, and report zero income. Thus, for $x < q(\sigma_R^*)$ the fraction of citizens who report income less than x is

$$\widehat{F}_R(x) = \alpha F_R(x) + (1-\alpha) F_R(q(\sigma_R^*)).$$

The function $\widehat{F}_R(x)$ is continuous. Furthermore, the quantity $(1-\alpha)F_R(q(\sigma_R^*))$is equal to some constant β which, due to the equilibrium condition, is independent of R. We can therefore write

$$\widehat{F}_R(x) = \begin{cases} \alpha F_R(x) + \beta & \text{for } x < q(\sigma_R^*) \\ F_R(x) & \text{for } x > q(\sigma_R^*). \end{cases}$$

The inverse of $\widehat{F}_R$ reads

$$\widehat{F}_R^{-1}(p) = \begin{cases} F_R^{-1}\left(\frac{p-\beta}{\alpha}\right) & \text{for } p < F_R(q(\sigma_R^*)) \\ F_R^{-1}(p) & \text{for } p > F_R(q(\sigma_R^*)). \end{cases} \tag{10}$$

Denote the QQ plot of the reported income distributions by

$$\widehat{h}(x) \stackrel{\text{def}}{=} \widehat{F}_A^{-1}\left(\widehat{F}_W(x)\right).$$

When $x > q(\sigma_W^*)$ then $\widehat{F}_W(x) = F_W(x)$, and so

$$\begin{aligned} \widehat{h}(x) &= \widehat{F}_A^{-1}\left(\widehat{F}_W(x)\right) \\ &= \widehat{F}_A^{-1}(F_W(x)) \\ &= F_A^{-1}(F_W(x)) = h(x), \end{aligned}$$

where the second-to-last equality follows from (10) because here $F_W(x) > F_W(q(\sigma_W^*)) = F_A(q(\sigma_A^*))$. When $x < q(\sigma_W^*)$ then $\widehat{F}_W(x) = \alpha F_W(x) + \beta$, and so

$$\begin{aligned} \widehat{h}(x) &= \widehat{F}_A^{-1}(\alpha F_W(x) + \beta) \\ &= F_A^{-1}\left(\frac{(\alpha F_W(x) + \beta) - \beta}{\alpha}\right) \\ &= F_A^{-1}(F_W(x)) = h(x). \end{aligned}$$

where the second equality follows from (10) because here $\alpha F_W(x) + \beta < \alpha F_W(q(\sigma_W^*)) + \beta = F_W(q(\sigma_W^*)) = F_A(q(\sigma_A^*))$ (the first equality follows from the definition of β, and the second from the equilibrium conditions). This shows that $\widehat{h}(x) = h(x)$ for all x. ∎

Proposition 5 suggests a method which, under the assumption of our model, can be employed to recover the QQ plot of potential legal earnings starting from the distributions of reported earnings. It suffices to add to the sample the proportion of people who are not sampled because incarcerated, and count them as reporting zero income. Since we assume that the "lucky" criminals (who are in our sample) are already reporting zero income, this modification achieves a situation in which all citizens who are criminals report earnings of zero. The modified sample is then consistent with the assumptions of Proposition 5, and so yields the correct QQ plot even though the distributions of reported income suffer from selection.[18]

As an illustration of this methodology we can compute the QQ plot based on data from the March 1999 CPS supplement. We take the c.d.f.'s of the yearly earnings distributions of males of age 15-55 who reside in Metropolitan Statistical Areas of the US.[19] Figure 1 presents the c.d.f.'s of reported earnings. It is clear that the stochastic dominance requirement is satisfied. For each race we then add a fraction of zeros equal to the percentage of males who are incarcerated, and then compute the QQ plot.[20] Figure 2 presents the QQ plot between earning levels of zero and $100,000:[21] This picture suggests that the stretch requirement is likely to be satisfied, except perhaps in an interval of earnings for whites between 10,000 and 15,000. In order to get a derivative of the QQ plot that is stable, we first smooth the p.d.f.'s and then use them to construct a smoothed QQ plot.[22] Figure 3 presents the derivative of the

[18] Notice that, to implement this procedure, it is not necessary to know α or σ^*_R. In contrast, knowledge of these quantities would be necessary to recover the c.d.f.'s of legal earning opportunities from the c.d.f.'s of reported earnings.

[19] Excluding the disabled and the military personnel. Earnings are given by the variable PEARNVAL and include total wage and salary, self-employment earnings, as well as any farm self-employment earnings.

[20] Of 100,000 African-American (white) citizens, 6838 (990) are incarcerated according to data from the Department of Justice.

[21] We ignore those with legal earnings opportunity above $100,000 as these individuals are few and the forces in our model are unlikely to accurately describe their incentives to commit crimes.

[22] The smoothed p.d.f.'s are computed using a quartic kernel density estimator with a bandwidth of

smoothed QQ plot. The picture suggest that the derivative of the QQ plot tends to be smaller than one at earnings smaller than $100,000 for whites. The derivative approaches one around earnings of $12,000 for whites. It is also possible to get a sense of whether condition (8) in Proposition 3 is satisfied by plotting the expression $h'(x) f_A(h(x)) / \min_{y\in[h(x),x]} f_A(y)$. Whenever this fraction is smaller than one, that condition is satisfied. Figure 4 plots this fraction as a function of x.

8 Disrepute As a Cost of Being Searched

In this paper we posit that it is desirable to equalize search intensities across groups. This principle rests on the idea that being searched by police entails a cost of time or distress, and that this cost should be born equally (in expectation) by citizens of all groups.[23] It may seem harmless to extend this interpretation to costs that relate to the outside perception of the individual being searched. This is to capture the stigmatization, shame, or *disrepute*, that is attached to being singled out for search by the police. The idea is that being publicly searched entails a cost because other people (bystanders, neighbors, etc.) use the search event to infer something about the criminal propensity of the individual who is subject to search.[24] In this acception, disrepute is not a a primitive of the model. Rather, how much disrepute is entailed with being searched depends on the search strategy that police use.

The fact that the loss of reputation is endogenous to the model opens up the possi-

$9,000.

[23] These costs encompass "objective" and "subjective" intrusions, as defined in *Michigan Dep't of Police v. Sitz*, 496 U.S. 444 (1990).

[24] Slobogin (1991) calls this notion "false stigmatization" and says that "... the innocent individual can legitimately claim an interest in avoiding the stigma, embarrassment, and inconvenience of a mistaken investigation." This interest is listed, along with that of protection against unjustified government infringement of "privacy" and "autonomy," and that of avoiding government action designed to harass, as the three key interest of citizens in avoiding search and seizure. (p. 7.)

bility that, by employing different search strategies, the police might reduce the total amount of discredit in the economy and achieve a Pareto improvement. Alternatively, it seems possible that by altering their search strategy, the police might transfer some of the burden across racial groups. However, we argue that neither possibility is true.

To flesh out the argument, let us establish a minimal model in which it is meaningful to talk about disrepute. In addition to the observable characteristic race, there must be some characteristic which is unobservable to the bystanders but observable to police, and this characteristic must be correlated with criminal behavior. Denote this characteristic by c, which has some distribution $\Pr(c|R)$ in each race. For concreteness, we refer to this characteristic as the criminal record. Since the police condition their search behavior on c, upon seeing an individual being searched, the bystanders would update their prior about the c value of the individual being searched, and hence about his criminal propensity. Denote with $\sigma(c, R)$ the probability of being searched of a random individual with criminal record c and race R.

We define the *disrepute* η of an individual with race R as

$$\eta(R) = \Pr(R \text{ is criminal}) - \Pr(R \text{ is criminal} \mid \text{searched}),$$

where the event "R is criminal | searched" denotes the event that a randomly chosen individual of race R is criminal conditional on his being searched.[25] Thus, we define the cost of disrepute as the amount of updating that an observer does about an individual of race R upon seeing that he is being searched. This definition captures a notion of the humiliation connected with being publicly searched.

Consistent with the notion of updating, we must also take into account the fact that an observer updates about a random individual also when that individual is *not* searched. Individuals who are not searched experience a corresponding increase in status. We term this effect *negative disrepute* and we denote it by θ. Formally,

$$\theta(R) = \Pr(R \text{ is criminal}) - \Pr(R \text{ is criminal} \mid \text{not searched}).$$

[25] There is nothing special about the event "R is criminal." The argument in this section applies equally to any event of the form "the individual of race R has a value of $c \in E$," where E is any set.

The total disrepute effect (positive and negative) in race R is

$$\eta(R) \cdot \sigma(R) + \theta(R) \cdot (1 - \sigma(R)) = 0,$$

where $\sigma(R) = \sum_c \sigma(R, c) \Pr(c|R)$ denotes the probability that a randomly chosen individual of race R is searched. Equality to zero reflects the fact that posterior probabilities are martingales. This means that, in terms of total humiliation, the search strategy is *neutral within race* R. No matter what the search strategy, the disrepute effect has constant sum within a race.[26] Because changing the search strategy cannot change the average disrepute within a race (which is always zero), a notion of fairness that prescribes equating costs of being searched across races cannot be based on the cost of shame.

This is not to say that the prior $\Pr(R \text{ is criminal})$ is unaffected by the search strategy σ. As we have seen, increasing $\sigma(R)$ decreases the probability that members of group R are criminals. What this neutrality result says is that, when we evaluate the effects of a certain search strategy, we can ignore the distributive effects of disrepute across races.

9 Discussion of Modeling Choices

The objective function of police An important assumption in our model is that police choose whom to search so as to maximize the probability of catching criminals. We believe this assumption on police behavior is realistic; the notion that police wish to maximize successful searches captures an important aspect of real-world career incentives for police.[27] This statement is supported by the explicit reference to this

[26]Notice that this argument does not rely on any assumption on behavior of either police or citizens. It is purely a point about Bayesian updating.

[27]In our model we simplify by assuming that this is the only determinant of police incentives, but this is not essential for the result that the equilibrium is inefficient: what matters is that there is some such incentive.

point in the Verniero and Zoubek (1999), as well as by anecdotal evidence.[28] From a more theoretical standpoint, rewarding officers based on their performance is a natural way of dealing with the problem of unobservable effort on the part of police. In a world where an individual police officer's effort is too small to measurably affect the aggregate crime level, rewarding successful searches is the natural way to induce police to exert effort.

The fact that the efficient (crime-minimizing) outcome is not achieved does not mean that it cannot be achieved in this setup. The efficient allocation *can* be implemented, at the cost of giving race-dependent incentives to police. That is, efficiency will generally require giving police higher rewards for successful busts on citizens of a particular race. However, it is doubtful that it would be politically feasible, or ethically desirable, to set up such incentive schemes. In addition, as shown in Section 4, the efficient outcome may be very unfair.

Non-racially biased police In this paper we assume that police are not racist, i.e., that they do not derive greater utility from searching a member of group A than a member of group W. We do this for two reasons: First, because we are interested in how the system of police incentives (maximizing successful searches) results in disparate treatment, and the theoretically clean way to address this question is to abstract from any bigotry. Second, because at least in some practical instances of alleged disparate impact, the unfairness is ascribed to the incentive system, and not directly to racism of the police; this is the position in Verniero and Zoubek (1999) (see Part III D), and is also consistent with the findings in Knowles, Persico, and Todd (2000). Therefore, we think it is interesting to study our problem abstracting from the issue of bigotry. If police were racially biased, then that would be an additional factor determining the equilibrium but it would not necessarily make it more efficient to implement fairness (although one

[28]See for instance Verniero and Zoubek (1999), p. 43, for a discussion of the incentive effects of the "Trooper of the Year" Award, given to the trooper of the NJ police who made the most drug arrests and the largest drug seizures.

could argue that it would make fairness more desirable on moral grounds).

Notions of efficiency of interdiction We have taken the position that efficiency of interdiction is measured by the total number of citizens who commit crimes. According to this measure, given a total number of crimes, efficiency of interdiction is the same whether most of the crimes are committed by one racial group or equally by both racial groups. However, one may argue that if most of the crimes are commited by one racial group, then members of that group are also disproportionately going to be the targets of crime, and that therefore one should also try to reduce the inequality in the distribution of crime across groups. This is a valid argument which suggests some interesting questions, such as how much inequality in the crime rates should be tolerated across groups. Keeping the crime rate completely homogeneous across groups necessitates policing different groups with different intensity, which opponents of racial profiling find inappropriate. How to strike an appropriate balance between considerations of fairness in policing and disparities in crime rates across groups is an interesting question which we leave for future research.

10 Conclusion

The controversy on racial profiling focusses on the fact that, in deciding whom to search, police use race as a proxy for other unobservable characteristics. The resulting racial disparity is unfortunate, but it may be legal (e.g., consistent with the 4th amendment) if it is an unavoidable by-product of effective policing. In this paper we have presented the minimal model that allows us to investigate the issue. We have shown that the two goals of policing—fairness and effectiveness—are not necessarily in contrast; sometimes, forcing the police to behave more fairly can increase effectiveness of interdiction. We have given exact conditions under which, within our simple model, such a trade-off is present. These conditions are based on the distributions of legal earning opportunities of citizens, and are expressed as contraints on the QQ plot of these distributions. Thus, our model relates the trade-off between fairness and efficiency of policing to income disparities across races. In our model it is possible to recover the QQ plot of the distributions of legal earning opportunities using reported earnings, so our approach has the potential to deliver quantitative implications. Finally, we have suggested that a redistributive notion of fairness (all citizens should bear the same expected cost of being searched) may not be meaningful when the costs of being searched derives from the stigmatization connected with being singled out for search.

This paper has dealt with a very simple model where crime is endogenous and the decision to be a criminal reflects lack of legal earning opportunities. To highlight the basic forces in the simplest way, we chose to focus on few modeling elements. Of course, any realistic situation would be richer in detail than this stylized model, and a number of additional considerations may be relevant in each case. In specific situations of policing, more modeling structure may enrich the insights of the simple model presented here.

Appendix

Proof of Proposition 3.

Proof: Since F_W first-order stochastically dominates F_A, we have $\sigma_W^* \leq \overline{\sigma} \leq \sigma_A^*$. To construct the total variation in crime from implementing the fair outcome, write

$$
\begin{aligned}
& F_A\left(\overline{\sigma}\left(J-H\right)+H\right)-F_A\left(\sigma_A^*\left(J-H\right)+H\right) \\
= & -\int_{\overline{\sigma}}^{\sigma_A^*} \frac{\partial F_A\left(s\left(J-H\right)+H\right)}{\partial s} ds \\
= & \left|J-H\right| \int_{\overline{\sigma}}^{\sigma_A^*} f_A\left(q\left(s\right)\right) ds.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
& F_W\left(\sigma_W^*\left(J-H\right)+H\right)-F_W\left(\overline{\sigma}\left(J-H\right)+H\right) \\
= & \left|J-H\right| \int_{\sigma_W^*}^{\overline{\sigma}} f_W\left(q\left(s\right)\right) ds \\
= & \left|J-H\right| \int_{\sigma_W^*}^{\overline{\sigma}} \left[h'\left(q\left(s\right)\right) f_A\left(h\left(q\left(s\right)\right)\right)\right] ds.
\end{aligned}
$$

The total variation in crime (expression 9) reads

$$
TV = \left|J-H\right| \cdot \left[N_A \int_{\overline{\sigma}}^{\sigma_A^*} f_A\left(q\left(s\right)\right) ds - N_W \int_{\sigma_W^*}^{\overline{\sigma}} \left[h'\left(q\left(s\right)\right) f_A\left(h\left(q\left(s\right)\right)\right)\right] ds\right].
$$

Now, denoting $m = \min_{\sigma \in \left[\overline{\sigma}, \sigma_A^*\right]} f_A\left(q\left(\sigma\right)\right)$, the first integral in the above expression is greater than $\int_{\overline{\sigma}}^{\sigma_A^*} m\, ds$, and so

$$
\begin{aligned}
TV & \geq \left|J-H\right| \cdot \left[N_A\left(\sigma_A^*-\overline{\sigma}\right) m - N_W \int_{\sigma_W^*}^{\overline{\sigma}} \left[h'\left(q\left(s\right)\right) f_A\left(h\left(q\left(s\right)\right)\right)\right] ds\right] \\
& = \left|J-H\right| \cdot \int_{\sigma_W^*}^{\overline{\sigma}} \left[N_A \frac{\sigma_A^*-\overline{\sigma}}{\overline{\sigma}-\sigma_W^*} m - N_W h'\left(q\left(s\right)\right) f_A\left(h\left(q\left(s\right)\right)\right)\right] ds \\
& = \left|J-H\right| \cdot \int_{\sigma_W^*}^{\overline{\sigma}} \left[N_W m - N_W h'\left(q\left(s\right)\right) f_A\left(h\left(q\left(s\right)\right)\right)\right] ds \qquad (11)
\end{aligned}
$$

where to get the last equality we used the identity $N_W\left(\overline{\sigma}-\sigma_W^*\right) = -N_A\left(\overline{\sigma}-\sigma_A^*\right)$. To verify this identity write

$$
N_A\left(\overline{\sigma}-\sigma_A^*\right)
$$

$$
\begin{aligned}
&= N_A \left(\frac{\sigma_A^* N_A + \sigma_W^* N_W}{N_A + N_W} - \sigma_A^* \right) \\
&= \frac{N_A N_W}{N_A + N_W} \left(\sigma_W^* - \sigma_A^* \right)
\end{aligned}
$$

and thus, symmetrically

$$
\begin{aligned}
N_W \left(\overline{\sigma} - \sigma_W^* \right) &= \frac{N_A N_W}{N_A + N_W} \left(\sigma_A^* - \sigma_W^* \right) \\
&= -N_A \left(\overline{\sigma} - \sigma_A^* \right).
\end{aligned} \tag{12}
$$

Now, let us return to expression (11); from the definition of m we have

$$
\begin{aligned}
m &= \min_{y \in \left[q\left(\sigma_A^* \right), q(\overline{\sigma}) \right]} f_A(y) \\
&= \min_{y \in \left[h\left(q\left(\sigma_W^* \right) \right), q(\overline{\sigma}) \right]} f_A(y),
\end{aligned}
$$

where to get from the first to the second line we use the fact that the equilibrium σ_A^*, σ_W^* is interior, as shown in the proof of Lemma 1. Now, fix any $s \in [\sigma_W^*, \overline{\sigma}]$. Since $s \geq \sigma_W^*$ and $q(\cdot)$ is a decreasing function we have $q(s) \leq q(\sigma_W^*)$ and so $h(q(s)) \leq h(q(\sigma_W^*))$. Also, since $s \leq \overline{\sigma}$ we have $q(s) \geq q(\overline{\sigma})$. Thus,

$$
\text{for any } s \in [\sigma_W^*, \overline{\sigma}] \text{ we have } m \geq \min_{y \in [h(q(s)), q(s)]} f_A(y) \;.
$$

Substituting for m into expression (11) we get

$$
TV \geq |J - H| \, N_W \cdot \int_{\sigma_W^*}^{\overline{\sigma}} \left[\min_{y \in [h(q(s)), q(s)]} f_A(y) - h'(q(s)) \, f_A(h(q(s))) \right] ds,
$$

which is positive under the assumptions of the proposition. This shows that crime increases as a result of implementing the completely fair outcome. ∎

References

[1] *Drug Policy and the Criminal Justice System: Fact Sheet* The Sentencing Project, available online at www.sentencingproject.org/brief/5047.htm

[2] Anderson, Elijah *StreetWise: Race, Class, and Change in an Urban Community.* The University of Chicago Press, Chicago 1990.

[3] Arnold, Barry C. *Majorization and the Lorenz Order: A Brief Introduction.* Lecture Notes in Statistics, 43, Springer-Verlag Berlin, Heidelberg 1987.

[4] Beck, Phyllis W., and Patricia A. Daly "State Constitutional Analysis of Pretext Stops: Racial Profiling and Public Policy Concerns." *Temple Law Review* 72, Fall 1999, p. 597.

[5] Becker, Gary S. (1968) "Crime and Punishment: An Economic Approach" *Journal of Political Economy* 76(2) Mar.-Apr. 1968, 169-217.

[6] Coate, Stephen; Loury, Glenn C. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review.* Vol. 83 (5). p 1220-40. December 1993.

[7] Ehrlich, Isaac (1973) "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation" *Journal of Political Economy* 81(3) May-Jun. 1973, 521-565.

[8] Ehrlich, Isaac. "Crime, Punishment, and the Market for Offenses." *Journal of Economic Perspectives* Vol. 10 (1). p 43-67. Winter 1996.

[9] Gibeaut, John "Profiling: Marked for Humiliation," *American Bar Association Journal*, Feb. 1999, pp. 46-7. Available online at http://www.abanet.org/journal/feb99/02RPROFI.HTML.

[10] Kennedy, Randall. *Race, Crime and the Law*, New York: Pantheon Books, 1977.

[11] Knowles, John, Nicola Persico, and Petra Todd "Racial Bias in Motor-Vehicle Searches: Theory and Evidence." *The Journal of Political Economy* 109(1), February 2001, pp. 203-229.

[12] Lehmann, E. "Comparing Location Experiments," *The Annals of Statistics*, 16 (1988), 521-533.

[13] Levitt, Steven D. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime" *The American Economic Review* Vol. 87 (3). p 270-90. June 1997

[14] Mailath, George J; Samuelson, Larry; Shaked, Avner. "Endogenous Inequality in Integrated Labor Markets with Two-Sided Search." *American Economic Review* Vol. 90 (1). p 46-72. March 2000.

[15] Mauer, Mark "Intented and Unintended Consequences: State Racial Disparities in Imprisonment," The Sentencing Project (1997), available online at www.sentencingproject.org/policy/9050.htm

[16] Moro, Andrea and Peter Norman "Affirmative Action in a Competitive Economy" Mimeo, University of Minnesota 2000.

[17] Norman, Peter "Statistical Discrimination and Efficiency," Mimeo, University of Wisconsin, June 2000.

[18] Polinsky, A. M. and S. Shavell, "The Fairness of Sanctions: Some Implications for Optimal Enforcement Policy," Harvard Olin Center Discussion Paper no. 247, 12/98.

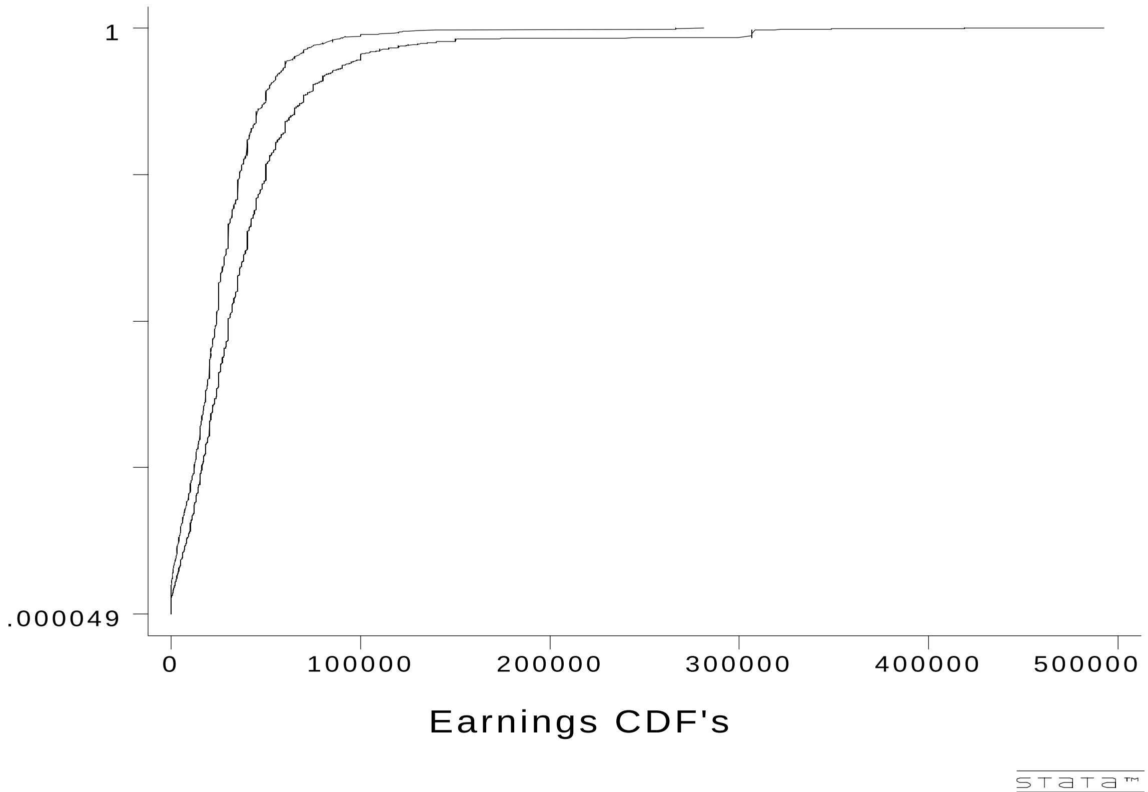
[19] Slobogin, Christopher "The World Without a Fourth Amendment." *UCLA Law Review* 1, October 1991, p.39.

[20] Thompson, Anthony C. "Stopping the Usual Suspects: Race and the Fourth Amendment." New York University Law Review 74, October 1999, p. 956.

[21] Verniero, Peter, and Paul Zoubek *Interim Report of the State Police Review Team Regarding Allegations of Racial Profiling.* April 20, 1999.

[22] Wilk, M. B. and R. Gnanadesikan "Probability Plotting Methods for the Analysis of Data" *Biometrika* Vol. 55, No. 1. (Mar., 1968), pp. 1-17.

[23] Wohlstetter, Albert and Sinclair Coleman. "Race Differences in Income." Rand Publication R-578-OEO, October 1970. (Available on-line at http://www.rand.org/publications/classics/wohlstetter/R578/).



1
.000049
0
100000
200000
300000
400000
500000
Earnings CDF's

Figure 1

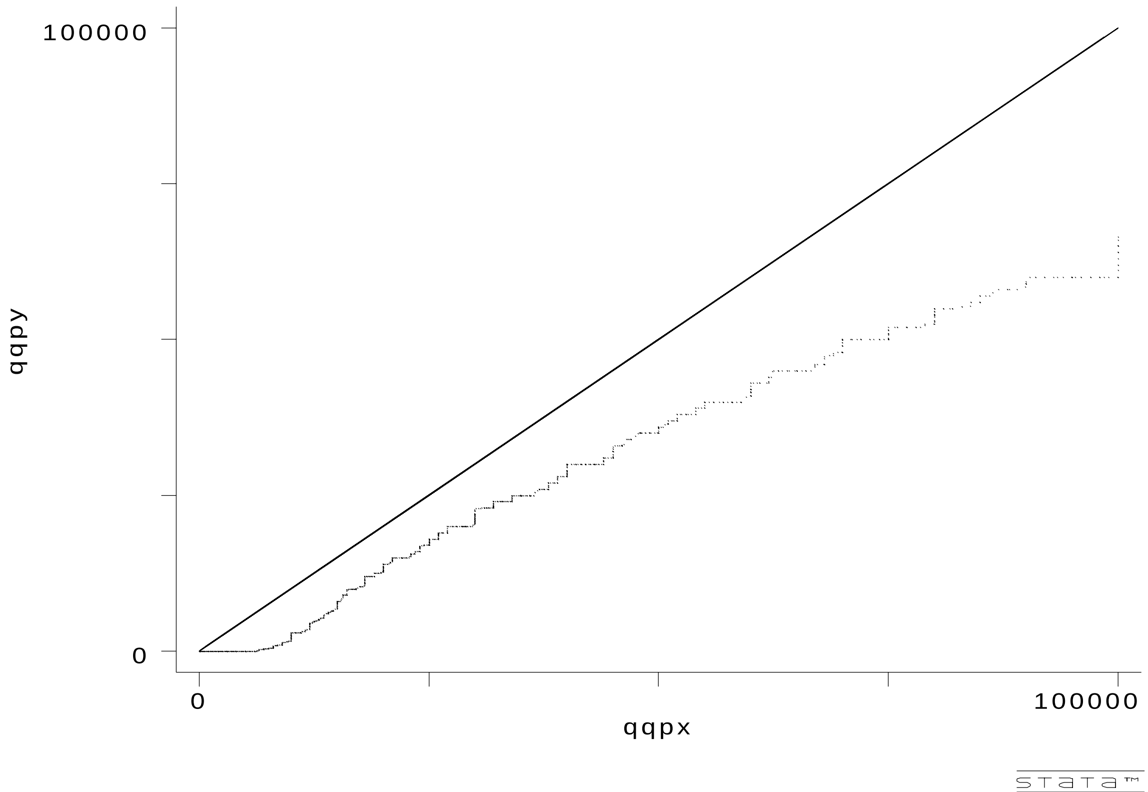


Figure 2

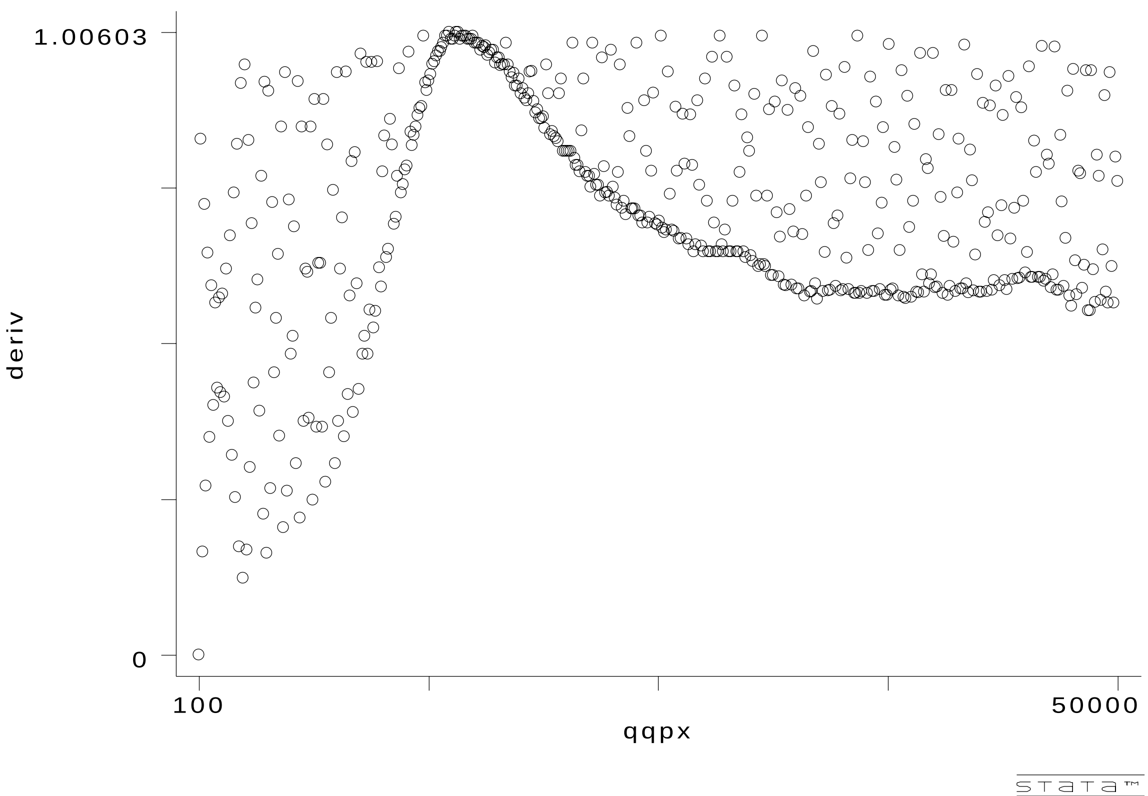


Figure 3

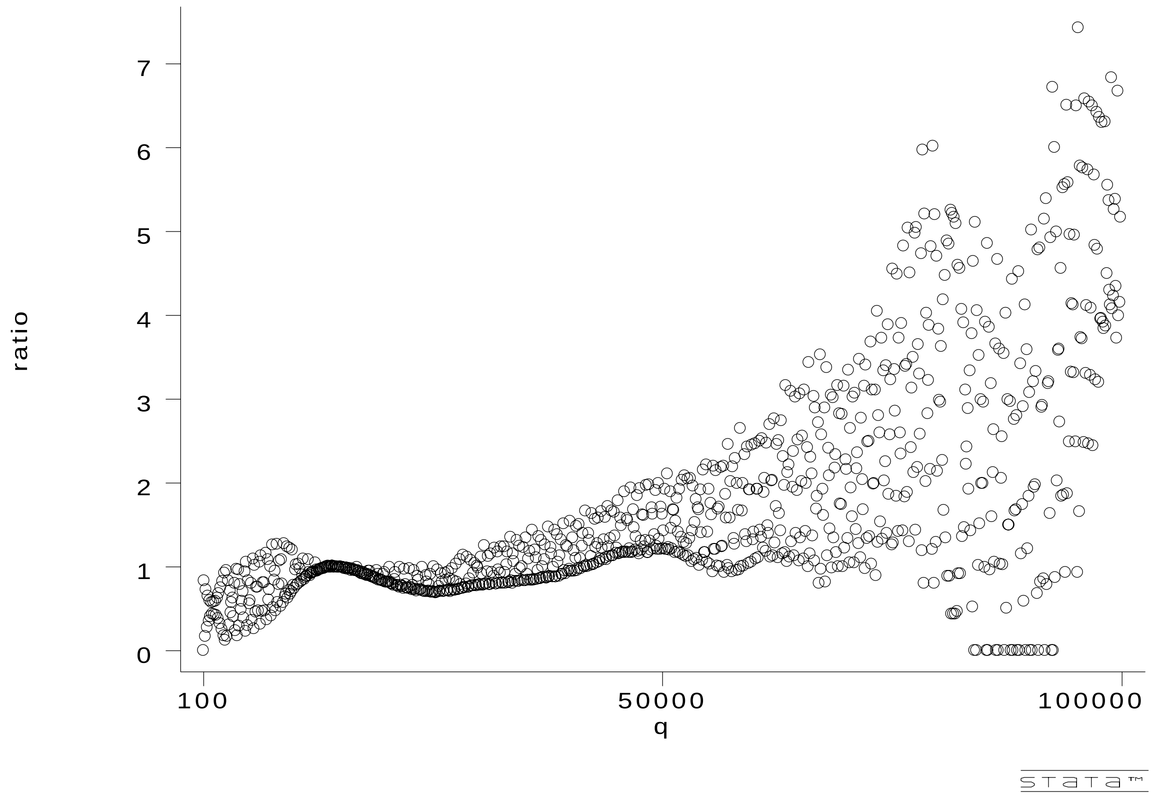


Figure 4