Racial Profiling, Fairness, and Effectiveness of Policing

By NICOLA PERSICO*

Citizens of two groups may engage in crime, depending on their legal earning opportunities and on the probability of being audited. Police audit citizens. Police behavior is fair if both groups are policed with the same intensity. We provide exact conditions under which forcing the police to behave more fairly reduces the total amount of crime. These conditions are expressed as constraints on the quantile–quantile plot of the distributions of legal earning opportunities in the two groups. We also investigate the definition of fairness when the cost of being searched reflects the stigma of being singled out by police. (*JEL* J71, J78)

Law enforcement practices often have disparate impact on different ethnic and racial groups. To take one well-publicized example, 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 in connection with customs searches at airports, and in a number of other situations involving policing.[1] Such racial disparities are an undeniable reality of policing and are widely perceived to be discriminatory. Public scrutiny 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.[2]

However, forbidding the police from using some characteristics may reduce the effectiveness of policing.[3] Often, those who engage in certain criminal activities tend to share certain characteristics relating to specific socioeconomic and ethnic backgrounds. Forcing the police to disregard such characteristics may lead to less effective policing and to increased crime.[4]

These conflicting considerations reflect a fundamental tension between the principle of equal treatment under the law and the practical demands of law enforcement. This tension colors the current judicial position, which tends to uphold the use of racial and other characteristics

* Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104 (e-mail persico@ssc.upenn.edu) URL: ⟨http://www.ssc.upenn.edu/~persico⟩. This work develops a line of inquiry that stems from joint work with John Knowles and Petra Todd; I am grateful to both of them. Thanks also to Aldo Colussi for research assistance. The input of several anonymous referees helped improve the paper. The financial support of the National Science Foundation, under Grants SBR-9905564 and SES-0078870, is gratefully acknowledged.

[1] See John Knowles et al. (2001) for data that are representative of the disparity in highway searches. Concerning airport searches, John Gibeaut (1999) reports that, of all passengers who were strip-searched by customs officials in 1997, 60 percent were black or Hispanic. 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 Elijah Anderson (1990) for a more general view of the relationship between the police and African-Americans.

[2] In response to charges of racial profiling against New Jersey police, the former attorney general of New Jersey 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 Peter Verniero and Paul H. Zoubek, 1999, pp. 52–53). 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.

[3] See Phyllis W. Beck and Patricia A. Daly (1999) for a discussion of the public policy argument in favor of discretional searches of motorists.

[4] There are two logically distinct scenarios under which members of one racial group may be searched disproportionally often by racially unbiased police. One is that race is correlated with other characteristics which are unobservable to police and are correlated with criminal behavior; in this case, race constitutes a proxy for other characteristics. The second scenario is that citizens are searched based on other (nonrace) characteristics which are observable to police and are more frequently present in that race. The results presented in this paper are valid under both scenarios; although the model is formally stated to describe the proxy scenario, it can readily be reinterpreted to account for the other scenario.

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.[5] Mindful of the established judicial line and of public policy concerns for effectiveness of interdiction, the courts have generally been cautious in imposing remedies on the police. In so doing, the courts implicitly have placed less weight on the plight of innocent citizens who bear a disproportionate amount of the policing effort. This position is defensible only insofar as the two considerations, fairness and effectiveness of policing, are in conflict. It is therefore extremely important to establish whether a conflict exists in practice.

This paper studies the trade-off between fairness and effectiveness of interdiction in the context of a simple model in which citizens of two groups choose whether to engage in illegal activities, and police audit (or search) citizens. The model is too stylized to have direct policy implications. Despite this fact, the model offers a more sophisticated view than is generally aired in the public debate. The first point of this paper is that, in principle, there need not be a trade-off. Constraining police to behave in a fair manner—in our context, searching both groups with the same intensity—need not per se result in more crime. This is because, under very plausible assumptions about the incentives of police, the equilibrium allocation does not coincide with the crime-minimizing allocation, and so placing constraints on the behavior of the police may actually increase the effectiveness of policing.[6]

To understand this statement it is necessary to understand why in equilibrium the effectiveness of policing is generally not maximal. 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, for the sake of the argument, that police can distinguish only two groups, A and W, which partition the population. 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. But then members of that group would respond by stepping up their illegal activity, and then police would want to search the group. This equilibrating process results in the key equilibrium condition: in equilibrium the two groups must have the same fraction of criminals.[7] Crucially, this equilibrium condition is independent of the elasticity of crime to policing. In contrast, the elasticity to policing plays a major role in determining the conditions for effective policing. To see this, take any allocation (including the equilibrium one), and suppose an increase of 1 percent in the search probability deters members of group W more than members of group A; then, maximal effectiveness 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 effective allocation of policing effort.[8]

[5] See Randall Kennedy (1997) for an exposition of this view, and more recently Anthony C. Thompson (1999) for a critical analysis. It may seem surprising that the Supreme Court would sanction such a narrow interpretation of the equal protection clause. However, racial profiling and other policing cases are often tried as fourth amendment cases (protection from unreasonable search and seizure) and not as fourteenth amendment cases (protection from racial discrimination). The distinction is crucial because the standard of evidence is lower in fourth amendment cases: police are only required to prove "reasonableness" of their behavior, instead of being subject to the "strict scrutiny" standard that applies in racial discrimination cases. The Supreme Court 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 [see *U.S.* v *Whren,* 517 U.S. 806, 116 S.Ct. 1769 (1996)]. The appropriateness of applying fourth amendment law to racial profiling cases is now hotly debated, and intense public scrutiny is changing the legal landscape in this area.

[6] In this paper effectiveness of interdiction is synonimous with "crime minimization." This point is discussed further in Section I, subsection A, in terms of the generality of the assumptions under which effectiveness and fairness do not coincide.

[7] Equalization of fraction guilty is predicted only for those crimes which are the object of interdiction. Furthermore, equalization of crime rates does not imply that equilibrium incarceration rates are the same for the two groups: the group which is searched more intensely will have a higher fraction of its members being incarcerated. Finally, the crime rate is equalized only among those groups that are searched in equilibrium. These points are discussed in subsection A.

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

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. Suppose in equilibrium group W is policed with less intensity; then, forcing the police to behave more fairly entails increasing the intensity with which group W is policed. Since group W is more responsive to policing, this increases the effectiveness of interdiction. 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: in this case, there is no trade-off between fairness and effectiveness.

In our model, then, constraining police may or may not increase the effectiveness of interdiction, depending on the relative elasticity to policing in the two groups. We thus seek theoretical conditions to compare that elasticity across the two groups. Measuring the elasticity directly is a difficult exercise, which at a minimum requires some sort of exogenous variation in policing (see Steven D. Levitt, 1997, who also reviews the literature). We outline an alternative approach based on measuring the primitives of our simple model of policing, specifically, legal earning opportunities. The reason why earning opportunities matter is that, given a certain intensity of policing, citizens whose legal earning opportunities are higher are less likely to engage in crime. In our stylized formalization, citizens above a given legal earning threshold will not engage in criminal activities, and those below the threshold will. Thus the amount of criminal activity—and hence also the elasticity of crime to policing—depends on the distribution of legal earning opportunities. The advantage of this approach is that it does not require observing any variation in policing. Instead, the answer to the main question—whether the group which in equilibrium is policed more intensely is also the group in which the elasticity of crime to policing is higher—lies in the shape of the distributions of legal earning opportunities in the two groups. Mathematically, the key condition is expressed as a constraint on the QQ (quantile–quantile) plot of these distributions. The QQ plot $h(x)$ is easily constructed; given the cumulative distribution functions (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 suggested above, in these circumstances moving slightly towards fairness reduces the effectiveness of policing. A similar condition allows us 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.

At this point, it is important to emphasize that our modeling strategy will lead us to abstract from many factors that can affect the decision to engage in crime, and focus on one—the legal alternative to crime. By isolating this factor, our model enables us to make progress with the theoretical conditions that link fairness and effectiveness of policing. However, many other factors also play important roles in the question we analyze. One such factor is the intensity of punishment that citizens of different groups incur if found guilty. If, for example, citizens of group A expect to receive longer prison sentences if convicted relative to citizens of group W, the deterrence effect of policing will be higher within group A even if that group had the same legal earning opportunities as group W. Similarly, our conclusions are modified if the objective function of police is not simply to catch criminals, but reflects some concern for the overall crime level. These generalizations are investigated in Section VIII. Extending the model in these directions changes the form of some results, but the QQ plot remains the key statistic in the analysis. Therefore, it is interesting to seek empirical equivalents for the QQ plot of legal earning opportunities.

Although an agent's legal earning opportunity may be unobservable if he/she decides to engage in crime and forgo legal occupation, and so F_A and F_W are in principle unobservable, we do have surveys of earnings reported by those citizens who are not incarcerated. We give conditions under which the QQ plot based on surveys of reported earnings coincides, in the equilibrium of our stylized model, with the QQ plot of *potential* legal earnings, which is the object of interest. Purely as an illustration of this methodology, and with no direct policy application in mind, we construct the QQ plot based on empirical distributions of reported earnings by African-American and white male residents of metropolitan statistical areas in the United States. Intriguingly, we find that this QQ plot indeed appears to have slope smaller than 1 (although the slope appears to be quite close to 1 for plausible values of the parameters). In the context of our model, a slope smaller than 1 would imply that the elasticity to policing at equilibrium is higher in the A group than in the W group, suggesting that if one were to move toward fairness by constraining police to shift resources from the A group to the W group, then the total amount of crime would not fall. Thus, our back-of-the-envelope calculation cannot rule out the presence of a trade-off between fairness and effectiveness of policing. Of course we do not imply that our stylized model is a close approximation of the real-world problem: the purpose of the calculation is to illustrate how the methodology can be applied to the data.

Finally, we explore the foundations of our notion of fairness through an inquiry on the nature of the cost of being searched. The notion of fairness adopted in this paper focuses on equating across groups the expected costs of being searched. This notion of fairness speaks to the concerns of critics of racial profiling, and embodies the idea that society should guarantee equal treatment to citizens of all racial groups. This notion is redistributional in nature because it is based on the idea that society can (and should be prepared to) 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 formal definition of stigmatization, which we call *disrepute*. We define disrepute as the rational updating about an individual's characteristics due to the knowledge that the individual is singled out for search. The reason that disrepute is particularly salient among the various costs of being searched is that this component is not a primitive of the model. Because of this fact, police may be particularly limited in its ability to affect this component of the cost. To the extent that disrepute is perceived as an important component of the cost of being searched, a policy maker might conceivably attempt to influence its magnitude and/or its distributional impact. However, we show that the total amount of disrepute within a group is independent of the police's search behavior, and consequently, changes in interdiction strategy have no effect on the amount of disrepute suffered by any group. For the purpose of distribution across groups, therefore, our analysis suggests that we can ignore that part of the cost of being searched which is due to disrepute. Of course, this finding should not be taken to imply that a redistributional notion of fairness is unjustified. Disrepute is just one determinant of the cost of being searched. To the extent that the cost of being searched reflects "exogenous" costs, such as loss of time, or other inconvenience associated with being searched, it is perfectly reasonable to attempt to equate those costs across races. We view our result as cause for optimism, in the sense that it allows us to focus attention on these nonreputational components of the cost of being searched. Such costs are arguably within the control of a stringent remedial policy.

Related Literature

The formal economic literature on crime and policing is vast, starting from Gary S. Becker (1968) and Isaac Ehrlich (1973); we refer to Ehrlich (1996) for a survey. In this literature, the issue of fairness is generally posed in terms of relationship between crime and sanction. For example, in A. Mitchell Polinsky and Steven Shavell (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 et al., 2001) a model similar to the one used here was employed in an empirical analysis of highway searches for drugs. That paper asks whether the police force is racially prejudiced, and it does not address issues of fairness.

We study 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 (Stephen Coate and Glenn C. Loury, 1993; George Mailath et al., 2000; Andrea Moro and Peter Norman, 2000; 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 Kenneth J. Arrow's (1973) 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 unique equilibrium, and so our comparative statics results are not subject to the multiple-equilibria criticism.

There is a literature on the effect of penalties on the crime rate. This literature generally relies on exogenous variation in the interdiction effort or in the penalty to estimate the elasticity of crime to interdiction (see Levitt [1997] and Avner Bar-Ilan and Bruce Sacerdote [2001], who also provide a good review of the literature on this topic). To our knowledge this literature does not address the different elasticity of different racial groups to crime. An exception is Bar-Ilan and Sacerdote (2001), who find that as the fine is increased for running a red light in Israel, the total decrease in tickets is much larger for Jews than for non-Jews.

I. Model

The following model is a stylized representation of police activity and its impact on potential criminals. The model is designed to more closely portray that area of interdiction which is highly discretionary in nature, in that it is the police who choose whom to investigate, as opposed to low-discretion interdiction in which police basically respond to requests for intervention (for example, domestic violence or rape). Examples of high-discretion interdiction are highway searches of vehicles, "stop and frisk" neighborhood policing, and airport searches. Allegations of racial disparities are made mainly in reference to high-discretion interdiction.

We emphasize that attention is restricted to the type of crimes that are the object of interdiction: for example, transporting drugs in the case of vehicular searches. Thus, although in what follows we will sometimes for brevity refer to "crime" without qualifying its type, the model does not claim to have implications for aggregates such as the overall crime rate. The model does, however, have implications for the fraction of motorists found guilty of transporting drugs on highways (if the model is applied to vehicular searches).

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

There is a continuum of citizens partitioned into two groups, A and W, with measures N_A and N_W respectively. Throughout, the index $r \in \{A, W\}$ will denote the group. If a citizen engages in crime he receives utility (in monetary terms) of $H > 0$ if not searched and $J < 0$ if searched. If a citizen does not engage in crime, he is legally employed and receives earnings $x > 0$ independent of being searched.

Every citizen knows his own legal earning opportunity x, which is a realization from random variable X_r. Denote with F_r and f_r the c.d.f. and density of X_r. 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 groups. A citizen's legal earning opportunity x is unobservable to the police when they choose whether to search him. The police can, however, observe the group to which the citizen belongs.

In the aggregate there is a measure $\bar{S}$ of searches that is allocated between the two groups. Denote with S_r the measure of searches of citizens of group r. Feasibility requires that $S_A + S_W = \bar{S}$. Denote with $\sigma_r = S_r/N_r$ the search intensity in group r (i.e., the probability that a random individual of group r is searched). The feasibility constraint can be expressed as

$$N_A\sigma_A + N_W\sigma_W = \bar{S}. \quad (1)$$

Given a search intensity σ_r, a citizen of group r with legal earning opportunity x commits a crime if and only if

$$x < \sigma_r J + (1 - \sigma_r)H$$
$$= \sigma_r(J - H) + H \equiv q(\sigma_r).$$

The function $q(\sigma)$ represents the expected reward of being a criminal when the search intensity equals σ. The fraction of citizens of group r that engages in criminal activity is $F_r(q(\sigma_r))$.

We now give our definition of effectiveness. Consistent with the objective of this paper, we use a definition of effectiveness that captures *effectiveness of interdiction*.[9] When defining effectiveness of interdiction we should account for the costs and benefits of the interdiction activity. In our model the cost of interdiction is $\bar{S}$ 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 effective* 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 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 designed 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 innocent citizens who are searched and that it is desirable to equalize the expected cost across races.[10] This definition is essentially comparative, since it focuses on the predicament of two citizens who belong to different 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).

II. Equilibrium and Effectiveness of Interdiction

A. *Equilibrium Conditions*

We use a superscript asterisk (*) to denote equilibrium quantities. A police officer who maximizes the number of successful searches will search only members of the 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 and so both groups must have the same fraction of criminals. Thus, at an interior equilibrium (σ^*_A, σ^*_W), the following condition must be verified:

$$F_A(q(\sigma^*_A)) = F_W(q(\sigma^*_W)). \tag{2}$$

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) are uniquely determined in equilibrium. The condition that pins them down is that, if a group is searched by the police, then it must have a fraction of criminals at least as large as the other group. Leaving aside uninteresting constellations of parameters,[11] this condition uniquely pins down the equilibrium 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 groups are searched with positive probability. From the empirical viewpoint, this is a reasonable restriction. To

[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.

[10] More on this in Section VII.

[11] If $\bar{S}$ is very small or very large 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.

guarantee interiority of equilibrium in the forthcoming analysis, we will henceforth implicitly maintain the following assumption.

ASSUMPTION 1:

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

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

An equilibrium prediction of our model is that interdiction will be more intense for the group with more modest legal earning opportunities. Indeed, suppose that the distribution of legal earning opportunities of one group, say group W, first-order stochastically dominate that of group A. Then $F_W(x) \leq F_A(x)$ for all x, so equation (2) requires $\sigma_A^* \geq \sigma_W^*$. Another, tighter equilibrium prediction is given by equation (2) itself: if the equilibrium is interior the success rates of police searches should be equalized across groups. The equalization is a direct consequence of our assumption that police maximize successful searches.[12]

Both these equilibrium predictions are consistent with the evidence gathered in a number of interdiction environments. In the case of highway interdiction, Knowles et al. (2001) present evidence that on Interstate 95 African-American motorists are roughly six times more likely to be searched than white motorists, and that the success rate of searches is not significantly different between the two groups (34 percent for African-Americans and 32 percent for whites). In that paper, the success rate of searches is shown to be very similar (not significantly different in the statistical sense) across other subgroups, such as men and women (despite the fact that women are searched much less frequently), motorists searched at day and at night, drivers of old and new cars, and drivers of own vehicles versus third-party vehicles.

Similarly, in the case of precinct policing in New York City, Eliot Spitzer (1999) reports African-Americans to be over six times more likely to be "stopped and frisked" than whites, but the success rate of searches is quite similar in the two groups (11 percent versus 13 percent). Finally, in the case of airport searches, a report by U.S. Customs concerning searches in 1997–1998 reveals that 13 percent of those searched were African-Americans and 32 percent were whites,[13] and "6.7 percent of whites, 6.3 percent of Blacks, ... had contraband."[14] The evidence in these disparate instances of policing reveals a common pattern that is consistent with the predictions of our model. In particular, the finding that success rates are similar across racial groups is consistent with the assumption that police are indeed maximizing the success rate of their searches. We take this evidence as supportive of our modeling assumptions.

The finding of equal success rates of searches might seem to contradict the disparity in incarceration rates. However, equalization of success rates does not translate into equalization of incarceration rates, or probability of going to jail, across groups. To verify this, observe that those who are incarcerated in our model are the criminals who are searched by police. Thus the equilibrium incarceration rate in group r equals $\sigma_r^* F_r(q(\sigma_r^*))$. Since the term $F_r(q(\sigma_r^*))$ is independent of r [see equation (2)], the incarceration rate is highest in the group for which σ_r^* is highest (i.e., the group that is searched more intensely). To the extent that the interdiction effort is more focused on African-Americans (as is the case for highway drug interdiction), the model's implication for incarceration rates is consistent with the observed greater incarceration rates within the African-American population, relative to the white population.

Finally, we point out that in the equilibrium of our model it is possible that some groups have a lower crime rate than others. The model predicts equal crime rates only among the groups that are searched. To see this, consider a richer model with more than two groups (say, old grandmothers and young males of both groups), and in which old grandmothers are

[12] At an interior equilibrium individual police officers are indifferent about the fraction of citizens of group W that they search, and so they do not have a strict incentive to engage in racial profiling. The incentive is strict only out of equilibrium.

[13] See *Report on Personal Searches by the United States Customs Service* (U.S. Customs Service, p. 6)

[14] Cited from Deborah Ramirez et al. (2000, p. 10).

unlikely to engage in crime even if they are not policed at all. In such a model, police will only search young males of both groups and will not search grandmothers. The crime rate among grandmothers will be lower than among young males, and the success rates of searches will be equalized among all groups that are searched (young males of group A and W). In light of this discussion, it may be best to interpret equation (2) as equating the success rate of searches, instead of the aggregate crime rates.

B. *Maximal Effectiveness Conditions*

We denote maximal effectiveness (minimum crime) outcomes by the superscript "EFF." A social planner who minimizes the total amount of crime subject to the feasibility constraint solves

$$\min_{\sigma_A,\sigma_W} N_A F_A(\sigma_A(J-H)+H) + N_W F_W(\sigma_W(J-H)+H)$$

subject to $N_A\sigma_A + N_W\sigma_W = \bar{S}$, or, using the constraint to eliminate σ_W,

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

The derivative with respect to σ_A is

$$(3)\qquad N_A(J-H)[f_A(\sigma_A(J-H)+H) - f_W(\sigma_W(J-H)+H)]$$

and equating to zero yields the first-order condition for an interior optimum:

$$(4)\qquad f_A(q(\sigma_A^{\text{EFF}})) = f_W(q(\sigma_W^{\text{EFF}})).$$

This condition is necessary for maximal effectiveness. The densities f_A and f_W measure the size of the population in the two groups whose legal income is exactly equal to the expected return to crime. If the returns to crime are slightly altered, f_A and f_W represent the rates at which members of the two groups will switch from crime to legal activity or vice versa. We can think of f_A and f_W as elasticities of crime with respect to a change in policing. At an interior allocation that maximizes effectiveness, the elasticities are equal. Condition (4) is different from condition (2) and thus will generally not hold in equilibrium; therefore, the equilibrium allocation of policing effort is generally not the most effective.

III. 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 most effective allocation is very asymmetric and the symmetric (and hence completely fair) allocation *minimizes* effectiveness. This shows that the most effective allocation can be highly unfair, and that this feature does not depend on the heterogeneity of income distributions (i.e., on the difference between groups).[15] We interpret this result as an indication that there is little logical relationship between the concept of effectiveness and properties of fairness.

PROPOSITION 1: *Suppose $F_A = F_W \equiv F$. Then the equilibrium allocation is completely fair. If in addition F is a concave function on its support then the equilibrium allocation minimizes effectiveness, and at the most effective allocation one of the two groups is either not policed at all or is policed with probability 1.*

PROOF:

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

To verify the second statement, observe that,

[15] A similar conclusion that the efficient allocation can be unfair in a symmetric environment was reached by Norman (2000) in the context of a model of statistical discrimination in the workplace à la Arrow (1973). There, race-dependent (asymmetric) investment helps alleviate the inefficiency caused by workers who test badly and end up being assigned to low-skilled jobs despite having invested in human capital.

since f_A and f_W are decreasing, the only pair (σ_A, σ_W) that solves the necessary condition (4) for an interior most effective allocation together with the feasibility condition (1), is $\sigma_A^* = \sigma_W^*$. Consider the second derivative of the planner's objective function of Section II, subsection B, evaluated at the symmetric allocation,

$$N_A(J-H)^2\left[f_A'(\sigma_A^*(J-H)+H) + \frac{N_A}{N_W}f_W'\left(\frac{\bar{S}-N_A\sigma_A^*}{N_W}(J-H)+H\right)\right]$$

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

Recall that maximizing effectiveness requires minimizing the objective function. Since by assumption f_A' and f_W' are negative, the second-order conditions for effectiveness maximization are not met by (σ_A^*, σ_W^*). Indeed, the allocation (σ_A^*, σ_W^*) meets the conditions for effectiveness *minimization*. Finally, to show that the most effective allocation is a corner solution, observe that (σ_A^*, σ_W^*) is the unique solution to conditions (4) and (1). Since (σ_A^*, σ_W^*) achieves a minimum, the allocation that maximizes effectiveness must be a corner solution.[16]

The reason why it may be effective to move away from a symmetric allocation is as follows. Suppose for simplicity the two groups have equal size. At the symmetric allocation $\sigma_A^* = \sigma_W^* \equiv \sigma^*$, the size of the population who is exactly indifferent between their legal income and the expected return to crime is $f(q(\sigma^*))$ in both groups. Suppose that a tiny bit of interdiction is shifted from group W to group A. In group A, a fraction of size approximately $\varepsilon \times f(q(\sigma^*) - \varepsilon)$ will switch from crime to legal activity, while in group W a fraction of size approximately $\varepsilon \times f(q(\sigma^*) + \varepsilon)$ will switch from legal activity to crime. If f is decreasing (i.e., F is concave), fewer people take up crime than switch away from it, and so total crime is reduced relative to the symmetric allocation.

The fact that the most effective allocation can be completely lopsided is a result of our assumptions on the technology of search and policing. In a more realistic model, it is unlikely that the most effective allocation would entail complete abstention from policing one group of citizens. We discuss this issue in Section VIII, subsection F.

Proposition 1 highlights the fact that fairness and effectiveness may be antithetical, and that this is true for reasons that have nothing to do with differences across groups. Our main focus, however, is on the effectiveness 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.

IV. Small Adjustments Toward Fairness and Effectiveness of Interdiction

In this section we present conditions under which a small change from the equilibrium toward fairness (i.e., toward equalizing the search intensities across groups) decreases effectiveness relative to the equilibrium level. To that end, we first introduce the notion of quantile–quantile (QQ) plot. The QQ plot is a construction of statistics that finds applications in a number of fields: descriptive statistics, stochastic majorization and the theory of income inequality, statistical inference, and hypothesis testing.[17]

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

The QQ plot is an increasing function. Given any income level x with a given percentile in the income distribution of group W, the number $h(x)$ indicates the income level with the same percentile in group A. If, for

[16] An alternative method of proof would be to employ inequalities coming from the concavity of F. This method of proof would not require F to be differentiable. I am grateful to an anonymous referee for pointing this out.

[17] See, respectively, M. B. Wilk and R. Gnanadesikan (1968), Barry C. Arnold (1987, p. 47, ff.), and Eric L. Lehmann (1988). A related notion, the Ratio-at-Quantiles plot, was employed by Albert Wohlstetter and Sinclair Coleman (1970) to describe income disparities between whites and nonwhites.

example, 10 percent of members of group W have income smaller than x, then 10 percent of members of group A have income less than $h(x)$. In other words, the function $h(x)$ matches quantile x in group W to the same quantile $h(x)$ in group A.

We are interested in quantiles of the income distribution because citizens are assumed to compare the expected reward of engaging in crime to their legal income. If the quantiles of the income distribution are closely bunched together, a large fraction of the citizens has access to legal incomes that are close to each other and, in particular, close to the income of those citizens who are exactly indifferent between committing crimes or not. Thus, a large fraction of citizens are close to indifferent between engaging in criminal activity or not, and a small decrease in the expected reward to crime will cause a large fraction of citizens to switch from criminal activity to the honest wage. If, instead, the quantiles of the legal income distribution are spaced far apart, only a small fraction of the citizens are close to indifferent between engaging in criminal activity or not, and so slightly changing the expected rewards to crime only changes the behavior of a small fraction of the population.

Since we consider the experiment of redirecting interdiction power away from one group toward the other group, what is relevant for us is the *relative* spread of the quantiles. If, for example, the quantiles of the legal income distribution in group W are more spread apart than in group A, shifting interdiction power away from group A will cause a large fraction of honest group-A citizens to switch to illegal activity, while only few group-W criminals will switch to legal activities. The relative spread of quantiles is related to the slope of h. Suppose for example that x and y represent the 10 and 11 percent quantiles in the W group. If $h'(x)$ is smaller than 1, then x and y are farther apart than the 10- and 11-percent quantiles in the A group. In this case, we should expect group W to be less responsive to policing than group A. This is the idea behind the next lemma.

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 $\sigma_A^* = \bar{S}/N_A$ and $\sigma_W^* = 0$ because then we would have

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

where the inequality follows from Assumption 1. This inequality cannot hold in equilibrium since then a police officer's best response would be to search only group W, and this is inconsistent with $\sigma_W^* = 0$. A similar argument shows that it is not an equilibrium to have $\sigma_W^* = \bar{S}/N_W$ and $\sigma_A^* = 0$, since by stochastic dominance,

$$F_A(H) \geq F_W(H) > F_W\left(\frac{\bar{S}}{N_W}(J-H)+H\right).$$

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 total amount of crime, 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 yields

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

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

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

This equality must hold in equilibrium. Now, evaluate expression (3) at the equilibrium and use (6) to rewrite it as

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

This formula represents the derivative at the equilibrium point of the total amount of crime with respect to σ_A. If $h'(q(\sigma_W^*)) < 1$ this expression is negative (remember that $J - H$ 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 toward fairness entails decreasing σ_A^*.

To this end, note that, since F_W first-order stochastically dominates F_A, then $h(x) \leq x$ for all 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, the A group is searched more intensely than the W group in equilibrium, and so moving from the equilibrium toward fairness requires decreasing σ_A^*.

Expression (7) reflects the intuition given before Lemma 1; 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.

Definition 4: F_W is a stretch of F_A if $h'(x) \leq 1$ for all x.

The notion of "stretch" is novel in this paper. If F_W is a stretch of F_A, quantiles that are far apart in X_W will be close to each other in X_A. Intuitively, the random variable X_A is a "compression" of X_W because it is the image of X_W through a function that concentrates, or shrinks, intervals in the domain. Conversely, it is appropriate to think of X_W as a stretched version of X_A. 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.[18]

The QQ plot can be used to express relations of stochastic dominance. It is immediate to check that F_W first-order stochastically dominates F_A if and only if $h(x) \leq x$ for all x. But the property that, for all x, $h(x) \leq x$, is implied by the property that, for all x, $h'(x) \leq 1$.[19] In other words, if F_W is a stretch of F_A then a fortiori F_W first-order stochastically dominates F_A (the converse is not true). This observation allows us to translate Lemma 1 into the following proposition.

PROPOSITION 2: *Suppose that F_W is a stretch of F_A. Then a marginal change from the equilibrium toward fairness decreases effectiveness.*

V. Implementing Complete Fairness

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

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

(8)
$$h'(x) \leq \frac{\min_{z \in [h(x), x]} f_A(z)}{f_A(h(x))} \qquad \text{for all } x \text{ in } [\underline{q}, \bar{q}].$$

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

PROOF:

The proof of this proposition is along the lines of Lemma 1, but is more involved and not very insightful and is therefore relegated to the Appendix.

Since $\min_{z \in [h(x), x]} f_A(z) \leq f_A(h(x))$, condition (8) is more stringent than simply requiring that $h'(x) \leq 1$. It is sometimes easy to check whether condition (8) holds.[20] However, condi-

[18] 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 random variables, the QQ plot has a linear form $h(x) = a + bx$. In this case $h' < 1$ if and only if $b < 1$ if and only if $\mathrm{Var}(X_A) = b^2\mathrm{Var}(X_W) < \mathrm{Var}(X_W)$. The notion of stretch is related to some conditions that arise in the study of income mobility (Roland Benabou and Efe A. Ok, 2000) and of income inequality and taxation (Ulf Jakobsson, 1976).

[19] The implication makes use of the assumption that the infima of the supports of F_A and F_W coincide.

[20] 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 1 satisfies the conditions in Proposition 3 for all x in $[0, K]$. Indeed, in this case

tion (8) can only hold on a limited interval $[\underline{q}, \bar{q}]$. This interval cannot encompass the entire support of the two distributions, and in general there will exist values of $\bar{S}$ such that either $q(\sigma_A^*)$ or $q(\sigma_W^*)$ lie outside the interval $[\underline{q}, \bar{q}]$. More importantly, there generally exist values of $\bar{S}$ for which the conclusion of Proposition 3 is overturned and there is no trade-off between complete fairness and effectiveness. This is shown in the next proposition.

PROPOSITION 4: *Assume $F_A(H) = 1$ or $F_W(H) = 1$ or both. Assume further that the suprema of the supports of F_A and F_W do not coincide. Then there are values of $\bar{S}$ such that the completely fair outcome is more effective than the equilibrium outcome.*

PROOF:

Denote with $\hat{q}$ the supremum of the support of F_A. We can restate the assumption on the suprema of the supports as $F_W(\hat{q}) < 1 = F_A(\hat{q})$ (this is without loss of generality, as we can switch the labels A and W to ensure that this property holds).

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

$$(9) \quad \text{TV} = N_A[F_A(q(\bar{\sigma})) - F_A(q(\sigma_A^*))] + N_W[F_W(q(\bar{\sigma})) - F_W(q(\sigma_W^*))].$$

Case $F_W(H) < 1$: In this case our assumptions guarantee that $F_A(H) = 1$. Thus, for $\bar{S}$ sufficiently small we get a corner equilibrium in which $\sigma_W^* = 0$ and $\sigma_A^* = \bar{S}/N_A$. Therefore, $\sigma_W^* < \bar{\sigma} < \sigma_A^*$. Further, when $\bar{S}$ is sufficiently small $q(\sigma_A^*) = \sigma_A^*(J - H) + H$ approaches H. In turn, H is larger than $\hat{q}$ since $F_A(H) = 1$. Therefore, for small $\bar{S}$ we must have $q(\sigma_A^*) > \hat{q}$. In sum, for $\bar{S}$ sufficiently small we have

$$\hat{q} < q(\sigma_A^*) < q(\bar{\sigma}) < q(\sigma_W^*) = H.$$

These inequalities imply that $F_W(q(\bar{\sigma})) < F_W(H)$ and also, since $F_A(\hat{q}) = 1$, that $F_A(q(\sigma_A^*)) = F_A(q(\bar{\sigma})) = 1$. Consequently $\text{TV} < 0$.

Case $F_W(H) = 1$: In this case, pick the maximum value of $\bar{S}$ such that in equilibrium all citizens engage in crime. Formally, this is the value of $\bar{S}$ 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 $\hat{q}$ we are guaranteed that $\hat{q} = q(\sigma_A^*) < q(\bar{\sigma}) < q(\sigma_W^*)$. Expression (9) becomes

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

It is important to notice that Proposition 4 holds for most pairs of c.d.f.'s, including those for which $h' < 1$. Comparing with Proposition 2, one could infer that it is easier to predict the impact on effectiveness of small changes in fairness rather than the effects of a shift to complete fairness.[21] By showing that the conclusions of Proposition 3 can be overturned for a very large set of c.d.f.'s (by choosing $\bar{S}$ appropriately), Proposition 4 highlights the importance of knowing whether the equilibrium values $q(\sigma_r^*)$ lie inside the interval $[\underline{q}, \bar{q}]$; this is additional information, beyond the knowledge of F_A and F_W, which was not necessary for Proposition 2.

VI. Recovering the Distribution of Legal Earning Opportunities

The QQ plot is constructed from the distributions of *potential* legal earning opportunities. Potential legal earnings, however, are sometimes unobservable. This is because, according to our model, citizens with potential legal income below a threshold choose to engage in crime. These citizens do not take advantage of their legal earning opportunities. Some of these citizens will be caught and put in jail, and are

$f_A(y)/f_A(h(x)) = 1$ for all y in $[h(x), x]$ so long as $x \in [0, K]$.

[21] Although in the proof of Proposition 4 we have chosen the value of $\bar{S}$ 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 0 and 1.

thus missing from our data, which raises the issue of truncation. The remaining fraction of those 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 in our data.

In this section we present a simple extension of the model that is more realistic and deals with these selection problems. Under 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.

Consider a world in which a fraction α of citizens of both groups 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 *strategic* (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 strategic. 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 group r who engage in crime is $(1 - \alpha)F_r(q(\sigma_r))$. The analysis of Section II goes through almost unchanged in the extended model, and it is easy to see that:

(i) For any $\alpha > 0$ the equilibrium and effectiveness conditions coincide with conditions (2) and (4).
(ii) Consequently, the equilibrium and most effective allocations are independent of α.
(iii) The analysis of Sections IV and V goes through verbatim.

This means that also in this augmented model, our predictions concerning the relationship between effectiveness and fairness depend on the shape of the QQ plot of (unobservable) legal earning opportunities, exactly as described in Propositions 2–4. Therefore, in terms of equilibrium analysis the augmented model behaves exactly like the case $\alpha = 0$.

Furthermore, the augmented model is more realistic, since it allows for a category of people who do not engage in crime even when their legal earning opportunities are low. This extended model allows realistically to pose the question of recovering the distribution of legal earning opportunities. To understand the key issue, observe that in the extreme case of $\alpha = 0$ which is discussed in the preceding sections, all citizens with legal earning opportunities below a threshold become criminals and do not report their potential legal income. If we assume that they report zero, the resulting c.d.f. of reported income has a flat spot starting at zero. This stark feature is unrealistic and is unlikely to be verified in the data.

When $\alpha > 0$, the natural way to proceed would be to recover the c.d.f.'s of legal earning opportunities from the c.d.f.'s of reported income. However, this exercise may not be easy if we do not know α and the equilibrium values $q(\sigma_r^*)$. The next proposition offers a procedure to bypass this obstacle, by directly computing the QQ plot based on reported incomes. The resulting plot is guaranteed to coincide, under the assumptions of our model, with the one based on 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 $\hat{F}_r(x)$, the distribution of reported earnings. All citizens of group r with potential legal earnings of $x > q(\sigma_r^*)$ do not engage in crime regardless of whether they are decent or strategic. They realize their potential legal earnings, and report them truthfully. This means that $\hat{F}_r(x) = F_r(x)$ when $x > q(\sigma_r^*)$.

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

$$\hat{F}_r(x) = \alpha F_r(x) + (1 - \alpha)F_r(q(\sigma_r^*)).$$

The first term represents the decent citizens who report their legal income, the second term those strategic citizens who engage in crime and, by assumption, report zero income. The function $\hat{F}_r(x)$ 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

$$\hat{F}_r(x) = \begin{cases} \alpha F_r(x) + \beta & \text{for } x < q(\sigma_r^*) \\ F_r(x) & \text{for } x > q(\sigma_r^*). \end{cases}$$

The inverse of $\hat{F}_r$ reads

$$(10) \quad \hat{F}_r^{-1}(p) = \begin{cases} F_r^{-1}\left(\dfrac{p - \beta}{\alpha}\right) & \text{for } p < F_r(q(\sigma_r^*)) \\ F_r^{-1}(p) & \text{for } p > F_r(q(\sigma_r^*)). \end{cases}$$

Denote the QQ plot of the reported income distributions by

$$\hat{h}(x) \equiv \hat{F}_A^{-1}(\hat{F}_W(x)).$$

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

$$\hat{h}(x) = \hat{F}_A^{-1}(\hat{F}_W(x)) = \hat{F}_A^{-1}(F_W(x)) = F_A^{-1}(F_W(x)) = h(x)$$

where the second 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 $\hat{F}_W(x) = \alpha F_W(x) + \beta$, and so

$$\hat{h}(x) = \hat{F}_A^{-1}(\alpha F_W(x) + \beta) = F_A^{-1}\left(\frac{(\alpha F_W(x) + \beta) - \beta}{\alpha}\right) = F_A^{-1}(F_W(x)) = h(x)$$

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 $\hat{h}(x) = h(x)$ for all x.

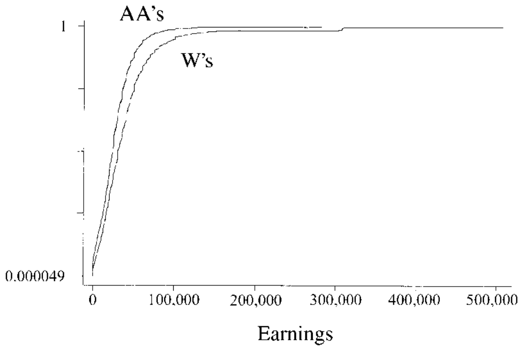

FIGURE 1. CUMULATIVE DISTRIBUTION FUNCTIONS OF EARNINGS OF AFRICAN-AMERICANS (AA'S) AND WHITES (W'S)

Proposition 5 suggests a method which, under the assumptions 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 already assume that "lucky" criminals (who are in our sample) report zero income, this modification achieves a situation in which all citizens who engage in crime 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. Notice that, to implement this procedure, it is not necessary to know α or σ_r^*.

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 earning distributions of males of age 15–55 who reside in metropolitan statistical areas of the United States.[22] Figure 1 presents the c.d.f.'s of reported earnings (earnings are on the horizontal axis).

[22] We exclude the disabled and the military personnel. Earnings are given by the variable PEARNVAL and include total wage and salary, self-employment earnings, and any farm self-employment earnings.

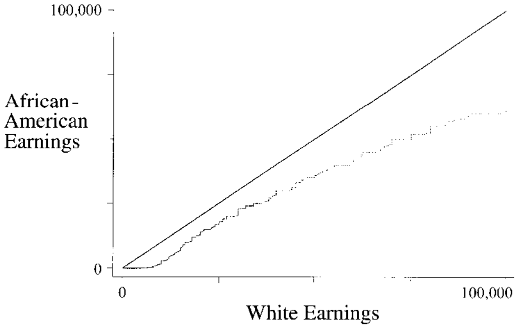

FIGURE 2. QQ PLOT BASED ON EARNINGS OF AFRICAN-AMERICANS AND WHITES

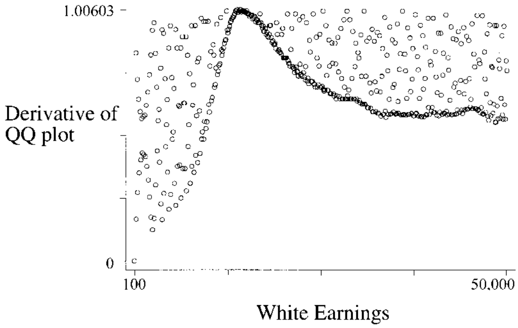

FIGURE 3. NUMERICAL DERIVATIVE OF THE QQ PLOT

It is clear that the earnings of whites stochastically dominates those of African-Americans. 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.[23] Figure 2 presents the QQ plot between earning levels of zero and \$100,000: earnings in the white population are on the horizontal axis.[24]

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 numerical derivative of the QQ plot that is stable, we first smooth the probability density functions (p.d.f.'s) and then use the smoothed p.d.f.'s to construct a smoothed QQ plot.[25] Figure 3 presents the numerical derivative of the smoothed QQ plot.

The picture suggests that the derivative of the QQ plot tends to be smaller than 1 at earnings smaller than \$100,000 for whites. The derivative approaches 1 around earnings of \$12,000 for whites. In the context of our stylized model, this finding implies that if one were to move toward fairness by constraining police to shift

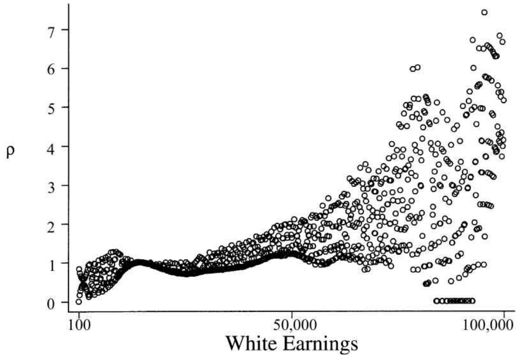

FIGURE 4. PLOT OF ρ

resources from the A group to the W group, then the total amount of crime would not fall. We emphasize that we should not interpret this as a policy statement, since the model is too general to be applied to any specific environment. We view this back-of-the-envelope computation as an illustration of our results.

It is interesting to check whether condition (8) in Proposition 3 is satisfied. That condition is satisfied whenever

$$\rho(x) \equiv h'(x) f_A(h(x)) / \min_{z \in [h(x), x]} f_A(z) < 1.$$

Figure 4 plots this ratio (vertical axis) as a function of white earnings (horizontal axis).

The ratio does not appear to be clearly below 1 except at very low values of white earnings, and clearly exceeds 1 for earnings greater than \$50,000. Thus, condition (8) does not seem to

[23] Of 100,000 African-American (white) citizens, 6,838 (990) are incarcerated according to data from the Department of Justice.

[24] In the interest of pictorial clarity we do not report the QQ plot for those whites with legal earnings above \$100,000. These individuals are few and the forces captured by our model are unlikely to accurately describe their incentives to engage in crime.

[25] The smoothed p.d.f.'s are computed using a quartic kernel density estimator with a bandwidth of \$9,000.

be useful in this instance to sign the effect of a switch to complete fairness.

VII. 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 citizens of all groups should bear this cost equally (in expectation).[26] It may seem harmless to include in this computation those costs which relate to the loss of reputation 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.

Being publicly singled out for search entails a cost because external observers (bystanders, neighbors, etc.) use the search event to infer something about the criminal propensity of the individual who is subject to search. For instance, if police have knowledge of the criminal record of an individual at the time of the search and an external observer does not, the observer should use the search event to update his opinion about the criminal record of the individual who is searched.[27] Christopher 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 as a key interest of citizens in avoiding search and seizure.

An interesting feature of disrepute as we have defined it is that it is endogenous, in the sense that it is the result of Bayesian updating and is not a primitive of the model. The amount of disrepute that is attached to being searched will generally depend on the search strategy that police use. The fact that the loss of reputation is endogenous opens up the possibility 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 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. Since we wish to capture a notion of updating on the part of bystanders, in addition to the observable characteristic "group," 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. For concreteness we refer to this characteristic as the criminal record, and we denote by $\Pr(c|r)$ its distribution in each group. Since the police condition their search behavior on c, upon seeing an individual being searched, the bystanders would update their prior about the c of the individual being searched, and hence about his criminal propensity. Denote with $\sigma(c, r)$ the probability of being searched for a random individual with criminal record c and group r.

Define the *disrepute* δ of a member of group r as

$$\delta(r) \equiv \Pr(r \text{ is criminal}) - \Pr(r \text{ is criminal}|\text{searched})$$

where the event "r is criminal|searched" denotes the event that a randomly chosen individual of group r is criminal conditional on being searched.[28] Thus, we define the cost of disrepute as the amount of updating that an observer does about an individual of group r upon seeing that he is being searched. This definition captures the humiliation connected with being publicly searched.

Consistent with the notion of updating, we must also take into account the fact that an

[26] These costs encompass "objective" and "subjective" intrusions, as defined in *Michigan Department of Police* v. *Sitz,* 496 U.S. 444 (1990).

[27] We refer to the criminal record for concreteness. In highway interdiction, for example, police often have knowledge of the criminal record because they radio in the driver's license of the individual who is stopped before deciding whether to initiate a search. Of course, our analysis is equally relevant if police are more informed than the external observer about *any* characteristic of the individual who may be searched.

[28] 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 group r has a value of $c \in \mathcal{E}$," where $\mathcal{E}$ is any set.

observer updates about a random individual also when that individual is *not* searched. Individuals who are not searched experience a corresponding increase in status in the eye of an external observer. We term this effect *negative disrepute,* and we denote it by $\bar{\delta}$. Formally,

$$\bar{\delta}(r) \equiv \Pr(r \text{ is criminal}) - \Pr(r \textit{ is criminal} | \text{not searched}).$$

The total disrepute effect (positive and negative) in group r is given by

$$(11) \quad \delta(r) \times \sigma(r) + \bar{\delta}(r) \times [1 - \sigma(r)] = 0$$

where $\sigma(r) = \sum_c \sigma(r, c)\Pr(c|r)$ denotes the probability that a randomly chosen individual of group r is searched. Equality to zero follows from the fact that conditional probabilities are martingales. Notice that equation (11) holds true for any value of $\sigma(r)$; regardless of how the search intensity is distributed between groups, the disrepute effect has constant sum (zero) within a group.[29] This means that, in terms of disrepute, the search strategy is *neutral within group r*: changing the search strategy cannot change the average disrepute within a group.

This neutrality result says that redistributing search intensity across groups has no effect on the disrepute portion of the average (within a group) cost of being searched. Thus, when we evaluate the effects of a certain search strategy, we can ignore the distributive effects of disrepute across groups. This finding is in contrast with the conventional view that "reputational" costs are an important element of disparity in policing. This view relies on a logical failure to take into account the complete effects of Bayesian updating (i.e., the effect on those who are *not* searched). After taking into account the full effects of Bayesian updating, a redistributive notion of fairness seems harder to justify purely on the basis of reputational costs.

As mentioned in the introduction, this line of inquiry does not lead to questioning the appropriateness of refocusing search intensity from African-Americans to whites. Disrepute is but one of the components of the cost of being searched. To the extent that being searched involves important nonreputational costs (such as loss of time, risk of being abused, etc.), it is perfectly reasonable to wish to equate these costs across races. The point of our line of inquiry is to suggest that, if nonreputational costs are what causes the unfairness, there is probably room for improvement through remedial policies such as sensitivity training for police. These policies can reduce the cost of being searched, but only if the cost is nonreputational.

[29] This phenomenon is purely a consequence of Bayesian updating and does not hinge on any assumption about the behavior of either police or citizens.

VIII. Discussion of Modeling Choices

A. *The Objective Function of Police*

An important assumption in our model is that police choose whom to search so as to maximize successful searches. In the equilibrium analysis, this assumption is reflected in the equalization across groups of the success rates of searches. This equalization is observed in a number of instances of interdiction, including vehicular searches, "stop and frisk" neighborhood patrolling, and airport searches, as discussed in Section II, subsection A. We view this evidence as supportive of our assumptions about the motivation of police, since equality of the crime rates across different categories of citizens is not likely to arise under other assumptions about the motivation of police.

Additional support for the assumption that police maximize successful searches is provided, for the case of highway interdiction, by the explicit reference to this point in Verniero and Zoubek (1999), who describe the trooper's incentive system as follows:

> [E]vidence has surfaced that minority troopers may also have been caught up in a system that rewards officers based on the quantity of drugs that they have discovered during routine traffic stops.... The typical trooper is an intelligent, rational, ambitious, and career-oriented professional who responds to the prospect of rewards and promotions as much as to the threat of discipline and punishment. The system of organizational rewards, by definition and design, exerts a powerful

influence on officer performance and enforcement priorities. The State Police therefore need to carefully examine their system for awarding promotions and favored duty assignments, and we expect that this will be one of the significant issues to be addressed in detail in future reports of the Review Team. It is nonetheless important for us to note in this Report that the perception persisted throughout the ranks of the State Police members assigned to the Turnpike that one of the best ways to gain distinction is to be aggressive in interdicting drugs. This point is best illustrated by the Trooper of the Year Award. It was widely believed that this singular honor was reserved for the trooper who made the most drug arrests and the largest drug seizures. This award sent a clear and strong message to the rank and file, reinforcing the notion that more common rewards and promotions would be provided to troopers who proved to be particularly adept at ferreting out illicit drugs.[30]

B. *Alternative Objectives of Police*

A system of rewards based on successful searches helps motivate police to exert effort in a world where an individual police officer's effort is too small to measurably affect the aggregate crime level. On this basis, we should believe that police incentives should be based at least partly on success maximization, and that these incentives will partly affect the racial distribution of searches. At the same time, other factors may also play a role in the allocation of police manpower across racial groups; in city policing, for example, crime in certain wealthy (and disproportionately white) neighborhoods may have great political salience, so those neighborhoods will be allocated more interdiction power and will experience lower crime rates. It is therefore interesting to consider an extension of the model which reflects these considerations.

Consider a situation of city policing in which police incentives are shaped by political pressure to stamp out crime committed in nonminority neighborhoods. Given any allocation of police manpower across neighborhoods, the base model of Section I would apply within each neighborhood. If, however, neighborhoods are segregated, making the allocation more fair may require shifting manpower across neighborhoods. This is a case that is not contemplated in our base model. The model can nevertheless be adapted to approximate this situation if we imagine two completely segregated neighborhoods, the A and the W. Citizens (criminals or not) cannot escape their neighborhood. Assume that police are rewarded more highly for catching a criminal belonging to neighborhood W. Under this assumption, the equilibrium crime rate (and the success rate of searches) will no longer be constant across the two groups. Instead, the equilibrium crime rate will be higher in neighborhood A. At the same time, as long as the rewards for catching a neighborhood-W criminal are not too much larger than those for catching a neighborhood-A criminal, the W neighborhood will be searched with less intensity than the A neighborhood. There is some realistic appeal in this combination of higher search intensity and higher crime rate (and success rate of searches) in the A neighborhood. Starting from this premise, we ask what happens to the crime rate if police are required to behave more fairly.

To answer this question, denote with σ_r^{**} the equilibrium search intensities in the augmented model in which police receive a higher reward for busting a neighborhood-W criminal. Since at an interior equilibrium police must be indifferent between searching a member of the two neighborhoods, the expected probability of success must be lower in neighborhood W, that is $F_A(q(\sigma_A^{**})) > F_W(q(\sigma_W^{**}))$. This inequality can be rewritten as

$$q(\sigma_A^{**}) > h(q(\sigma_W^{**})).$$

In addition, we have posited that the equilibrium search intensity in neighborhood W does not exceed that in neighborhood A, or equivalently that

$$q(\sigma_A^{**}) \leq q(\sigma_W^{**}).$$

Equipped with these inequalities, let us write down the total crime rate. Assume for simplicity that the two neighborhoods are of the same size

[30] Cited from Verniero and Zoubek (1999, pp. 42–43).

(all the steps go through almost unchanged, and the conclusion is the same, if we remove this simplifying assumption). The crime rate is then

$$F_A(q(\sigma_A^{**})) + F_W(q(\sigma_W^{**})).$$

Transfering one unit of interdiction from the A to the W neighborhood increases total crime if and only if

$$f_W(q(\sigma_W^{**})) \le f_A(q(\sigma_A^{**})).$$

Since $h(q(\sigma_W^{**})) < q(\sigma_A^{**}) \le q(\sigma_W^{**})$, this condition will hold if

$$f_W(q(\sigma_W^{**})) \le \min_{z \in [h(q(\sigma_W^{**})), q(\sigma_W^{**})]} f_A(z).$$

Rewrite this inequality substituting x for $q(\sigma_W^{**})$, and divide through by $f_A(h(x))$ to obtain

$$h'(x) \le \frac{\min_{z \in [h(x) x]} f_A(z)}{f_A(h(x))}.$$

This condition coincides with condition (8) from Proposition 3. We have therefore shown that in the augmented model in which police rewards are greater for busting neighborhood-W criminals, condition (8) is a sufficient condition for there to exist a trade-off between fairness and effectiveness.

C. *Nonracially Biased Police*

In this paper we assume that police are not racially biased (i.e., that they do not derive greater utility from searching members of group A rather than members of group W). We do this for two reasons: First, 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, at least in some practical instances of alleged disparate impact, the unfairness is ascribed to the incentive system, and not directly to racial bias on the part of police; this is the position in Verniero and Zoubek (1999; see their part III-D) and is also consistent with the findings in Knowles et al. (2001). 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 affecting the equilibrium, but in terms of the focus of this paper, the presence of racial bias on the part of the police need not necessarily make it more effective to implement fairness—although it would probably make fairness more desirable on moral grounds.

In this connection, it is worth emphasizing that while our model focuses on the economic incentives to commit crime, it ignores the possibility that unfairness of policing per se encourages crime among those who are unfairly treated. According to Robert Agnew's (1992) general strain theory, the anticipated presentation of negative stimuli results in strain. Strain, in turn, causes individuals to resort to illegitimate ways to achieve their goals. If we take this viewpoint, the expectation of being unfairly treated by police can produce strain. There is evidence that some groups anticipate being frustrated in the interaction with police (see Anderson, 1990). According to this view, interdiction power can be not a deterrent but a cause of crime if it is unfairly applied.

D. *Differential Deterrence*

We have assumed that the deterrence effect is identical for both groups. This is a reasonable assumption as long as crime's gain and punishment are similar across groups. It is possible, however, that convicted criminals from a given group are likely to incur more severe punishment (longer prison sentences), leading to a smaller value of J for that group. On the other hand, the social stigma from being convicted, presumably also a component of J, could be weaker in a given group, leading to larger values of J. Along the same lines, it could be argued that H is larger for one group when being a (successful) criminal does not carry much social stigma in that group. To allow for the possible difference in the deterrence effect of policing, we now explore what happens to the key result in this paper, Lemma 1, if we assume that J and H are group-specific.[31]

[31] To some extent, the effects mentioned above may countervail themselves. Consider, for example, a thought experiment in which the social stigma attached to being a criminal decreases. In our formalization, this decrease will result in higher values of both J and H, but the difference

Consider an augmented model in which J_r and H_r denote the group-specific cost and benefit of crime, but which is otherwise the same as the one of Section I. Denote by σ_r^{***} the equilibrium search intensities in this augmented model. The police maximization problem will still require that, in equilibrium, the fraction of criminals is the same in the two groups. Thus, the equilibrium search intensities must solve

$$F_A(q_A(\sigma_A^{***})) = F_W(q_W(\sigma_W^{***}))$$

where $q_r(\sigma) = \sigma(J_r - H_r) + H_r$. One can then replicate the analysis of Lemma 1 and find that its conclusion holds if and only if

$$h'(q_W(\sigma_W^{***})) < \frac{|J_A - H_A|}{|J_W - H_W|}. \tag{12}$$

In other words, marginally shifting the focus of interdiction towards the W group increases the total amount of crime if and only if this inequality holds. Condition (12) differs from the condition in Lemma 1 in the right-hand side of the inequality, which is no longer necessarily equal to 1. The quantity $|J_r - H_r|$ represents the drop in utility that a criminal of group r experiences if caught; it captures the deterrence effect of policing on group r. If $|J_A - H_A| = |J_W - H_W|$, the deterrence effect of policing is the same in both groups and condition (12) reduces to the one in Lemma 1. If, however, $|J_A - H_A| < |J_W - H_W|$, the prospect of being caught deters citizens of group W more than citizens of group A. In this case condition (12) is stronger (more restrictive) than the condition in Lemma 1, reflecting the fact that shifting the focus of interdiction towards the W group is less likely to result in increased crime.

Condition (12) shows how our key result is changed with differential deterrence. The fact that there is a change highlights the importance of the maintained assumption of equal deterrence. At the same time, from the methodological viewpoint we find that the basic analysis straightforwardly generalizes to this richer environment, and equation (12) confirms the central role played by the QQ plot even in the case of differential deterrence.

$J - H$ may not be affected. This difference is what matters for the analysis, as shown in what follows.

E. *Changing Characteristics*

We have assumed that the composition of the groups is fixed and cannot be changed. In this subsection we consider the case in which criminals can change their characteristics to reduce the risk of being caught. In a model with more than two groups, we could think of many ways in which criminals of a highly targeted group might attempt to disguise some suspicious characteristic (by driving certain types of cars, by dressing differently, etc.) to blend in with members of a different group. Even in our simple model with just two groups, we could get the same effect; if one racial group is searched more frequently by highway patrols, drug traffickers who belong to that group may decide to employ couriers, or "mules," of the other racial group in order to reduce the probability of their cargo being discovered.[32] Knowles et al. (2001) point out that the key equilibrium prediction of the base model, namely, equality of success rates of search, is preserved in the more general model in which characteristics can be changed.

To see how the argument works in our two-group model, imagine that members of group A can, at cost c, change their appearance to that of a group-W member. Assume further that F_A stochastically dominates F_W, so that in the model with fixed characteristics we have $\sigma_A^* > \sigma_W^*$. Paying c allows a group-A member to enjoy the lower search intensity σ_W^* instead of σ_A^*. If c is not too high, all those group-A citizens who engage in crime find it profitable to pay c and change their appearance. In this case, it is no longer an equilibrium for police to search those who look like they belong to group A, since there are no criminals among those citizens. This means that the search intensity must be different in the equilibrium of the model with changeable characteristics. Yet, if (σ_A^*, σ_W^*) was an interior equilibrium when characteristics are fixed, then a fortiori the equilibrium will be interior if citizens can change their appearance.[33] That is, "interiority" of equilibrium is maintained

[32] We assume that the drug traffickers, as well as the mules, will go to jail if their cargo is found by police.

[33] Equivalently, and perhaps easier to see: if an allocation $(\tilde{\sigma}_A, \tilde{\sigma}_W)$ is a noninterior equilibrium in the model in which citizens can change their appearance, then a fortiori it is an equilibrium if citizens cannot change their appearance.

if we allow citizens to change characteristics. Interiority implies that police must obtain the same success rate in searching either group. Thus, that key implication of our model, equality of success rates of searches among groups (now, among those citizens who *appear to belong* to the different groups) holds even if we allow citizens to choose their characteristics.[34]

The elasticity of crime to policing will depend on the opportunity for groups to disguise themselves. Nevertheless, some of our results still apply concerning the trade-off between effectiveness and fairness. To see why, let us evaluate the effect of a small shift in policing intensity starting from the equilibrium of the model with changeable characteristics. It is easiest to reason starting from an allocation that is slightly more fair than the equilibrium one; we will then use continuity of the crime rate in the search intensity to draw implications about the equilibrium allocation. At this slightly fairer allocation, no group-A criminal finds it expedient to change appearance. In this respect the situation is similar to the base model in which characteristics cannot be changed. However, the allocation is more fair than the equilibrium in that base model: group A must be policed less intensely, and group W more intensely, in order to make group-A citizens almost indifferent between switching groups. We are, then, in an environment that is identical to the one analyzed in subsection B. The same condition, condition (8) from Proposition 3, will be sufficient to identify a trade-off between effectiveness and fairness. Under this condition, a small increase in fairness comes at the cost of increased crime. Using the fact that the crime rate varies continuously with intensity of interdiction, we then extend this result to a nearby allocation, the equilibrium one. We conclude that if condition (8) is met, making the equilibrium slightly more fair will increase the crime rate even if citizens can choose to switch away from group A.

[34] One may be interested in the complete description of equilibrium in this case. Denote the equilibrium search intensities by $(\tilde{\sigma}_W, \tilde{\sigma}_A)$. Group-$A$ criminals are indifferent between switching group or not, so we must have $\tilde{\sigma}_A(J - H) = \tilde{\sigma}_W(J - H) - c$. Group-$A$ criminals will randomly switch group with a probability that is designed to equate the fraction of criminals in group A to the fraction of criminals among those citizens who look like group W. This choice of probability makes police willing to search the two groups with search intensities $(\tilde{\sigma}_A, \tilde{\sigma}_W)$.

F. *The Technology of Search*

An important simplifying assumption underlying our analysis is that search intensity can be shifted effortlessly across groups. Perfect substitutability is a strong assumption. In particular, achieving a very high search intensity of one group may be very costly in terms of police manpower. Imagine, for example, that troopers wanted to stop and search almost all male drivers. To achieve that high level of search intensity, troopers would have to continually monitor traffic 24 hours a day, and that may be very costly. The last unit of search intensity of males is probably more costly than (cannot be achieved by forgoing) just one unit of search intensity of females. To the extent that the cost of achieving a given search intensity differs across groups, the success rates of searches will depart from equalization across groups. Ultimately, the question of scale efficiencies in policing is an empirical one. The result in Knowles et al. (2001) suggest that, in highway interdiction, scale effects are not sufficiently important to undermine equality of success rates. In other situations, however, scale efficiencies may well play an important role.

G. *Achieving the Most Effective Allocation*

It is important to recognize that this paper takes a *positive* approach—in the sense that it takes as given a realistic incentive structure and investigates the comparative statics properties of the model and whether there is a conflict between various desirable properties that may be required of allocations. Because police are assumed not to care about deterrence, but only about successful searches, in the equilibrium of the model crime is not minimized.

The approach in this paper is not *normative,* in the sense that we do not pursue the question of how to achieve particular allocations. For example, the fact that the most effective (crime-minimizing) outcome is generally not achieved in equilibrium does not mean that in this setup the most effective outcome cannot be implemented. In our simple model, the most effective allocation *could* be implemented—at the cost of giving race-dependent incentives to police. By giving police higher rewards for successful busts on citizens of a particular race, it is possible to induce *any* allocation of search intensity on the part of police. In particular, it will be possible to implement the most effective allocation and the completely fair allocation.[35] This argument shows that our model is simplified in such a way that asking the implementation question is not interesting—although the model is well suited to asking other questions. A model that were to attempt an interesting analysis of policing from the viewpoint of implementation would probably have to provide foundations for the costs and benefits of interdiction, as well as of legal and illegal activities. Such an analysis is not the goal of the present work.

Nevertheless, it is interesting to inquire about the reason for why police might be given incentives that are out of line with crime reduction. In addition to the fact that minimizing crime might require incentive schemes that are highly unfair and therefore politically risky, we can offer another possible reason: the crime rate is sometimes not a direct concern of the police force who does the policing. For example, only a relatively small fraction of the illegal drugs transported on I-95 originates from, or is directed to, Maryland. Most of it originates from Florida and is directed to the large metropolitan areas of the Northeast. One might therefore expect that the Maryland police force policing I-95 would only have a partial stake in the crime rate generated by the I-95 drug traffic. In such cases, it is not hard to believe that a police department would maximize the number of drug finds and place little emphasis on crime minimization.

H. *Notions of Effectiveness of Interdiction*

We have taken the position that effectiveness of interdiction is measured by the total number of citizens who commit crimes. According to this measure, given a total number of crimes, effectiveness of interdiction is the same regardless of 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 committed by one racial group then, depending on the type of crime, members of that group may be targets of crime more often. This is a valid argument which suggests that one should also try to reduce the inequality in the distribution of crime across groups. This raises some interesting questions, such as how much inequality in crime rates should be tolerated across groups. Going to one extreme and keeping the crime rate completely homogeneous across groups necessitates policing different groups with different intensity (this is the outcome that obtains in the equilibrium of our model when police are unconstrained). This, critics 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.

IX. Conclusion

The controversy on racial profiling focuses on the fact that some racial groups are disproportionately the target of interdiction. This feature is not peculiar to highway interdiction: it arises in many other policing situations, such as neighborhood policing and customs searches. The resulting racial disparities are unfortunate but, as discussed in the introduction, are not per se illegal if they are an unavoidable by-product of effective policing. The fundamental question, therefore, is whether there is necessarily a tradeoff between fairness and effectiveness in policing. This question is coming to the fore once again in regards to the profiling of airline passengers. At the moment, airport security mostly relies on screening all passengers with metal detectors (which, incidentally, means that women and men are treated equally despite the fact that female terrorists seem to be rare.) To achieve more effective interdiction, the U.S. government is planning to establish a centralized data base of ticket reservations, to be mined for unusual or suspect patterns.[36] At the moment,

[35] 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 III, the efficient outcome may be very unfair. We implicitly subscribe to some of these considerations by focusing on fairness.

[36] See Robert O'Harrow, Jr. (2002).

public opinion seems willing to tolerate some amount of profiling in exchange for greater security.[37]

In this paper the trade-off between fairness and effectiveness in policing has been investigated from a theoretical viewpoint, by employing a rational choice model of policing and crime to study the effects of implementing fairness. Our notion of fairness is appealing on normative grounds and speaks to the concerns of critics of racial profiling. We have shown that the goals of fairness and effectiveness are not necessarily in contrast: sometimes, forcing the police to behave more fairly can increase the effectiveness of interdiction. We have given exact conditions under which, within our simple model, the contrast is present.

These conditions are based on the distributions of legal earning opportunities of citizens, and are expressed as constraints on the QQ plot of these distributions. Thus, our model relates the trade-off between fairness and effectiveness of policing to income disparities across races. In our model it is possible directly to recover the QQ plot of the distributions of legal earning opportunities by using reported earnings, so our approach has the potential to deliver quantitative implications. We view the close relationship between the model and data as a desirable feature.[38] At the same time, we are aware that we have ignored many factors that are important in the decision to engage in crime, an issue to which we will return.

Finally, we have suggested that a redistributive notion of racial fairness such as the one we adopt (the average citizen should bear the same expected cost of being searched regardless of his/her race) may not be meaningful when the cost of being searched reflects the stigmatization connected with being singled out for search. Therefore, we conclude that there may be theoretically valid reasons to ignore the cost of stigmatization in a discussion of racial fairness. This conclusion may be cause for optimism. The finding that the main costs of being searched is due to "physical" aspects of the search (such as loss of time, improper behavior of police, etc.) is encouraging since aggressive policy action can presumably ameliorate these problems. This is in contrast to the costs due to stigmatization, which are endogenous to the model and, therefore, potentially beyond the reach of policy instruments. To the extent that the physical costs of being searched can be reduced by remedies such as better training for police, remedial action can focus on police behavior during the search, rather than on constraining the search strategy of police.

One contribution of this paper is to give a rigorous foundation to the idea that fairness and effectiveness of policing are not necessarily in contrast, and to show that due to a "second best" argument, constraining police behavior may well increase effectiveness of interdiction. On the other hand, we have not shown that this is the case in practice. Although we do not claim conclusively to have answered the question, we hope that, if the framework proposed here proves to be convincing, it can provide an analytical foundation for the public debate.

This paper has dealt with a very simple model where crime is endogenous and the decision to engage in crime reflects a lack of legal earning opportunities. To highlight the basic forces in the simplest way, we chose to focus on few modeling elements. In most of the paper, for example, race is the only characteristic that is observable to police. Also, we have assumed that the rewards to crime are independent of one's legal earning opportunities, whereas in reality the two may be correlated. Analogous simplifying assumptions are common to much of the theoretical literature on discrimination, and we have exploited the simplifications to obtain theoretically clean results. At the same time, we have ignored many interesting and possibly controversial questions that would arise in a more complex model, for example, the question of the right definition of "fairer interdiction" in a world in which there are more than two characteristics. Due to the many simplifications and to the generality of the model, no part of the analysis should be understood to

[37] See Henry Weinstein et al. (2001).

[38] Some interesting recent papers in the discrimination literature (see especially Hanming Fang [2000] and Moro and Norman [2000]) have focused on estimating models of statistical discrimination à la Arrow (1973). In statistical discrimination models, individuals differ in their cost of undertaking a productive investment. Since this characteristic is unobservable, estimating its distribution in the population is a difficult endeavor.

have direct policy implications. Any realistic situation will be richer in detail than this stylized model, and a number of additional considerations will be relevant in each case. In each specific situation of policing, the additional modeling structure will likely enrich the insights of the simple model presented here. In order to obtain policy implications, the model needs to be tailored to specific situations of interdiction.

The model developed here could be applied to tax auditing. In this application, the two groups of citizens would represent groups of taxpayers that are observably different. These different groups of citizens will reasonably differ in their elasticity to auditing; small taxpayers, for example, more than large corporations who employ tax lawyers may dread the prospect of being audited. The role of police would now be played by the IRS, who is assumed to maximize expected recovery. There is a large literature on auditing models (see, e.g., Parkash Chander and Louis L. Wilde, 1998). Most of it assumes that there is only one observable group of taxpayers, with the exception of Susan Scotchmer (1987) who does not focus on the disparity in auditing rates. We hope that the questions asked in the present paper can stimulate further research into the issue of auditing with different groups, but we leave this topic for future research.

Appendix

PROOF OF PROPOSITION 3:

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

$$F_A(\bar{\sigma}(J-H)+H) - F_A(\sigma_A^*(J-H)+H)$$

$$= -\int_{\bar{\sigma}}^{\sigma_A^*} \frac{\partial F_A(s(J-H)+H)}{\partial s}\, ds$$

$$= |J-H| \int_{\bar{\sigma}}^{\sigma_A^*} f_A(q(s))\, ds.$$

Similarly,

$$F_W(\sigma_W^*(J-H)+H) - F_W(\bar{\sigma}(J-H)+H)$$

$$= |J-H| \int_{\sigma_W^*}^{\bar{\sigma}} f_W(q(s))\, ds$$

$$= |J-H| \int_{\sigma_W^*}^{\bar{\sigma}} [h'(q(s)) f_A(h(q(s)))]\, ds.$$

The total variation in crime [expression (9)] reads

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

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

(A1)

$$\mathrm{TV} \geq |J-H| \times \left[N_A(\sigma_A^* - \bar{\sigma})m - N_W \int_{\sigma_W^*}^{\bar{\sigma}} [h'(q(s)) f_A(h(q(s)))]\, ds \right]$$

$$= |J-H| \times \int_{\sigma_W^*}^{\bar{\sigma}} \left[N_A \frac{\sigma_A^* - \bar{\sigma}}{\bar{\sigma} - \sigma_W^*} m - N_W h'(q(s)) f_A(h(q(s))) \right] ds$$

$$= |J-H| \times \int_{\sigma_W^*}^{\bar{\sigma}} [N_W m - N_W h'(q(s)) f_A(h(q(s)))]\, ds$$

where to get the last equality we used the identity

$N_W(\bar{\sigma} - \sigma_W^*) = -N_A(\bar{\sigma} - \sigma_A^*)$. To verify this identity, write

$$N_A(\bar{\sigma} - \sigma_A^*) = N_A\left(\frac{\sigma_A^* N_A + \sigma_W^* N_W}{N_A + N_W} - \sigma_A^*\right)$$

$$= \frac{N_A N_W}{N_A + N_W}(\sigma_W^* - \sigma_A^*)$$

and thus, symmetrically

$$\text{(A2)} \quad N_W(\bar{\sigma} - \sigma_W^*) = \frac{N_A N_W}{N_A + N_W}(\sigma_A^* - \sigma_W^*)$$

$$= -N_A(\bar{\sigma} - \sigma_A^*).$$

Now, let us return to expression (A1); from the definition of m we have

$$m = \min_{z \in [q(\sigma_A^*), q(\bar{\sigma})]} f_A(z)$$

$$= \min_{z \in [h(q(\sigma_W^*)), q(\bar{\sigma})]} f_A(z)$$

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 $s \in [\sigma_W^*, \bar{\sigma}]$. 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 $s \leq \bar{\sigma}$ we have $q(s) \geq q(\bar{\sigma})$. Thus, for any $s \in [\sigma_W^*, \bar{\sigma}]$ we have

$$m \geq \min_{z \in [h(q(s)), q(s)]} f_A(z).$$

Substituting for m into expression (A1) we get

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

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

REFERENCES

Agnew, Robert. "Foundation for a General Strain Theory." *Criminology*, February 1992, *30*(1), pp. 47–87.

Anderson, Elijah. *StreetWise: Race, class, and change in an urban community.* Chicago: University of Chicago Press, 1990.

Arnold, Barry C. *Majorization and the Lorenz order: A brief introduction.* Heidelberg: Springer-Verlag Berlin, 1987.

Arrow, Kenneth J. "The Theory of Discrimination," in O. Ashenfelter and A. Rees, eds., *Discrimination in labor markets.* Princeton, NJ: Princeton University Press, 1973, pp. 3–33.

Bar-Ilan, Avner and Sacerdote, Bruce. "The Response to Fines and Probability of Detection in a Series of Experiments." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 8638, December 2001.

Beck, Phyllis W. and Daly, Patricia A. "State Constitutional Analysis of Pretext Stops: Racial Profiling and Public Policy Concerns." *Temple Law Review*, Fall 1999, *72*, pp. 597–618.

Becker, Gary S. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, March–April 1968, *76*(2), pp. 169–217.

Benabou, Roland and Ok, Efe A. "Mobility as Progressivity: Ranking Income Processes According to Equality of Opportunity." Mimeo, Princeton University, 2000.

Chander, Parkash and Wilde, Louis L. "A General Characterization of Optimal Income Tax Enforcement." *Review of Economic Studies*, January 1998, *65*(1), pp. 165–83.

Coate, Stephen and Loury, Glenn C. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, December 1993, *83*(5), pp. 1220–40.

Ehrlich, Isaac. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy*, May–June 1973, *81*(3), pp. 521–65.

______. "Crime, Punishment, and the Market for Offenses." *Journal of Economic Perspectives*, Winter 1996, *10*(1), pp. 43–67.

Fang, Hanming. "Disentangling the College Wage Premium: Estimating a Model with Endogenous Education Choices." Mimeo, Yale University, April 2000.

Gibeaut, John. "Profiling: Marked for Humiliation." *American Bar Association Journal*, February 1999, pp. 46–47.

Jakobsson, Ulf. "On the Measurement of the Degree of Progressivity." *Journal of Public Economics*, January–February 1976, *5*(1–2), pp. 161–68.

Kennedy, Randall. *Race, crime and the law*. New York: Pantheon Books, 1977.

Knowles, John; Persico, Nicola and Todd, Petra. "Racial Bias in Motor-Vehicle Searches: Theory and Evidence." *Journal of Political Economy*, February 2001, *109*(1), pp. 203–29.

Lehmann, Eric L. "Comparing Location Experiments." *Annals of Statistics*, 1988, *16*(2), pp. 521–33.

Levitt, Steven D. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *American Economic Review*, June 1997, *87*(3), pp. 270–90.

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

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

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

O'Harrow, Robert, Jr. "Intricate Screening of Fliers in Works: Database Raises Privacy Concerns." *Washington Post*, February 1, 2002, p. A1.

Polinsky, A. Mitchell and Shavell, Steven. "The Fairness of Sanctions: Some Implications for Optimal Enforcement Policy." Harvard Olin Center Discussion Paper No. 247, December 1998.

Ramirez, Deborah; McDevitt, Jack and Farrell, Amy. *A resource guide on racial profiling data collection systems: Promising practices and lessons learned.* U.S. Department of Justice, Washington, DC, November 2000. [Available online at 〈http://www.ncjrs.org/pdffiles1/bja/184768.pdf〉.]

Scotchmer, Susan. "Audit Classes and Tax Enforcement Policy." *American Economic Review*, May 1987 (*Papers and Proceedings*), *77*(2), pp. 229–33.

Slobogin, Christopher. "The World Without a Fourth Amendment." *UCLA Law Review*, October 1991, *39*(1), pp. 1–107.

Spitzer, Eliot. *A report to the people of the state of New York from the office of the attorney general.* New York: Civil Rights Bureau, December 1, 1999.

Thompson, Anthony C. "Stopping the Usual Suspects: Race and the Fourth Amendment." *New York University Law Review*, October 1999, *74*(4), pp. 956–1013.

U.S. Customs Service. *Report on personal searches by the United States Customs Service, personal search review commission.* Personal Search Review Commission, Washington, DC. [Available online at 〈http://www.customs.gov/personal_search/pdfs/full.pdf〉.]

Verniero, Peter and Zoubek, Paul H. *Interim report of the state police review team regarding allegations of racial profiling*. NJ Attorney General's Office, April 20, 1999.

Weinstein, Henry; Finegan, Michael and Watanber, Teresa. "Racial Profiling Gains Support as Search Tactic." *Los Angeles Times*, September 24, 2001, p. A1.

Wilk, M. B. and Gnanadesikan, R. "Probability Plotting Methods for the Analysis of Data." *Biometrika*, March 1968, *55*(1), pp. 1–17.

Wohlstetter, Albert and Coleman, Sinclair. "Race Differences in Income." RAND Publication R-578-OEO, October 1970.