Abstract: Is racial profiling an efficient police practice? The answer depends on the goal of policing, which might reasonably be thought of as minimizing the amount of illicit activity. Other plausible objectives, however, include maximizing arrests or convictions. We examine police behavior under different objectives, within the context of drug-law enforcement. If police are motivated to arrest and convict carriers of illicit drugs, pursuit of this goal can result in wide disparities in stops-and-searches among racial groups. If the social goal is minimizing total drug possession, however, the disparities may need to be reduced or eliminated. A further concern is whether the goal of policing is implemented at the level of the police force or by individual officers. Divergence between the incentives facing individual police officers and the objective of the police force as a whole leads to inefficiently extensive profiling. Targeted groups are “over-exploited” from a social point of view, as individual police officers noncooperatively pursue a goal of maximizing arrests or convictions. A further contributor to socially excessive profiling is the undermining of public cooperation with the police that can develop from police stops of innocent people.

An earlier version of this paper was presented at the Midwest Economic Association Meetings, Cleveland, March 2001, and at Indiana University. We would like to thank Marcus Stanley, Brent Never, and seminar participants for helpful comments.
Racial Profiling
Michael Alexeev and Jim Leitzel

1. Introduction

“Racial profiling” in crime control is the use of race as one of the characteristics upon which police make decisions concerning whether to stop, question, or search an individual. Typical racial profiling involves targeting racial groups that are perceived to have higher average propensities to commit a certain type of crime. In recent years, racial profiling, particularly as it is directed at African-Americans in the enforcement of drug laws, has become a controversial practice in the United States.¹ Some police forces have explicitly disowned race-based policies, though casual observation as well as statistical information suggests that implicit racial profiling remains a common police routine.² High measured crime rates among African-Americans are occasionally pointed to as evidence that racial profiling is an efficient policy, and on the whole might even benefit African-Americans, who are disproportionally victims of crime.

Racial profiling also brings with it a variety of costs, however. Most important, members of targeted groups have more frequent, and often more unpleasant, encounters with police.³ These encounters can generate hostility between members of minority communities and the police – hostility which can be reflected in an unwillingness to cooperate with the police in terms of reporting crime, providing testimony, or accepting police testimony when called upon to serve as jurors.⁴ Members of targeted groups also take pains to alter their behavior to minimize the potential for unpleasant police contact, by avoiding certain neighborhoods where race-based stops are likely, for instance, or by driving cars that do not attract attention.⁵

A determination of the efficiency properties of racial profiling depends on the social goal of policing. One possibility is that the police, with respect to a particular crime such as drug possession, should attempt to maximize (justified) arrests. In this
case, standard economic reasoning suggests that at the optimum, the marginal impact on arrests of a police-citizen encounter should be equal across racial groups. Nevertheless, in policy discussions, the efficiency of profiling typically is examined not in terms of marginal impacts, but in terms of averages: a comparison of the percentages of stops of various racial groups that lead to arrests. Given the available data, such a focus on averages is understandable, but in general this approach cannot resolve the issue of whether racial profiling is efficient. Further, maximizing arrests is only one potential goal for the police, and is surely not the overall social objective. Profiling that is efficient from the point of view of maximizing arrests may well be excessive with respect to minimizing the underlying amount of illicit activity.

The incentives for “the police” in general and the incentives for an individual police officer can differ. As in standard provision-of-public-goods settings, an individual officer might rationally perceive his or her behavior as being inconsequential relative to the overall behavior of the force, though in the aggregate, of course, the behavior of the police force is determined by the behavior of each of the individual officers. Specifically, individual officers might take the probability of drug carrying by different racial groups as given, though each stop-and-search by an officer affects the overall deterrence provided by the enforcement regime. As a result, police stops of citizens lead to too much deterrence of the targeted group: seemingly optimal behavior from the point of view of individual officers will involve socially excessive profiling. With this consideration, efficiency in policing can be improved by a reduction or elimination of race-based stops, whether the social goal is to minimize drug possession or to maximize the number of convictions.

Arrests and convictions are not the same thing, even if the arrests are justified in terms of the behavior of the arrested individual – that is, even if drugs are not planted on
arrestees by police. While the police can arrest offenders, convictions occur through the court system. But stopping and searching individuals can lead to reduced cooperation between the public and the police. The reduced cooperation, in turn, indicates that some arrested drug carriers will not be convicted at trial, through a form of jury nullification. The presence of this “conviction externality” implies that police efforts to maximize convictions should involve less profiling than occurs when arrest is equivalent to conviction.

We thus identify three mechanisms through which claims of efficient profiling can falter. First, efficiency in serving the goal of maximizing arrests does not imply efficiency in terms of minimizing illegal behavior. Second, efficient profiling by individual police officers can lead to socially excessive profiling in the aggregate. And third, to the extent that profiling reduces cooperation between the targeted group and the police, efficient profiling from the point of view of maximizing arrests will be inefficient in terms of maximizing convictions, or in minimizing illegal activity.

Our model potentially applies to all forms of criminal profiling, not just profiles based on race. Any visible characteristic upon which the police can, in part, base their actions, and which generally is not a choice variable of individuals – sex and age, for instance – could be addressed within the framework of the simple model developed here. While generally we identify efficiency arguments that suggest typical profiling decisions by police will lead to socially excessive stops of the targeted group, we also identify circumstances – particularly when the police have relatively large resources – when the social optimum involves substantial profiling.

The economics literature on racial profiling is quite recent. Knowles, Persico, and Todd (2001; hereafter KPT) examine racial profiling by developing a model in the tradition of the economics of discrimination (Becker, 1957). KPT posit (p. 205, footnote
omitted) that “the police maximize the number of successful searches, net of the cost of searching motorists.” A non-discriminating police officer is one who views the cost of searching to be the same across all racial groups. An unbiased police officer, then, will choose to search that racial subgroup where a search would yield the highest probability of an arrest for carrying contraband (or, that subgroup with the highest probability of a search leading to a “successful” arrest, if the amount of drugs discovered must be substantial for the arrest to count as a success.) As a result, all searched subgroups should yield the same proportion of (successful) arrests, in equilibrium. Using data on searches by Maryland state police in the late 1990s, KPT find that the probability of an arrest is similar across racial groups, so that the hypothesis that the disproportionate stops of African-Americans represent efficient policing cannot be rejected.7

Within our framework, the KPT analysis corresponds to a determination of whether individual police officers are behaving efficiently with respect to the goal of maximizing arrests. Even if the answer to this question is yes, however, profiling can still be socially inefficient, principally for two reasons. First, rational behavior by individual officers (i.e., searching those drivers with the highest probability of carrying drugs) does not imply that the overall police force is maximizing arrests; the individually optimal police behavior leads to equal average arrest rates across groups, irrespective of the marginal impact of a search on arrests. Second, maximizing arrests (net of search costs) is clearly not the overall social goal, and the pursuit of this goal need not minimize the amount of illegal activity. So while the KPT (2001) analysis provides useful evidence that Maryland state police officers on average were not systematically discriminating against African-American drivers, it does not address the question of the social optimality of the police behavior. An extended set of the Maryland data also forms the basis of the legal analysis of racial profiling provided by Gross and Barnes (2002). In a
“revealed preference”-style argument, these authors conclude that the goal of the Maryland car stops and searches seems to be to maximize the chance of uncovering a major drug trafficker.

Another relevant article is Borooah (2001). Again starting from an economics of discrimination model, Borooah assumes that the police are motivated to maximize arrests. The marginal condition for optimality is identified (p. 26), but an assumption (that the elasticity of arrest rates per stop rates is the same across racial groups) is introduced that justifies a focus on average arrest rates. Borooah’s empirical results are similar to those in KPT, as he finds that the disproportionate stops of racial minorities in England in the late 1990s were not inconsistent with efforts to maximize arrests.

Persico (2002) offers a theoretical approach that is in some important respects similar to ours, especially in noting the disjunction between arrest maximization and crime minimization, which she examines from the perspective of individual officers. Persico concentrates on the issue of how marginal deviations from the equilibrium of the individual policeman problem (see our section 2.3) affect the amount of crime, and she looks at how differential opportunities for legal earnings between racial groups influence the marginal deterrence that a police stop provides. She does not, however, examine the incentives of the police force as a whole, nor does she address the potential conviction externality created by racial profiling.

We develop the basic framework in section 2, and demonstrate how minimization of illegal activity leads to different profiling characteristics than attempts by either individual police officers or the entire police force to maximize arrests. We also show that, under plausible circumstances, crime minimization might require targeting those groups that are relatively unlikely to engage in illegal activity. In section 3 we examine profiling in more detail within a simplified version of our framework, one that restricts
marginal deterrence among the racial groups to be equal (and hence, in terms of minimizing carrying, involves no underlying predisposition to stop one group more frequently than another.) As in section 2, in section 3 we again take into account the potential difference in motivation between individual police officers and the police force as a whole. In section 4, we extend the simplified model to include the conviction externality, in the form of jury nullification. Section 5 offers brief conclusions.

2. The Framework

“Drugs” are prohibited substances, but due to a large informal market, people can still procure drugs. The police are interested in apprehending “carriers,” those in possession of one of the proscribed drugs. The threat of arrest and conviction serves as a (generally imperfect) deterrent against carrying drugs.

Let $p \in [0,1]$ be the probability that an individual is stopped by the police. If the individual is carrying drugs when stopped by the police, he or she will be arrested for possession of an illicit substance, and convicted. The individual controls the variable $x \in \{0,1\}$, where $x=1$ represents the decision to carry the drug and $x=0$ represents the decision not to carry the drug. The magnitude of the disutility of being convicted of drug possession is denoted by $v$, where $v \geq 0$. The payoff from not being convicted, $u(x) \geq 0$, depends on whether the individual carries the drug ($x=1$) or not ($x=0$). Individuals have the same utility of carrying, $u(1)$ and disutility of being convicted, but differ in terms of their payoffs from not carrying, $u(0) \equiv u$. An individual’s (expected) utility, then, is:

$$W(x) = (1 - p \cdot x)u(x) - p \cdot x \cdot v.$$ 

Under these assumptions, a person chooses to carry iff $u \leq u(1) - p(v + u(1))$. If a drug carrier is certain to be arrested and convicted ($p=1$), then clearly no carrying takes place.
Now suppose that there are two groups of individuals, “high carriers” and “low carriers.” Police can visually differentiate between high and low carriers before stopping them. The police cannot determine if an individual is carrying, however, unless the person is stopped and searched.

Assume that for the high carriers, the payoff from not carrying \( u \) is random, and is distributed according to density \( f(u) \). For the low carriers, alternatively, \( u \) is distributed with density \( g(u) \). (In section 3, we examine a special case of this set-up that is particularly easy to work with, where \( f(u) \) and \( g(u) \) are uniform distributions over different domains.) High carriers are more likely to have a small payoff from not carrying than are low carriers. Specifically, we assume that there is a constant \( K, 0 < K < u(1) \), such that \( f(u) \geq g(u) \) for all \( u \leq K \) and \( f(u) < g(u) \) for all \( u > K \): the payoffs to not carrying are “bunched more closely” to zero for the high carriers than for low carriers. This assumption is somewhat stronger than the assumption that the distribution of \( u \) for low carriers first-order stochastically dominates the distribution of \( u \) for high carriers.

Let \( \pi_h \) represent the probability that a high carrier chooses to carry drugs, and let \( \pi_l \) represent the probability that a low carrier will carry drugs. The probabilities (or population proportions) of carrying for the high and low carriers, respectively, can be calculated from the cumulative distribution functions \( F(\@) \) and \( G(\@) \) of \( f(\@) \) and \( g(\@) \) evaluated at the payoff to carrying \( u(1) - p(v+u(1)) \):

\[
\pi_h = F(T) = \pi_h(p) \quad \text{and} \quad \pi_l = G(T) = \pi_l(p),
\]

where \( T = u(1) - p(v+u(1)) \). We will assume that these probabilities are twice continuously differentiable. The derivatives of these probabilities with respect to \( p \), the “marginal deterrents”\(^{10} \) associated with police stops, are

\[
\pi_h' = -f(T)(v+u(1)) < 0 \quad \text{and} \quad \pi_l' = -g(T)(v+u(1)) < 0.
\]
Assuming that the disutilities of conviction $v$ and the payoffs from avoiding conviction while carrying $u(1)$ are the same for both groups of carriers, 

$|\pi_h'| \leq |\pi_l'|$ for all $T \geq K$ or, equivalently, as long as $p \leq [u(1)-K]/[v+u(1)] \equiv P$. So, for a low probability of apprehension ($p$ sufficiently small), the absolute value of the marginal deterrence of high carriers does not exceed that of low carriers, when measured at the same probability $p$. In contrast, when $p \geq P$, $|\pi_h'(p)| \geq |\pi_l'(p)|$. Notice also that given the properties of $f(\widehat{\eta})$ and $g(\widehat{\eta})$ assumed above, $F(T) > G(T)$ for all $T$ and, therefore, $\pi_h(p) > \pi_l(p)$ for all $p$. For drug law enforcement on the streets and highways, the probability of apprehension of a given carrier in any single “trip” appears to be quite small.\(^{11}\) As a low probability of apprehension appears to be much more empirically relevant, in the analysis below we will assume that $K$ is sufficiently small and police resources are sufficiently limited so that for either group of carriers, $p \leq P$.

Under our assumptions, enforcement provides deterrence. That is, both $\pi_h'$ and $\pi_l'$ are negative: the higher the probability of conviction, the less likely the carrying of illicit drugs. Is the further restriction, that $|\pi_h'| \leq |\pi_l'|$ for sufficiently small probabilities of detection, sensible? Targeted groups, such as African-Americans, typically get stopped more frequently on a per capita basis than other groups. This claim is consistent both with anecdotal evidence and official data; for instance, the Maryland state police data used in KPT indicate that 63% of the searches were conducted on African-Americans, as opposed to 29% on whites, despite the number of white drivers greatly exceeding the number of African-American drivers. In England and Wales in 1998, per 1000 individuals of the specified race, 145 blacks were stopped by the police, 45 Asians, and 19 whites (Borooah (2001, p. 18)). It appears, then, that the probability of being stopped by the police is much higher among blacks than whites, both in Britain and in the US. Nevertheless, arrests rates (per stop) are not lower on average for blacks, for the samples
in both KPT and Borooah (2001). This suggests that the condition $|\pi_h'| \leq |\pi_l'|$ is plausible, when the “high carrying” group is associated with blacks in Britain or America, and the “low carrying” group is associated with whites.\textsuperscript{12}

Assume that the number of high carriers is equal to the number of low carriers, and that both are normalized to 1. (The total non-police population is thus of size 2.) Let the number of persons of each group stopped and searched by the police be $H$ and $L$, respectively, with $H, L \in [0,1]$. The probability that a high carrier is stopped ($p_h$), therefore, is also $H$, and the corresponding probability for a low carrier ($p_l$) is $L$. The police have limited resources, and cannot stop everyone. Specifically, let $N$ equal the maximum number of stops that the police can make: $H + L \leq N$.

2.1 Minimizing illegal activity

The police must decide how many of their $N$ stops to devote to low carriers and how many to devote to high carriers. Let the $(L, H)$ combination that minimizes the total amount of carrying ($B_h + B_l$) be denoted by $(L^*, H^*)$. That is, $(L^*, H^*)$ solves

\begin{equation}
\min \{B_h(H) + B_l(L)\}, \text{ subject to } H + L \leq N.
\end{equation}

Proposition 1: Let the probability of apprehension for both groups be sufficiently small ($p \leq P$). Then, (i) if problem (*) above has an interior solution and $\pi_h''(p) > 0$ for all $p$ between $L^*$ and $H^*$, then $L^* \geq H^*$; (ii) if problem (*) has an interior solution and $\pi_h''(p) < 0$ for all $p$ between $L^*$ and $H^*$, then $L^* \leq H^*$; and (iii) if problem (*) has a corner solution, then $L^* = N$ and $H^* = 0$.

Proof: Note first that for $p \leq P$, the probability of carrying is positive for both groups. Therefore, at the carrying-minimizing optimum, the constraint $H + L \leq N$ is binding. The problem for the police, then, is to choose $H$ to minimize $B_h(H) + B_l(N-H)$. The first order condition for this problem requires that $\pi_h' = \pi_l'$. Recall also that we assume that when $p \leq P$, $|\pi_h'(p)| \leq |\pi_l'(p)|$.

(i) Let the interior minimum exist and let $\pi_h''(p) > 0$ for all $p$ between $L^*$ and $H^*$. At an interior minimum, $\pi_h' = \pi_l'$. Suppose $L^* < H^*$. Then given the sign of the second
derivative, \( \pi_h'(L^*) < \pi_h'(H^*) = \pi_l'(L^*) \), but by assumption, \( \pi_h'(L^*) \not\geq \pi_l'(L^*) \) for \( p \leq P \); so, we have a contradiction. Therefore, \( L^* \geq H^* \).

(ii) Let the interior minimum exist and let \( \pi_h''(p) < 0 \) for all \( p \) between \( L^* \) and \( H^* \). (The interior minimum in this case implies that \( \pi_h''(L^*) > 0 \) and \( |\pi_h''(H^*)| < |\pi_l''(L^*)| \).) Suppose \( L^* > H^* \). Then given the sign of the second derivative, \( \pi_h''(L^*) < \pi_h''(H^*) = \pi_l''(L^*) \), but by assumption, \( \pi_h'(L^*) \not\geq \pi_l'(L^*) \) for \( p \leq P \); so, we have a contradiction. Therefore, \( L^* \leq H^* \).

(iii) A corner solution could entail either \( L^* = N, H^* = 0 \) or \( L^* = 0, H^* = N \). Given that \( |\pi_h'(p)| \leq |\pi_l'(p)|, \pi_h(0) - \pi_h(N) \leq \pi_l(0) - \pi_l(N) \). Therefore, \( \pi_h(0) + \pi_l(N) \leq \pi_h(N) + \pi_l(0) \), implying that the carrying-minimizing corner solution is \( L^* = N, H^* = 0 \). Q.E.D.

Proposition 1 establishes a set of sufficient conditions for profiling either the low carriers (part (i)) or the high carriers (part (ii)), when the police force has limited resources and is motivated to minimize the amount of carrying. The optimal profiling pattern depends upon the sign of \( \pi_h''(p) \) in the relevant range. The sign of this second derivative is determined by the slope of \( f(T) \). If \( \pi_h''(p) > 0 \), the marginal deterrence of high carriers declines in absolute value as the probability of arrest increases, or, in other words, there are diminishing marginal returns to stops and searches of high carriers. If \( \pi_h''(p) < 0 \), there are increasing marginal returns to searching high carriers. Notice that \( \pi_h''(p) < 0 \) makes it more likely that the second order condition for minimization is not satisfied. In that case no interior solution exists, but we show in (iii) that the corner solutions will entail more stops of low carriers.

The above framework assumes that the shares of high and low carriers in the population are equal to 50 percent each. Dropping this assumption would affect the absolute numbers of high and low carriers who are stopped in the solution of problem (*), but the equilibrium probabilities with which members of each group are stopped are independent of the relative size of these groups.
2.2 Maximizing arrests

In the previous section, the social goal of policing is taken to be the minimization of illegal activity. It might be difficult, however, for the police to operationalize this goal in practice. An alternative that is more easily implemented is to direct the police to maximize the number of arrests and convictions. Formally, the arrest-maximizing problem for the police is:

\[(**) \max \{H \mathcal{B}_h(H) + L \mathcal{B}_l(L)\}, \text{ subject to } H + L \leq N.\]

Denote the solution to the above problem as \(H^{**}\) and \(L^{**}\). If police resources are sufficiently limited so that the constraint is binding, we can eliminate the constraint by substituting \(L = N - H\) into the objective function. The first order condition for this problem is:

\[H \mathcal{B}_h' + L \mathcal{B}_l' - (N - H) \mathcal{B}_l = 0,\]

where \(\mathcal{B}_l\) and \(\mathcal{B}_l'\) are evaluated at \(L = N - H\). Assuming that \(N\) is sufficiently small so that \(p \leq P\), the left-hand side of the above expression is positive when evaluated at \(H = N/2\), because \((\mathcal{B}_h - \mathcal{B}_l) > 0\) and \((\mathcal{B}_h' - \mathcal{B}_l') > 0\). Therefore, if problem \((**)*\) has a unique interior solution where the resource constraint is binding, \(H^{**} > L^{**}\).

Compare the solutions to problems (*) and (**). While minimizing carrying may or may not target the high carriers – depending on the sign of \(\mathcal{B}_h''\) – in the set-up employed here, a police force interested in maximizing arrests and convictions will always choose to focus its resources on the high carriers.

2.3 The individual police officer and the police force

So far we have modeled the behavior of the police force as a whole, in the sense that, given the objective (minimizing illicit activity or maximizing convictions), the
police force takes into account the consequences of its actions for the probability of carrying. An individual officer, however, presumably would view his or her actions as being inconsequential for overall deterrence, because those individual actions are just a drop in the vast bucket of police force activity. In other words, police officers are not unlike taxpayers in a public goods provision problem, who rationally view the total amount of the public good that is provided as essentially independent of their actions—even though it is the aggregate behavior of taxpayers that determines the extent of provision of the public good. Here, it is the effect of an officer’s actions on the probability of carrying that is rationally ignored by the officer.

Let an individual officer, then, take the probabilities of carrying to be fixed, for both high and low carriers. (In general, these fixed probabilities will differ between high and low carriers.) Instructing an officer to try to minimize carrying, then, would provide no guidance at all, as by assumption the officer views the total amount of carrying to be independent of his or her actions. Arrests and convictions, however, will depend on whom the officer chooses to stop. Assume, then, as in section 2.2, that the individual officer would like to maximize the number of convictions that result from the stops that he or she makes. (We denote this problem as (I**), to indicate that it is the individual police officer’s variant of problem (**).) The police officer’s problem is to choose the probability (denoted r, with $0 \leq r \leq 1$) with which the officer stops a high carrier:

$$\text{(I**)} \quad \max_{0 \leq r \leq 1} \ r \cdot \pi_h + (1-r) \cdot \pi_l$$

Interior solutions to problem (I**) occur only if $\pi_h = \pi_l$; otherwise, the officer would always stop the type that yields the greater expected arrest rate conditional upon a stop, $\pi_i$. If all police officers behave in this way, an equilibrium where members of both groups are stopped also can occur only if $\pi_h = \pi_l$. Such an interior equilibrium will be stable: if
somehow the equilibrium is disturbed and $\pi_h > \pi_l$, then all officers begin to stop only high carriers. This behavior would tend to lower $\pi_h$ and increase $\pi_l$. A similar argument applies, *mutatis mutandis*, when $\pi_h < \pi_l$, until the equilibrium where $\pi_h = \pi_l$ is reestablished.

At the equilibrium, each officer is indifferent between stopping a high carrier and stopping a low carrier. But a symmetric equilibrium, in which all of the (identical) police officers are behaving identically, further requires the “market equilibrium” (or “market clearing”) condition, $r = H/N$ and $(1-r) = L/N$.

An officer attempting to maximize arrests will concentrate his or her efforts on whatever group is most likely to be carrying, the group for which $\pi$ is highest. Given our assumption that $G(u)$ first-order stochastically dominates $F(u)$, $\pi_h$ must be greater than $\pi_l$ when both probabilities are evaluated at the same number of stop-and-searches. An officer who thinks that both groups are being stopped with equal probability, then, will target the high carriers. As the number of stop-and-searches of high carriers increases relative to that of low carriers, the carrying probabilities converge. The equilibrium occurs when the carrying probabilities are identical ($\pi_h = \pi_l$), which can happen only when the high carriers are being stopped more frequently than the low carriers ($H > L$).

The discussion in sections 2.2 and 2.3 can be summarized in the following proposition:

Proposition 1A. The solution to the individual policeman problem ($I^{**}$) involves more stop-and-searches of high carriers than of low carriers. Moreover, if the police force-level arrest maximization problem ($**$) has a unique interior maximum, then its solution also satisfies $H^{**} > L^{**}$.

Propositions 1 and 1A demonstrate that when either the police or individual policemen are motivated to maximize arrests, they have incentives to engage in “standard” racial profiling. Although under some conditions targeting high carriers may
also be efficient from a crime minimizing perspective, generally speaking, there will be
tension between the goals of arrest maximization and crime reduction. The tension arises
because carrying minimization (at an interior solution) involves using stops to equate
marginal deterrence between high and low carriers, whereas arrest maximization by
individual police officers leads to equating the average probability of carrying.\textsuperscript{13}

3. The Linear Model

Proposition 1 looked at optimal police behavior when the goal of the police force
is to minimize drug carrying. The discussion of Proposition 1 noted that the relative
values of “marginal deterrence” ($B_h'$ and $B_l'$) and the sign of the derivative of marginal
deterrence for the high carriers ($B_h''$) are important factors in how police stops should be
allocated among high and low carriers. This section examines police behavior when
marginal deterrence is constant and equal for both high and low carriers, so that any
racial targeting is undertaken for reasons divorced from marginal deterrence
considerations.

Consider a more restricted version of the earlier model, where payoffs to not
carrying are drawn from uniform distributions, and where we make further assumptions
concerning the payoffs of carrying and of being convicted.\textsuperscript{14} Specifically, assume that the
payoff to not carrying, $u$, is uniformly distributed on $[0, R_h]$ for high carriers, and on $[0, R_l]$ for low carriers, with $0 < R_h < R_l < 1$. Under these simplifications, the probability of
carrying for low carriers becomes $\pi_l = T/R_l$, while the probability of carrying for high
carriers is $\pi_h = T/R_h$, where as before, $T = u(1)-p(v+u(1))$.

By having $R_h$ and $R_l$, the maximum payoffs to not carrying, also figure in the
payoffs to avoiding conviction while carrying, we can reduce our set-up to a particularly
simple case. Specifically, let the payoff to “high” carriers from avoiding conviction,
though carrying, be \( u_h(1) = R_h \) while the disutility of conviction for high carriers is \( v_h = 0 \). The payoff to “low” carriers from avoiding conviction, though carrying, is \( u_l(1) = R_l a \), with \( 0 < a < 1 \), and the disutility of conviction for low carriers is \( v_l = R_l (1 - a) \). Given this set-up, expected utility maximization implies that the probability of carrying for the high carriers reduces to \( \pi_h = 1 - p \), and the low carriers choose to carry with probability \( \pi_l = a - p \), as long as \( p \neq a \). We term this case the linear model, because the probability of carrying is a linear function of the probability of apprehension. Further, there is no social “bias” towards stopping either high or low carriers for the purposes of minimizing carrying, as the marginal deterrence (\( d\pi_h/dp \) and \( d\pi_l/dp \)) is constant and the same for high and low carriers at all interior solutions (i.e., as long as carrying is not eliminated within either group.) We will concentrate on the linear model in what follows, both because of its simplicity and because in this case, marginal deterrence considerations alone provide no rationale for profiling, if carrying minimization is the social goal.

As in section 2, let the number of high carriers equal the number of low carriers, and let both be normalized to 1. Again, \( H \) and \( L \) (\( H + L \neq N \)) represent the number of high and low carriers stopped by the police, and also the respective probabilities of being stopped. The probabilities that high and low carriers choose to carry, therefore, are \( \pi_h = 1 - H \) and \( \pi_l = a - L \), with \( 0 < a < 1 \).

### 3.1 Minimizing carrying

Suppose that the social goal is to minimize the overall amount of drug carrying. The problem facing the police, then, is to allocate their \( N \) stops between \( H \) and \( L \) in such a way as to minimize \( B_h + B_l \).

Proposition 2: The total amount of carrying, \( B_h + B_l \), is independent of how the police allocate their \( N \) stops between high carriers and low carriers, except at corner solutions.
Proof: Total carrying is $B_h + B_l$; but, $B_h = 1-H$ and $B_l = a-L$, as long as both $1-H$ and $a-L$ are non-negative. So total carrying equals $1+a-H-L$. With $H+L = N$, at interior solutions, total carrying equals $1+a-N$, irrespective of the values of $H$ and $L$. Q.E.D.

Proposition 2 obviously arises from the linearity in the model that results in the same marginal deterrence effect of a stop upon either high or low carriers: $d\pi_h/dp = d\pi_l/dp = -1$.

In the remainder of sections 3 and 4, we will examine police behavior in four separate versions of the linear model: (1) police force-level arrest maximization; (2) individual police officer arrest maximization; (3) police force-level conviction maximization when jury nullification can result in “not guilty” verdicts for factually guilty defendants (the “conviction externality”); and (4) individual police officer conviction maximization under the conviction externality. The solutions (number of high carriers stopped, number of low carriers stopped) to each of these four problems will be denoted as $(H^A, L^A)$; $(H^{AI}, L^{AI})$; $(H^C, L^C)$; and $(H^{CI}, L^{CI})$, respectively.

3.2 Maximizing arrests

Assume that the police are interested in maximizing the number of arrests and convictions for drug possession. (In this section, arrest is equivalent to conviction, as all those who are arrested are found guilty in court.) The problem the police face is:

$$\begin{align*}
\max & \quad H(1-H) + L(a-L) \\
\text{Subject to:} & \quad H + L \neq N.
\end{align*}$$

As long as $N$ is not too large ($N<(1+a)/2$), the constraint will be binding at the optimum: more police resources lead to more arrests and convictions.\textsuperscript{15} For $N$ sufficiently small ($N<(1-a)/2$), $H^A=N$ and $L^A=0$, i.e., there is a corner solution involving extreme profiling of the high-carrying group.
The interior solution to optimization problem (1) yields \( H^A = (2N+1-a)/4 \) and \( L^A = (2N-1+a)/4 \), for \((1-a)/2 < N < (1+a)/2\). The rate at which high carriers actually choose to possess drugs, at the policing optimum, is \( B_h^A = 1-H^A = (3-2N+a)/4 \), and the corresponding probability for low carriers is \( B_l^A = a-L^A = (3a-2N+1)/4 \).

For instance, if \( a=.5 \) and \( N=.5 \), then \( H^A = 3/8 \) and \( L^A = 1/8 \): the police will stop three times as many high carriers as low carriers, though in the absence of enforcement high carriers are only twice as likely to possess drugs as low carriers. (In the absence of enforcement, the propensity for high carriers to break the law relative to that of low carriers is \( 1/a \).) The “disproportionate” response by the police is even more glaring when compared with equilibrium probabilities of carrying, \( B_h^A \) and \( B_l^A \). Here, \( B_h^A=.625 \) and \( B_l^A=.375 \), so at the equilibrium, high carriers are about 1.79 times as likely to carry drugs as are low carriers, but the high carriers are stopped three times as often.

Proposition 3: Assume that the police are interested in maximizing arrests. For \((1-a)/2 < N < (1+a)/2\), the ratio of stops of high carriers to low carriers that solves problem (1), \( (H^A/L^A) \), is greater than one and exceeds both the ratio of carrying that would attain in the absence of enforcement \((1/a)\), and the ratio of carrying that takes place given the enforcement regime \((B_h^A/B_l^A)\). Furthermore, the ratio of actual carrying between high and low carriers given the enforcement regime is itself less than \((1/a)\), for \( N \) in this range.

Proof: See Appendix.

Proposition 3 indicates that at reasonable values of the parameters, a police force attempting to maximize arrests engages in rather extreme profiling. The rate at which the police stop high carriers can greatly exceed the “exogenous” excess propensity with which high carriers carry drugs, and exceed by an even greater margin the actual ratio of carrying between high and low carriers.

Together, Propositions 2 and 3 indicate that a police force focused on maximizing arrests and convictions might find it best to engage in extreme profiling, even though the
amount of drug carrying violations would be the same under any policing strategy, including a race-neutral one. This result is in accord with that in section 2: maximizing arrests leads to too much profiling, relative to what would occur if the police were motivated to minimize carrying. In section 2, however, the marginal deterrence of the low carrying group was assumed to (weakly) exceed in magnitude that of the high carrying group, at relatively low levels of policing. Here, the marginal deterrence that a police stop imposes on the two groups is taken as constant and equal.

3.3 The Individual Police Officer’s Problem

The previous section indicated that in many circumstances, an arrest-maximizing police force will engage in profiling that is excessive relative to the amount of profiling required by an effort to minimize carrying. Now we look at the potential for an individual police officer to have incentives that are different from those of the force as a whole. As in section 2.3, consider the situation from the point of view of an individual police officer, who, because of agency problems, is motivated to maximize arrests for carrying. The officer can safely ignore the impact of his own activity on the overall deterrence of carrying by the two groups.

Once again, an individual officer takes the probabilities of carrying to be fixed, for both high and low carriers. An arrest-maximizing officer’s problem is to choose the probability \( r \) with which to stop a high carrier:

\[
\max_{0 \leq r \leq 1} \quad r \cdot \pi_h + (1-r) \cdot \pi_l
\]

Recall that interior solutions to problem (2) occur only if \( \pi_h = \pi_l \). When all police officers behave in this way, an equilibrium where members of both groups are stopped also can
occur only if $\pi_h = \pi_l$. At a symmetric equilibrium, the “market equilibrium” condition $r = H/N$ (and hence $(1-r) = L/N$) must be met.

Once again, the probabilities of carrying are $\pi_h = 1-H$ and $\pi_l = a-L$. Equilibrium will be established, therefore, when $1-H = a-L$. Taking into account the resource constraint, $L + H = N$, we obtain $1-H = a-N+H$, and

$$H^{AI} = (N/2) + ((1-a)/2) > H^A = (N/2) + ((1-a)/4);$$
$$L^{AI} = N-H^{AI} = (N/2) - ((1-a)/2) < L^A = (N/2) - ((1-a)/4);$$
and $$r^{AI} = 0.5 + ((1-a)/(2N)),$$
establishing the following result:

Proposition 4: Assuming interior solutions, the behavior of individual officers attempting to maximize arrests results in an equilibrium where the number of stops of high carriers, $H^{AI}$, exceeds $H^A$, i.e., individually optimal behavior by police officers leads to more profiling than the amount that maximizes the total number of arrests.

So in the linear model, when there is no imperative to engage in any profiling at all in terms of minimizing illegal behavior, a police force that aims to maximize arrests will nevertheless engage in substantial profiling (Proposition 3). Furthermore, if stop and search decisions are being made by individual police officers who do not take into account their own actions on deterrence, then the outcome is even more extreme profiling (Proposition 4) – again, despite there being no reason to profile in order to minimize illicit activity. The individual policemen, in essence, look upon the high carriers as a particularly rich “common resource,” and hence “overgraze” that resource through their individual actions. This profiling inefficiency is manifest even if maximizing arrests is taken to be the social goal.

4. The Conviction Externality

So far, we have assumed that all those who are stopped while carrying drugs are convicted of drug possession. We will now replace that assumption with an alternative
specification in which those arrested for drug possession face trials to determine their legal guilt. No longer is an arrest equivalent to a conviction. The jury that will make the determination of guilt will consist of \( m \) “peers” of the defendant. A “peer” is here taken to be someone of the same “carrying class” as the defendant; i.e., apprehended high carriers are judged by \( m \) randomly selected high carriers, and apprehended low carriers are judged by \( m \) low carriers (though some mixing of juries could be permitted without altering the qualitative aspects of the analysis.)

To convict a defendant, the jury must unanimously vote to convict. We maintain the assumption that all those who are charged with drug possession actually were in possession of illegal drugs. Nevertheless, jurors will not always choose to vote to convict. In particular, we assume that those who have previously been stopped and searched by the police will choose to vote to acquit defendants. So, \( H \) percent of high carriers and \( L \) percent of low carriers will vote to acquit at a trial. In other words, the probability that a high carrier juror will vote to convict is \( J_h = 1 - H \), while the corresponding probability that a low carrier juror will vote to convict is \( J_l = 1 - L \). In an \( m \)-person jury, where unanimity is required to return a guilty verdict, the probability of conviction at trial equals \( (J_i)^m \), for \( i=h, l \).

This formulation of the probability of conviction has the undesirable feature that it implicitly assumes that even those previously convicted of drug possession are allowed to serve on juries (and it also ignores the obvious restriction that a defendant cannot serve on his or her own jury.) We justify our choice primarily by reason of tractability. Other formulations, such as restricting juries to those who have not been convicted, have similar features but are much less tractable, particularly for multi-person juries. The main requirement for the results is that the probability of conviction goes down as the number of stops of “peers” rises. Recall also that the approach to jury nullification adopted here is
meant to reflect, more broadly, reduced cooperation with the police by those who have
been “hassled” in the past. The reduced cooperation could take other forms, such as an
unwillingness to report crimes or an unwillingness to provide testimony or other
evidence. Many of these aspects of reduced cooperation would apply to those with
previous convictions, even if they are ineligible to serve on juries. Further, police stops
may have direct ramifications that go beyond the individual who is stopped. Some traffic
stops, for instance, involve passengers as well as the driver of a vehicle, and any police
encounter could influence attitudes towards the police by friends and family of the
stopped individual.

4.1 The police force as a whole

We will once again employ the linear model from section 3, but when making
decisions about whether or not to carry drugs, individuals respond to the probability of
conviction, and not the probability of being stopped and arrested. Furthermore, the goal
of the police will be taken to be to maximize the number of convictions, not the number
of arrests. Let $C_i$ represent the (unconditional) probability that a member of group $i$ will
be convicted of drug possession. $C_i$ is the probability that a member of group $i$ is stopped,
times the probability that the person is carrying, times the probability that someone who
is stopped while carrying will be convicted. Following the linear case developed earlier,
the probability of carrying for high carriers will be $B_h = 1 - H (J_h)^m$, while the probability
that a low carrier actually carries is $B_l = a - L (J_l)^m$, where $0 < a < 1$. (If an individual who is
stopped while carrying is certain of being convicted ($J_h = J_l = 1$), then we have the same
initial set-up as in section 3.)

So, $C_h = H \frac{B_h}{(1 - H)^m} = H \frac{a - L (J_l)^m}{(1 - H)^m}.$

Let $X = H \frac{a - L (J_l)^m}{(1 - H)^m}$; then, $C_h = X \frac{a - L (J_l)^m}{(1 - H)^m}.$
Similarly, \( C_i = L (B_i @ J_i)^m \), with \( J_i = 1 - L \) and \( B_i = a - L(1-L)^m \). Letting \( Y = L @ (1-L)^m \), \( C_i = Y @ (a-Y) \).

Total convictions are \( C_h + C_i = H(1 - H)^m [1-H(1-H)^m] + L(1 - L)^m [a-L(1-L)^m] = X(1-X) + Y(a-Y) \).

We will look both at the problem facing the overall police force (as in section 3.2) and that facing individual police officers (as in section 3.3). Consider the problem facing the police force if it attempts to maximize convictions:

\[
\begin{align*}
\text{(3)} \quad \max \quad & C_h + C_i = X(1-X) + Y(a-Y) \\
\text{Subject to:} \quad & H + L \not\leq N.
\end{align*}
\]

The “negative externality” associated with each additional stop, a reduction in the probability of conviction, suggests that the police resource constraint need not be binding. (This is contrary to the situation in section 3, where as long as \( N < (1+a)/2 \), additional police resources increase arrests and convictions.) The police might find it best to restrict the number of stops below \( N \), to maintain a sufficiently high probability of conviction. The irrelevance of the police resource constraint becomes more likely as the size of juries, \( m \), increases.

Proposition 5: In problem (3), if \( N > 2/(1+m) \), the constraint \( H + L \not\leq N \) will not be binding.

Proof: The unconstrained optimum occurs where \( H=L=1/(1+m) \), as long as the probability that a low carrier carries, \( B_i \), remains nonnegative at \( L=1/(1+m) \). In this case, total stops would be \( H+L=2/(1+m) \), so for \( N \) greater than or equal to \( 2/(1+m) \), additional police resources would not increase the number of convictions. If the unconstrained optimum as described above would violate the non-negativity constraint \( B_i \leq 0 \), then even fewer police resources are needed to maximize convictions. Q.E.D.

The existence of a negative externality of stops upon conviction rates implies that, for sufficiently small \( N \), fewer stops of high carriers should take place, relative to the amount
that would be chosen by the police when all apprehended carriers are convicted. Let $H^C$ represent the optimal choice of $H$ by the conviction-maximizing police force in problem (3).

Proposition 6: At interior solutions of problem (3), if $N < (1+a)/2$, or if $a \geq 0.2$, or if both conditions hold, then $H^C < H^A$.

Proof: See Appendix.

Proposition 6 indicates that the “efficiency” of racial profiling, in terms of maximizing convictions, is reduced if stops lead to less cooperation between the public and the police. If the “conviction externality” is severe enough (perhaps because the number of jurors, $m$, is large, so the possibility of at least one “holdout” is substantial), disproportionate stops ($H > L$) are unjustified from an efficiency perspective.

Recall our previous example, when $a=0.5$ and $N=0.5$, which led in problem (1) to the solution $H^A = 0.375$ and $L^A = 0.125$. Employing $m=1$ in problem (3), these parameters result in $H^C = 0.347$ and $L^C = 0.153$. In accordance with Proposition 6, the conviction externality reduces the disproportionality in stops between high and low carriers; instead of high carriers being stopped three times as often as low carriers, they are stopped about 2.27 times as frequently. As $m$ rises, the extent of profiling declines further, when the police are attempting to maximize convictions. With $a=N=0.5$, at $m=3$, the constraint $H + L \# N$ is just binding, so for $m \geq 3$, profiling (in the sense of a disproportional number of stops based on visual characteristics) disappears entirely: $H^C = L^C = 1/(1+m)$. Specifically, with $a=N=0.5$ and $m=3$, $H^C = L^C = 0.25$.

Recall that the amount of carrying in our original linear model is independent of the allocation of $H$ and $L$ (Proposition 2), so that racial profiling does not result in any advantage in terms of reduced drug possession. With a conviction externality, however,
where deterrence is based on probabilities of conviction as opposed to arrest, total carrying is no longer invariant to the distribution of stops. Nevertheless, the dissonance between a police force that attempts to maximize convictions and the social goal of minimizing carrying is not eliminated by the conviction externality: racial profiling still occurs even though it is not justified from the point of view of minimizing carrying, at least in the tractable case when jury size $m=1$.

With the conviction externality and $m=1$, let $H_{\text{carry}}$ be the number of high carrier stops that minimizes total carrying, and let $L_{\text{carry}}$ represent the number of stops of low carriers that minimizes total carrying: $L_{\text{carry}} = N - H_{\text{carry}}$ when the resource constraint is binding. As the following proposition shows, $H_{\text{carry}}$ is smaller than $H^C$, the number of high carrier stops that maximizes convictions.

**Proposition 7:** Let $m=1$. In the presence of the conviction externality ($J_h = 1 - H$, $J_l = 1 - L$, $B_h = 1 - HJ_h$, $B_l = a - L J_l$), the amount of carrying is minimized when $H_{\text{carry}} = L_{\text{carry}} = N/2$. Further, for interior solutions to problem (3), $H^C > H_{\text{carry}}$.

**Proof:** The total amount of carrying is $\pi_h + \pi_l = 1-H(1-H)+a-L(1-L)$. Given that $L=N-H$, the first order condition for minimizing the above expression is $4H-2N=0$, implying that $H_{\text{carry}} = L_{\text{carry}} = N/2$. Note that the second order condition for a minimum is satisfied. Evaluating the first order condition for problem (3) (this is expression (3A) in the Appendix) at $H=L=N/2$ we obtain

$$(1-2X)(1-2H) - (a-2Y)(1-2L) = (1-N)(1-a-2(X-Y)) = (1-N)(1-a) > 0$$

Given that the second derivative of the number of convictions with respect to $H$ is always negative (see the proof of Proposition 6), the above inequality implies that $H^C > N/2$. Q.E.D.

Summarizing, in problem (1) in section 3.2, attempting to maximize convictions leads to substantial racial profiling, without reducing total carrying. In this section, which incorporates the notion that stopping individuals diminishes cooperation with the police, the carrying-minimizing policy of the police will involve no profiling at all. The
alternative goal of maximizing convictions in problem (3) will lead to some profiling, albeit less than in model (1), though for large enough juries (or some other sufficiently large conviction externality), profiling is abandoned entirely (by Proposition 7).

4.2 Individual police behavior with a conviction externality

Let the jury size be fixed at m=1. As in sections 2.3 and 3.3, an individual police officer will choose r, the probability of stopping a high carrier, in such a way as to maximize convictions, without taking into account his or her own influence on overall deterrence of the conviction rate:

\[
\max_{0 \leq r \leq 1} r \cdot \pi_h \cdot J_h + (1-r) \cdot \pi_l \cdot J_l
\]

Interior solutions to problem (4) occur only if \(\pi_h \cdot J_h = \pi_l \cdot J_l\). A symmetric equilibrium involves, as in section 3.2, the “market clearing” condition that \(r=H/N\) and \((1-r)=L/N\). Interior equilibrium is obtained when \(1-H(1-H))(1-H) = (a-L(1-L))(1-L)\), or, recalling our definitions of X and Y, when

\[
(1-X)(1-H) = (a-Y)(1-L)
\]

Proposition 8: Within the framework of problem (4), assuming interior solutions and assuming that \(N \leq 1/2\), optimal behavior by individual officers results in an equilibrium where \(H^{Cl} > H^{e}\), i.e., individually optimal behavior by police officers leads to more profiling than the amount that maximizes the overall number of convictions. Given Proposition 7, with \(m=1\), this also implies that \(H^{Cl} > H^{carry}\).

Proof. See Appendix.

Proposition 8 indicates that even if society is interested in maximizing convictions (as opposed to minimizing illegal carrying), individually rational behavior by police officers will lead to excessive profiling. (If the goal is to minimize carrying,
individual police behavior tends to drive the result even further from the social optimum.) Once again, by ignoring the effect of their own actions on deterrence, police officers “overdeter” high carriers, from a social perspective. Recall our earlier example, with N=.5 and a=.5. For these parameter values, problem (2) does not produce an interior solution, but for slightly higher N, an interior equilibrium emerges with near complete profiling (H almost equal to N), when individual police officers maximize arrests in the absence of a conviction externality. This near-complete profiling can be compared with our earlier calculation from problem (1) of H^A=.375. Problem (4), with the parameters N=a=.5, produces H^C = .442$ $H^C = .347, in conformity with proposition 8.

This section indicates that conscientious policing by individual police officers need not produce a socially optimal outcome, even if the social goal is to maximize convictions. Here, the rationale is that the targeted group will be overdeterred in the aggregate by individually rational police behavior. As noted in the Introduction, there may be other reasons to suspect that profiling is sub-optimal. In particular, individual police officers, responding to a profile, may not recognize that other police officers are likely to respond similarly to the observable characteristics of a citizen. The same person, therefore, might be subject to a series of stops, each one of which appears to be rational from the point of view of the officer involved.\textsuperscript{16} The disutility (or other social costs) associated with police stops may well be convex in the number of times a given individual is stopped. In other words, for a fixed number of stops, and with all else equal, society might be better off if the stops are widely distributed, rather than concentrated on just a few individuals, “the usual suspects.” Indeed, a wider distribution of stops might make more palpable to citizens the costs incurred by those relatively few individuals who currently are habitually stopped.
5. Conclusions

Decreases in crime in the US in the past dozen years have been nothing short of astounding. Seeming inequities in the treatment of minorities by the police, however, are a continuing cause of concern. Inequities in the treatment of different racial groups by police are not haphazard; rather, they frequently have been the result of implicit or explicit policies to target minorities.

US crime statistics suggest that for many crimes, some racial groups are disproportionately likely to be both perpetrators and victims. A defense of racial profiling, then, is that it reflects efficient policing – there would seem to be little point in focusing police resources on unlikely perpetrators. The main point of our paper is that “efficiency” in policing can be a complicated concept. An arrest-maximizing policy might not lead to minimal levels of crime, or to maximum convictions, either. Further, individually rational behavior by police officers can lead to sub-optimal social outcomes, and too much profiling, whether crime minimization or the maximization of convictions is taken to be the social goal.

Our results indicate that racial profiling can represent a rational strategy for police, either individually or at the level of the force as a whole. Nevertheless, in many circumstances the degree of racial profiling undertaken by police results in sub-optimal outcomes, from the perspective of crime control. “Rational” racial profiling engaged in by unbiased police officers can be inferior to race-neutral policing policies.
Appendix

Proof of Proposition 3: \( H^A = (2N+1-a)/4 \) and \( L^A = (2N-1+a)/4 \). \( L^A > 0 \) implies that \((1-a)/2 < N. \( H^A/L^A = (2N+1-a)/(2N-1+a) \).

To establish that \((H^A/L^A)\) exceeds \(1/a\) in the specified range, define \( x = (H^A/L^A) - 1/a \), or \( x = [(2N+1-a)/(2N-1+a)] - 1/a \). Note that \( x \) approaches infinity as \( N \) approaches the lower bound \((1-a)/2 \) (or as \( 2N \) approaches \( 1-a \)) from above. \( x \) is monotonically decreasing in \( N \):

\[
\frac{M}{M'} = 2(2N-a) - 2(2N+1-a) / (2N-1+a)^2 = (4a-4) / (2N-1+a)^2 < 0. \]
At the upper bound, \( N = (1+a)/2, x = 0 \). So within the specified range, \( x \) is always positive, decreasing from near infinity to near zero.

We will now establish that the ratio of carrying in equilibrium \((B^A_h/B^A_l)\), is less than \((1/a)\), the ratio that would obtain in the absence of enforcement. As we have already shown that \((H^A/L^A)\) exceeds \((1/a)\) in the specified range, this demonstration will simultaneously show that \((H^A/L^A)\) exceeds \((B^A_h/B^A_l)\).

\( B^A_h = (3-2N+a)/4 \) and \( B^A_l = (3a-2N+1)/4 \). So, \((B^A_h/B^A_l) = (3-2N+a)/(3a-2N+1) \). Let \( y \) be the difference between \((1/a)\) and \((B^A_h/B^A_l)\): \( y = (1/a) - (3-2N+a)/(3a-2N+1) \). At the lower bound, where \( N = (1-a)/2, y = (1-a)/2a > 0 \). Furthermore, \( y \) is monotonically decreasing in \( N \):

\[
\frac{M}{M'} = - [(3a-2N+1)(-2) - (3-2N+a)(-2)] / (3a-2N+1)^2 = (4a - 4)/(3a-2N+1)^2 < 0.\]
At the upper bound, \( N = (1+a)/2, y = 0 \). So within the specified range, \( y \) is always positive, decreasing from near \((1-a)/2a\) to near zero. Q.E.D.

Proof of Proposition 6: Recall that we denoted \( X = H(1-H)^m \) and \( Y = L(1-L)^m \), in which case, \( H^C = \arg \max_h \{X(1-X) + Y(1-Y)\}, \) where \( L = N-H \). (We assume here that the resource constraint in problem (3) is binding. If the resource constraint is not binding, then a similar result holds: \( H^A = (1-a)/4 + N/2 \geq (1-a)/4 + (1/(m+1)) > 1/(m+1) = H^C \).

The first order condition (FOC) for this problem is

\[
(3A) \quad (1-2X)(1-2H) - (a-2Y)(1-2L) = 0.\]

From problem (1), at the optimum \( 1-2H^A = a-2L^A \). We will demonstrate that at these values of \( H^A \) and \( L^A \) the left hand side of \((3A)\) is negative.

\[
(1-2X)(1-2H^A) - (a-2Y)(1-2L^A) = (1-2H^A) - 2X + 4H^AX - a + 2aL^A + 2Y(1-2L^A) = a - 2L^A - 2X + 4HX - a + 2aL^A + 2Y(1-2L^A) = 2L^A(a-1) - 2X(1-2H^A) + 2Y(1-2L^A).\]
Note that $1-2L^A=1-a+a-2L^A=1-a+1-2H^A$. Therefore, the above expression can be rewritten as


Notice that if $X$ and $Y$ are evaluated at $H^A$ and $L^A$, $X-Y=(1-a)(1-N)/2$. (Recall that $H^A=(1-a+2N)/4$ and $L^A=N-H^A$.) Also, $Y<L^A$. Therefore, the latter expression is negative as long as $2H^A \leq 1$ or $N \leq (1-a)/2 = (1+a)/2$. It is also straightforward to show, using the values of $H^A$ and $L^A$, that this expression would be negative for all feasible $N$ as long as $a \geq 0.2$.

Therefore, we have shown that the left hand side of the first order condition (3A) is negative when evaluated at the optimum to problem (1), $H^A$. This implies that $H^C < H^A$ if the derivative of the left hand side of (3A) is also negative. This second derivative of the number of convictions with respect to $H$ is

$$-2[(1-2H)^2+(1-2X)+(1-2L)^2+(a-2Y)] < -2[1-2X+a-2Y] < 0$$

because

$$X = H(1-H) \leq 1/4 \text{ and } L(1-L) \leq 1/4.$$

**Proof of Proposition 8:** We need to demonstrate that the $H$ that solves (4 and 5) is greater than the $H^C$ that solves the first order condition for problem (3):

(3A) $(1-2X)(1-2H)-(a-2Y)(1-2L)=0$.

In order to accomplish this, we will show that the left hand side of (3A), when evaluated at the $H^C$ that solves (4), is negative. Given that the second derivative of the left hand side of (3A) is negative (see the proof of Proposition 6), this would imply that $H^C$ is smaller than $H^C$.

Let us first establish that (5) implies that $H > L$. If, alternatively, $H \# L$, then $(1-H)\#(1-L)$, and in order for (5) to hold, $1-X \# a-Y$. But $1-X \# a-Y$ implies that $1-X-a+Y \# 0$, so


Under the assumption that $H \# L$ and $N<1$, this inequality implies that $a \neq 1$, which is not the case in our setup. Therefore, $H^C > L^C$.


29
Condition (5) can be rewritten as \((1-H-X+XH) = a-aL-Y+YL\). Substituting this condition into the previous expression (that is, evaluating the left hand side of (3A) at \(H=H^C\)) yields

\[(a-aL-Y+YL)-H-X+3XH-a+2aL+2Y-4YL = aL+Y-3YL-H-X+3XH.\]

Recalling that (with \(m=1\)), \(X = H \frac{1}{1-H}\) and \(Y = L \frac{1}{1-L}\), this expression can be rewritten as

\[aL+L-4L^2-2H+4H^2-3(H^3-L^3)\]. We can now take this expression, and dropping the cubic term (which is negative) and replacing \(a\) (which is less than 1) with one, we get an expression which exceeds the previous one, i.e.,

\[aL+L-4L^2-2H+4H^2-3(H^3-L^3) < 2(L-H)-4(L^2-H^2) = 2(L-H)(1-2(H+L)) = 2(L-H)(1-2N) \leq 0,\]

as long as \(N \leq \frac{1}{2}\) and \(H > L\). Q.E.D.

(Alternatively, instead of the condition that \(N \leq \frac{1}{2}\), we could have imposed a condition on the parameter \(a\). Examining the left hand side of (3A),


Notice that \(2XH=2H^2(1-H)\) reaches its maximum at \(H=2/3\) and this maximum is equal to \(8/27\). Therefore, \((1-L)a-1+2XH < (1-L)a-21/27 < 0\) as long as \(a \leq 21/27\). Q.E.D.
References


Endnotes

1. See, for instance, Harris (1999, 2002). For the purposes of examining profiling, we take the illegality of certain drugs as given, though optimal regulatory schemes with respect to drugs are an item of contention. See, for instance, Miron and Zweibel (1995).

2. See, for instance, the anecdotal and statistical information compiled in Harris (2002), Meeks (2000), and Fagan and Davies (2000).

3. Most of those stopped via racial profiles are innocent citizens, too. For example, Harris (1999) notes the settlement of a lawsuit against Eagle County, Colorado, brought by the American Civil Liberties Union on behalf of 402 black and Latino motorists who were stopped because they matched a drug courier profile. None of the plaintiffs were ticketed or arrested for drug offenses.


5. For some of the costs of inequities in policing, see Harris (2002), Cole (1999), and Kennedy (1997).


7. Fagan and Davies (2000) analyze 1998 stop-and-frisk data from New York City, and find that the number of stops of black and Hispanic citizens per arrest was higher than for whites, particularly for stops premised on suspicion of weapons violations or violent crimes.

8. See also the comment and reply, Chakravarty (2002) and Borooah (2002).

9. Alternatively, individuals could differ with respect to \( v \), their disutility from being convicted. Such an approach leads to similar qualitative results, but the alternative assumption that \( u \) is random makes the analytics in section 3 more straightforward.

10. The phrase “marginal deterrence” is sometimes used for a different purpose, indicating that more serious crimes should involve higher punishments; otherwise, someone apprehended for a minor crime like jaywalking might find it worthwhile to attempt to use extreme violence, for instance, to escape from the police. In other words, the punishment structure should be such that those who engage in minor offenses will still be deterred “at the margin,” that is, from engaging in more serious crime. See Stigler (1970).

11. Gross and Barnes (2002), for instance, suggest that the probability that any given car being stopped during a trip through the 48.5 mile corridor of Interstate 95 for which they have data was approximately 1 in 1250 during 1997-99.

12. Of course, the fact that the arrest rates for blacks appear to be about the same as for whites despite a higher probability of being searched does not imply that marginal
deterrence at low levels of policing is, of necessity, greater for whites than for blacks, though it is suggestive. Alternatively, it might be the case that $\pi_h(0) > \pi_l(0)$ at zero levels of policing, so that arrests rates for blacks might still be as high as for whites even if blacks are more deterrable at the margin than are whites.

13. Persico (2002) establishes a result analogous to Proposition 1A.

14. While this setup is generally simpler than the one in the previous section, it also relaxes the assumption that $v$ and $u(1)$ are the same for both groups of carriers.

15. Note that carrying, and hence arrests, could be eliminated if $N \equiv (1+a)$. This outcome would not be attractive to a police force interested in maximizing arrests and convictions. For illicit drug possession and many other crimes, however, elimination is not a realistic option given the level of enforcement resources, and arrest maximization appears to be consistent with police behavior.

16. Consider the famous “San Diego walker” case (Kennedy (1998, p. 35).) Edward Lawson, a black man, was detained some 15 times in the space of less than two years, often while walking without identification in the “wrong” neighborhood. Kolendar v. Lawson 461 U.S. 352 (1983).