DEPARTMENT OF ECONOMICS

WORKING PAPER SERIES

2004-11

McMASTER UNIVERSITY

Department of Economics
Kenneth Taylor Hall 426
1280 Main Street West
Hamilton, Ontario, Canada
L8S 4M4

http://socserv.socsci.mcmaster.ca/~econ/

Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate*

David Bjerk

Working Paper
Department of Economics
McMaster University

*For most current version of this paper, go to http://socserv.socsci.mcmaster.ca/bjerk/papers.html

Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate

David Bjerk*
Department of Economics, McMaster University
bjerkd@mcmaster.ca

February 25, 2004

Abstract

Using a model similar to labor market models of statistical discrimination, I describe how and why racial profiling can arise even when law enforcement officers are racially unbiased. Specifically, if one racial group has a higher fraction of individuals who are at risk of committing the relevant type of crime than another, and if law enforcement officers can observe a noisy signal of guilt in addition to an individual's race, then it will be optimal for officers to treat observationally equivalent individuals of different races differently. Moreover, this model can be used to show how the effect of a racially colorblind policy on the overall crime rate for a particular type of crime will depend on the racial make-up of the relevant jurisdiction, the relative proportions of each racial group that are at risk of choosing to commit that crime, the proportion of the relevant population that officers can observe, the magnitude of the punishment for that particular type of crime, and distribution of the benefits to committing that particular crime. The implications coming from this analysis are then applied and analyzed with respect to two specific contexts—highway patrol vehicle searches for drugs or weapons, and border patrol investigations of foreign entrants for terrorist connections.

*Thanks to Seungjin Han and participants in the McMaster Economics Seminar for their helpful comments. JEL codes K14, K42, J71

1 Introduction

Racial profiling by law enforcement officers, or "using race as a factor in conducting stops, searches, and other investigative procedures," [Bush, 2001] has attracted a vast amount of attention over the last decade. In a Gallup Poll from 1999, more than half of Americans polled believed that police actively engage in the practice of racial profiling, including 56 percent of whites surveyed and 77 percent of blacks surveyed. Moreover, the same survey showed that 72 percent of black men between the ages of 18 and 34 believed they had been stopped because of race [Ramirez, McDevitt, and Farell, 2000]. These perceptions of widespread racial profiling, and the view that treating individuals differently based on their race is unethical and unconstitutional, have led to numerous efforts to implement policies that would eradicate racial profiling by law enforcement officers.[1]

If racially unequal behavior by law enforcement officers were soley the result of racial bias, then ethical concerns undoubtedly justify the imposition of such anti-profiling policies. However, racial profiling could also result from a form of statistical discrimination, where racially unbiased law enforcement officers use race as a signal of other unobservable traits that relate to criminal proclivity. While the ethics and constitutionality of this type of discrimination are also questionable, such statistical discrimination type behavior may persist because it is the most efficient way for officers to maximize the expected guilt rate among individuals they investigate or search. Hence, banning racial profiling may come at the cost of increasing the overall crime rate and/or increasing the cost of crime control.

This paper looks at this cost theoretically, describing how the overall crime rate may be affected by enforcing colorblind police behavior. In doing so, I first devise a model of law enforcement and individual behavior showing how and why statistical discrimination may arise in equilibrium. This model is similar to those used for examining statistical discrimination in the context of the labor market, most notably Coate and Loury [1993]. In the model used here, individuals decide whether or not to commit a particular crime based on the benefits they incur from doing so, the cost to getting caught for doing

[1]More than twenty states have passed legislation prohibiting racial profiling and/or requiring jurisdictions within the state to collect data on law enforcement stops and searches [Racial Profiling Data Collection Resource Center, 2003].

so, and their belief concerning the probability they will be caught for doing so. A law enforcement officer on the other hand, decides on a case by case basis whether or not to investigate each individual he observes based on an endogenous guilt signal emitted by the individual, the officer's belief regarding the overall guilt rate of the individual's racial group, the benefit the officer incurs from a successful investigation, and the cost to performing an investigation.

In equilibrium, all beliefs correspond to the truth and the optimal strategy for unbiased officers is to treat observationally similar members of two racial groups differently as long as one of the groups has a higher proportion of individuals who can be classified as "at risk" of committing the particular relevant crime. While this result in and of itself is relatively unsurprising, the model also shows that if statistical discrimination is the cause of racially unequal investigation rates, then it must be the case that the more frequently investigated race has a higher overall guilt rate than the other race, while the guilt rate for *investigated* members of the more frequently investigated race can be greater than, equal to, or less than the guilt rate for the *investigated* members of the other race. Both of these results are in direct contrast to the previous work that has attempted to examine statistical discrimination in the context of racial profiling in policing [Knowles, Persico, and Todd, 2001; Persico, 2002; Hernandez-Murillo and Knowles, 2003]. Moreover, because of these results, this paper reveals that in order to identify whether racially unequal investigation rates can be due to statistical discrimination rather than racial bias among law enforcement officers, it is necessary to have data regarding the overall guilt rate of each group, not just the guilt rate among the investigated members of each racial group (the only data that is currently available).

After describing the equilibrium when unconstrained law enforcement officer behavior results in statistical discrimination, I derive the equilibrium that would arise if a theoretical policy were implemented such that officers are constrained to behave in a perfectly colorblind manner. I show that the degree to which such a colorblind policy alters the crime rate will depend on the relevant crime in question and characteristics of the relevant jurisdiction. In general, if a group b has a higher fraction of individuals who are at risk of committing the relevant crime than group a, then a colorblind policy will lead to smaller increases (or possibly even decreases) in the overall crime rate (i) the smaller the proportion of the jurisdiction that is made up of group b, (ii) the more equal the two groups are in terms of their relative fractions of individuals who are at risk of committing the relevant crime, (iii) the smaller the proportion of the relevant population officers are able to observe, (iv) the smaller the penalty associated with the relevant crime, and (v) the smaller the proportion of individuals who incur a large benefit from committing the relevant crime relative to those who only incur a small benefit.

Finally, I discuss the results of this model in two contexts relevant for racial profiling—motor vehicle searches by highway patrol and investigation of foreign nationals by border patrol officers at U.S. airports and highways. The model suggests that the crime rate cost to a colorblind policy is likely to be quite small (and possibly even negative) with respect to highway motor vehicle searches, but may be substantially larger with respect to investigation of foreigners coming into a country such as the United States.

2 Previous Literature

Despite the fact that economists have long examined statistical discrimination in the context of labor markets [Phelps, 1972; Arrow, 1973; Aiger and Cain, 1977; Lundberg and Startz, 1983; Coate and Loury, 1993; Cornell and Welch, 1996; Moro and Norman, 2003], economists have just begun applying such thinking to racial inequality in other areas, such as police searches. Currently, there exist only a handful papers in this area, namely Farmer and Terrell [2001], Knowles, Persico, and Todd [2001], Persico [2002], and Hernandez-Murillo and Knowles, [2003]. Since these latter three papers all use similar underlying models to obtain their results, I will subsequently refer to these three papers as "KPT and the related papers," since Knowles, Persico, and Todd [2001] was the first of these to appear in print.

KPT and the related papers all essentially focus on the Nash equilibrium of a simultaneous move game, where police decide what proportion of each group to investigate in order to maximize the guilt rate among the investigated, and members of each group decide whether or not to commit the relevant crime (i.e. carry contraband) based on whether the benefit to doing so exceeds the expected cost. In Persico [2002], only race is observable by police, meaning groups simply refer to racial groups. In Knowles, Persico, and Todd [2001] and Hernandez-Murillo and Knowles [2003] however, police can also observe other exogenous characteristics besides race, meaning groups are made up of all individuals with similar race/characteristic combinations. In what follows, I will use "groups" to refer to the more general race/characteristic combinations, but the discussion would generally be the same using the more restrictive definition used in Persico [2002].[2]

The equilibrium that results from the environment proposed by KPT and the related papers is one in which police employ investigation rates (possibly differing) over each group such that, given these investigation rates, the fraction of each group that chooses to carry contraband (i.e. the "guilt rate" for each group) is equal across all groups and is such that police are indifferent between investigating and not investigating members of each group. Such behavior can be maintained in equilibrium because police are indifferent between investigating and not investigating members from each group, meaning any search rate between zero and one will be weakly optimal from the police officers' perspective. Moreover, this equilibrium is unique, because if any group had a crime rate higher than the rate at which police officers were indifferent between investigating and not investigating, then police would investigate all members of this group with probability one, and the construction of the model is such that, with an investigation probability of one, it would not be optimal for any members of this group to carry contraband. Hence, such a situation cannot be maintained as an equilibrium in this environment. A similar argument can be made for a situation where one group had a lower crime rate than that which makes police officers indifferent between investigating and not investigating.

This equilibrium implies that unequal investigation rates across races are consistent with racially unbiased officers as long as, for any given observable characteristics, members of one race generally incur higher benefits to carrying contraband (or incur lower costs of arrest) than the other. With such an assumption, officers have to investigate groups containing members of the more crime prone race at higher rates in order to make it optimal for all groups to carry contraband at the same equilibrium rate. This implies that, in equilibrium, a higher fraction of the more crime prone race will be investigated. Knowles, Persico, and Todd [2001] (as well as Hernandez-Murillo and Knowles

[2]Furthermore, in KPT and the related papers, police simply decide whether or not to search individuals. In this paper, I use the more general term "investigate," where a search is assumed to be the type of investigation relevant to highway patrol officers.

[2003]) use this equilibrium to motivate a "test" for whether or not observed racially unequal search rates in particular highway stretches are due to statistical discrimination or racial bias among state troopers. In particular, they claim that unequal search rates across racial groups are consistent with unbiased officers as long as guilt rates among the searched are equal across races (as required by their equilibrium with unbiased officers). If guilt rates differ across racial groups, then Knowles, Persico, and Todd [2001] argue that officers must be biased against the race with the lower guilt rate, regardless of which race is searched at a higher rate. Persico [2002] also uses the equilibrium result that guilt rates must be equal across all searched groups when proving his lemmas regarding how a policy that makes search rates more racially fair will affect the overall crime rate.

While these papers provide an important starting point for thinking rigorously about statistical discrimination in the context of policing, the equilibria discussed in these papers lead to some unappealing results for this context. First, since each group has the same guilt rate in equilibrium, police choose who to investigate at random. Not only is this behavior generally illegal in many real world contexts,[3] but it is also quite unrealistic, as it means that the guilt rate among those investigated should equal the guilt rate among those who are not investigated. In practice, this implies that officers' experiences do not make them any better at spotting law-breakers than a machine that picks individuals at random.

Second, and maybe even more importantly, the implicit beliefs necessary to maintain this equilibrium make the resulting behavior difficult to reconcile with officers being racially unbiased. Specifically, as discussed above, in order to maintain the necessary equilibrium investigation rates, officers must be indifferent between investigating and not investigating members of each group. This implicitly assumes that officers believe each racial group has equal guilt rates, making officers indifferent between investigating members of different races with respect to maximizing the guilt rate among the investigated. Given each officer is indifferent between which race to search in terms of maximizing successful investigation rates, it seems problematic to focus on an equilibrium where unbiased officers still choose to systematically investigate one race at a

[3] For example, before conducting a search, New Jersey State Troopers must "be prepared to point to specific facts and the reasonable inferences that can be drawn from those known facts to support the suspicion that this particular individual may be carrying a concealed weapon." [Verniero and Zoubek, 1999]

higher rate than the other, as such behavior can be argued to only be consistent with racial bias.[4]

In addition to these unappealing outcomes, the equilibrium and all the results coming from KPT and the related papers hinge on the fact that law enforcement officers will never find it optimal to investigate a particular group (i.e. all individuals with a particular race/characteristics combination) with probability one in equilibrium. As mentioned above, the reason this is true is because in these models, if officers investigate a particular group with probability one, every member of that group carrying contraband would be caught. Therefore, no member of that group would find it optimal to carry contraband, making it not optimal for police to use the strategy of investigating that group with probability one.

However, this necessary assumption that police never find it optimal to investigate any groups with probability one in equilibrium requires two rather strong implicit assumptions. First, it requires that police can observe or come in contact with every member of each group, so that when police choose to search members of a particular group with probability one, they will actually search every member of that particular group. To see why this implicit assumption is necessary, say law enforcement officers could only observe a small fraction of the population. If this was the case, then even if officers decide to investigate members of a particular group they observe with probability one, many individuals in that group may still find it optimal to carry contraband since there is a high probability they will not be observed, and therefore not investigated. Since a significant fraction of this group still chooses to carry contraband even if they know they will be investigated with probability one if they are observed, officers may still find it optimal to investigate every member of this group they observe with probability one. Since in many jurisdictions where racial profiling has become an issue (highways, stores) officers are unlikely to be able to observe every individual, this assumption may be problematic.

Second, for it to be true that police never investigate one group with probability one in equilibrium, all observed characteristics must be exogenous and deterministic.

[4]To state this point another way, say officers are defined to be racially unbiased only if they use similar investigate rates across races when they believe races to be equally likely of being guilty. Under this definition of unbiasedness, it is not possible for the equilibrium coming out of KPT to be maintained with racially unbiased officers.

To see why this is true, say characteristics arose endogenously, meaning an individual obtains a particular characteristic only if he is guilty of committing the relevant crime. Then, it will always be the case that officers investigate individuals with this characteristic with probability one. More generally, say relevant characteristics are endogenously probabilistic, meaning the probability that an individual exhibits a certain characteristic depends probabilistically on whether or not the individual is guilty of the relevant crime. In this case, even if individuals believe that if they exhibit a certain characteristic they will be investigated with probability one, a certain proportion of individuals who exhibit that characteristic may still be found to be guilty because they believed there to be some probability they would not exhibit that particular characteristic. Hence, police still might find it optimal to investigate every individual with that particular characteristic, as they still might find enough guilty individuals with that characteristic to make such a strategy optimal.

Generally, other than race, age, and gender, most characteristics relevant to police search behavior can be argued to be endogenously probabilistic. For example, in the case of motor vehicle searches, if there is a strong odor of marijuana smoke in a car, an officer will almost surely search the car. However, an individual smoking marijuana in the car may believe that if he rolls down the windows and sprays air freshener, a police officer might not smell the marijuana smoke. Hence, a strategy of searching every car that has a strong odor of marijuana may not deter all drivers from carrying or smoking marijuana in their car. A similar case could be made for drivers who are erratic, disoriented, visibly nervous, have prior records, or have visible contraband related materials (e.g. drug paraphernalia, empty beer cans, bullets, etc.).

The model presented in the next section proposes an environment where law enforcement officers can only investigate individuals they observe, and it is possible that they cannot observe all individuals. Furthermore, in addition to race, police can observe a characteristic that arises endogenously, where some realizations of this characteristic are more likely for individuals who committed the relevant crime. As will be shown, the equilibrium that arises in this environment results in several important differences from the equilibrium that arose in KPT and the related papers. In particular, the optimal police strategy is effective in the sense that the guilt rate among the investigated is higher than in the population at large, and unbiased police officers treat observationally equivalent individuals from different races differently only if it is a strictly dominant strategy to do so. This will be shown to have important consequences with respect to whether the tests used by Knowles, Persico, and Todd [2001] and Knowles and Hernandez-Murillo [2003] are valid for determining whether what racially unequal investigation rates in a particular jurisdiction are due to racial bias or statistical discrimination.

Moreover, I will also use the model proposed below to characterize how implementing a colorblind policy will affect the overall crime rate. In this way, this paper is similar to Farmer and Terrell [2001] and Persico [2002]. However, this model differs from these papers not only in its flexibility,[5] but also in its ability to describe how a variety of different jurisdictional characteristics may affect how the crime rate responds to a colorblind policing policy.[6]

3 Model of Racially Unequal Search Rates with Unbiased Officers

In developing this model, I first describe individual behavior, then police behavior, and finally I characterize the equilibrium.

3.1 Individual Behavior

In the relevant population/jurisdiction, let there be a continuum of individuals, where individuals can be divided into two racial groups, a and b, with a fraction β of the population being from race b.[7] Within each race, assume there are two types of individuals,

[5]For example, this model allows law enforcement officers to observe endogenously determined characteristics related to guilt (something not allowed in Persico [2002]), allows for deterrence (something not allowed for in Farmer and Terrell [2001]), and allows for the possibility that police cannot observe, and therefore cannot arrest, all individuals (something not allowed for in either of the previous papers).

[6]In Farmer and Terrell [2001], there are no parameters that can differ across jurisdictions. In Persico [2002], only the relative differences in the wealth distributions across races (implicitly the distributions of benefits to committing the relevant crime) can differ across jurisdictions.

[7]Note that "race" can actually be any directly observable exogenous characteristic, meaning any characteristic where the realization of this characteristic does not depend on whether or not an individual is guilty of committing the relevant crime and can be observed by a law enforcement officer with minimal effort. For example, age groups, gender groups, or race/gender/age group combinations can substitute for racial groups without loss of generality.

those individuals who are "at risk" for choosing to commit a particular relevant crime, and those who are "not at risk" for choosing to commit the relevant crime.[8] Denoting the net utility benefit to person i from committing the crime as ϵ_i, assume $\epsilon_i = 0$ for all *not at risk* individuals. Alternatively, for *at risk* individuals, assume ϵ_i is an i.i.d. random variable drawn from a distribution G defined over the support $(0, \bar{\epsilon}]$ (i.e. $\epsilon_i > 0$ for all *at risk* individuals).

Let the two racial groups differ in that a fraction λ_a of race a are *at risk* individuals, while λ_b of race b are *at risk* individuals, with $\lambda_a < \lambda_b$. There are a variety of reasons why one racial group may have a higher fraction of *at risk* individuals than the other. For example, if the opportunity cost of arrest is prohibitively high for high-income individuals, then $\lambda_a < \lambda_b$ will result if race a has a significantly higher proportion of high-income individuals than race b. Alternatively, $\lambda_a < \lambda_b$ may arise if race b has a higher fraction of young people than race a, and the utility benefit to committing the relevant crime is negligible for everyone after reaching a certain age. Or, there may be other more complicated mechanisms, such as race specific gangs or racial differences in religious beliefs, that account for racial differences in the fraction of each race that are at risk of participating in some particular criminal behavior. Clearly, whether or not it is reasonable to assume $\lambda_a < \lambda_b$ depends on the specific context in question.[9]

Next, assume that the utility cost to getting arrested for the criminal activity is equal to c for all individuals, and that each individual knows that if he is guilty of committing the crime he will be arrested if and only if a law enforcement officer chooses to investigate him.[10] However, each individual also knows that law enforcement officers can only investigate an individuals who they come in contact with, which is a fraction

[8] The particular relevant crime will depend on the jurisdiction in question. For example, if the relevant jurisdiction is the highway system, then the relevant crime is carrying contraband such as drugs, alcohol or weapons. Alternatively, if the relevant jurisdiction is a store, then the relevant crime is possessing shoplifted material, or if the relevant jurisdiction is border crossings, the relevant crime may be supporting or planning terrorist activity.

[9] Note that Persico's [2002] assumption that the benefits to crime are directly related to income, and the distribution of income differs across races, is just a special case of the assumption used here as discussed in the example given above.

[10] In the context of the highway patrol, an "investigation" is a motor vehicle or person search. In the broader sense however, an investigation could refer to everything from a security officer following a customer around a store, to an FBI or INS investigation of an individual for terrorist connections.

η of the total population.[11] Moreover, an individual of race j believes that if he is guilty of committing the crime and is observed by a law enforcement officer he will be investigated with probability $\widehat{p}_{g,j}$.[12]

Given this setup, an individual from race j will choose to commit the crime as long as his benefit from doing so exceeds the expected cost of doing so. More explicitly, an individual i will commit the crime when $\epsilon_i > \eta\widehat{p}_{g,j}c$. Therefore, the proportion of *at risk* individuals from race j who choose to commit the relevant crime equals $1 - G(\eta\widehat{p}_{g,j}c)$. This implies that the overall proportion of race j that is guilty of committing the crime is given by

$$\pi(\lambda_j, \widehat{p}_{g,j}) = \lambda_j[1 - G(\eta\widehat{p}_{g,j}c)] \tag{1}$$

Not surprisingly, this expression for $\pi(\lambda_j, \widehat{p}_{g,j})$ implies that the proportion of a race that is guilty of committing the relevant crime increases with the fraction of the race that are *at risk* (i.e. λ_j), and decreases with the believed probability of search given an individual is observed by police and is guilty of the relevant crime (i.e. $\widehat{p}_{g,j}$).

Finally, for the sake of completeness, assume $G(0) = 0$ and $G(c) > 0$. This simply assumes that all *at risk* individuals will commit the crime if they believe there is zero probability of being caught, and the utility cost of arrest is high enough such that at least some (and possibly all) of the *at risk* individuals will choose not to commit the relevant crime if they think that they will be caught with probability one.

3.2 Law Enforcement Officer Behavior

Recall that law enforcement officers can observe only a fraction η of all individuals in their jurisdiction. For each individual they do observe, they must decide whether or not to investigate the individual. If they choose to investigate an individual, they will arrest a guilty individual with probability one. The cost to investigating an individual is equal to q, while the benefit to arresting a guilty individual is denoted v, where $v > q$.

[11] For example, in the context of highway patrol, η is the fraction of all cars on the highway that one officer can observe. Similarly, in a department store, it is the fraction of all customers that a security officer can observe.

[12] As will be discussed in section 3.3, in equilibrium, this belief must correspond to the true probability a guilty individual from race j will be searched given he is observed by police.

If we assume that law enforcement officers are simply agents of society, then v can be interpreted as the benefit to society for apprehending an individual who is guilty of the relevant crime and q can be interpreted as the societal cost to investigating an individual.

When observing an individual, law enforcement officers can observe the individual's race and as well as a guilt signal denoted $\theta \in [0,1]$, where this signal is some characteristic of the individual that may be influenced by whether the individual is guilty (e.g. nervousness, clothing style/colors, driving patterns, unusual odors, etc.). Similar to Coate and Loury [1993], let θ be a random variable realized after the individual has decided whether or not to commit the crime, where the distribution of θ depends on whether or not the individual is guilty. Let $F_g(\theta)$ be the probability the signal does not exceed θ given that the individual is guilty of committing the crime, and $F_n(\theta)$ be the probability the signal does not exceed θ given that the individual is not guilty of committing the crime. Denoting the corresponding density functions as $f_g(\theta)$ and $f_n(\theta)$, the likelihood ratio at θ will equal $\frac{f_n(\theta)}{f_g(\theta)}$. Assume that this likelihood ratio is decreasing on [0,1], implying $F_g(\theta) \leq F_n(\theta)$ for all possible θ. Hence, higher values of the signal are more likely for guilty individuals.

As stated above, law enforcement officers can also distinguish the race of each individual they observe. This is important as law enforcement officers may have different beliefs concerning the proportion of each race that is guilty. Specifically, let these beliefs concerning race j be denoted $\widehat{\pi}_j$.[13]

Using Bayes' Rule, a law enforcement officer's posterior beliefs concerning the probability that an observed individual from race j who emits a signal θ is guilty will equal

$$\xi(\widehat{\pi}_j, \theta) = \frac{\widehat{\pi}_j f_g(\theta)}{\widehat{\pi}_j f_g(\theta) + (1-\widehat{\pi}_j) f_n(\theta)}.$$

Re-writing the above equation in a more convenient form we get,

$$\xi(\widehat{\pi}_j, \theta) = \frac{\widehat{\pi}_j}{\widehat{\pi}_j + (1-\widehat{\pi}_j)\frac{f_n(\theta)}{f_g(\theta)}}. \tag{2}$$

Because the likelihood ratio ($\frac{f_n(\theta)}{f_g(\theta)}$) was assumed to be decreasing in θ, taking derivatives of equation (2) will show that $\frac{\partial \xi(\widehat{\pi}_j,\theta)}{\partial \theta} > 0$ and $\frac{\partial \xi(\widehat{\pi}_j,\theta)}{\partial \widehat{\pi}_j} > 0$. In other words, the

[13] Once again, as will be discussed in section 3.3., these beliefs will correspond to the true proportion of each racial group that is carrying contraband in equilibrium.

posterior belief of guilt is increasing in the signal, and for any given signal, the posterior belief of guilt is greater the greater the prior belief of guilt. Moreover, $\xi(1, \theta) = 1$ and $\xi(0, \theta) = 0$ for all θ, meaning if officers believe all members of race j are guilty of the relevant crime, or believe no members of race j are guilty of the relevant crime, then the observed guilt signal does not influence the officers' posterior beliefs.

Given this setup, a law enforcement officer will investigate an observed individual of race j if and only if the individual emits a signal such that $\xi(\widehat{\pi}_j, \theta)v > q$, or only if

$$\frac{\widehat{\pi}_j}{\widehat{\pi}_j + (1 - \widehat{\pi}_j)\frac{f_n(\theta)}{f_g(\theta)}} \geq \frac{q}{v}. \tag{3}$$

We can now define a threshold level $\theta^*(\widehat{\pi}_j)$, such that for any belief $\widehat{\pi}_j$, equation (3) holds if and only if $\theta > \theta^*(\widehat{\pi}_j)$, where $\theta^*(\widehat{\pi}_j) = 0$ if $\widehat{\pi}_j$ is such that equation (3) holds for all θ, $\theta^*(\widehat{\pi}_j) = 1$ if $\widehat{\pi}_j$ is such that equation (3) never holds for any θ, and $\theta^*(\widehat{\pi}_j)$ equals the value of θ that causes (3) to hold at equality if possible.

Because $\frac{f_n(\theta)}{f_g(\theta)}$ was assumed to be non-increasing in θ, it can be easily confirmed that when $\theta^*(\widehat{\pi}_j)$ is between 0 and 1, it is strictly decreasing in $\widehat{\pi}_j$, or the greater the prior belief of guilt, the lower the signal necessary for investigation.

Given this search strategy, the probability a guilty individual from race j is arrested equals the probability he is observed, η, times the probability he will be investigated if observed, which equals

$$p_g(\widehat{\pi}_j) = [1 - F_g(\theta^*(\widehat{\pi}_j))]. \tag{4}$$

Finally, note that we can easily incorporate racial bias among law enforcement officers against race b into this model by allowing v (the benefit to a successful search) to be greater for race b than race a (or by allowing the cost to searching q to be lower for race a than race b). Equation (3) shows that the effect of such a bias on police behavior would be to further lower $\theta^*(\widehat{\pi}_b)$ relative to $\theta^*(\widehat{\pi}_a)$, which in turn would increase $p_g(\widehat{\pi}_b)$ relative to $p_g(\widehat{\pi}_a)$, all else equal.

3.3 Equilibrium

Equilibrium occurs when each individuals' beliefs concerning the probability of being investigated given he is guilty and observed by a law enforcement officer corresponds to the true probability of being investigated given he is guilty and observed by an officer, and when law enforcement officers' beliefs concerning the proportion of each race that are guilty of committing the relevant crime correspond to the truth. In other words, equilibrium is a pair of beliefs $\{\widehat{p}_{g,j}, \widehat{\pi}_j\}$ for each race j that simultaneously satisfy $\widehat{\pi}_j = \pi(\lambda_j, \widehat{p}_{g,j})$ and $\widehat{p}_{g,j} = p_g(\widehat{\pi}_j)$, where $\pi(\lambda_j, \widehat{p}_{g,j})$ is defined in equation (1) and $p_g(\widehat{\pi}_j)$ is defined in equation (4). The Appendix proves that for any λ_j, there exists a unique equilibrium pair of beliefs $\{\widehat{p}_{g,j}, \widehat{\pi}_j\}$. Denote this equilibrium pair as $\{p^e_{g,j}, \pi^e_j\}$. In the discussion to follow, unless otherwise stated, the equilibrium is assumed to be one in which officers are racially unbiased (i.e. v and q are the same for both races).

The first thing to note about this equilibrium is that if $\lambda_b > \lambda_a$, then $\pi^e_b > \pi^e_a$.[14] In other words, if one race has as higher proportion of *at risk* individuals than the other, then in equilibrium, the more *at risk* racial group will also have a higher proportion of actual law-breakers than the other racial group. While this result is quite intuitive, it is worth noting that this result is in direct contrast to the statistical discrimination equilibrium in KPT and the related papers. As discussed in Section 2, the key aspect of their equilibrium is that even if one race has a higher average benefit to carrying contraband than the other (which is analogous to a higher fraction of *at risk* individuals), in equilibrium, both races will still choose to carry contraband at the same rate, which in turn causes police to be indifferent between searching any two individuals of different races. In the equilibrium of the model presented here, police are only indifferent between investigating (or searching in the context of highway patrol officers) the marginal members of each race, where the marginal member of race j is defined to be an individual who emits a guilt signal equal to the race specific threshold level $\theta^*(\pi^e_j)$.

If we restrict our attention to parameterizations where $\theta^*(\pi^e_b) > 0$ and/or where $\theta^*(\pi^e_a) < 1$,[15] then because $\lambda_b > \lambda_a$ implies $\pi^e_b > \pi^e_a$, we also know that $\theta^*(\pi^e_b) < \theta^*(\pi^e_a)$. In other words, if race b has a higher proportion of *at risk* individuals than race a, then not only will race b have a higher equilibrium guilt rate than race a, but also unbiased officers will set a lower signal threshold necessary for investigating members of race b than race a. This is how statistical discrimination manifests itself in this

[14] See Appendix for proof of this assertion.

[15] In other words, assume that π^e_b is low enough such that police do not investigate all individuals from group b that they observe and/or π^e_w is high enough such that police have at least a positive probability of investigating an individual from group w.

model. In particular, unequal equilibrium crime rates mean it is strictly optimal for unbiased officers to treat some observationally similar individuals from different races differently (i.e. individuals who emit a guilt signal between $\theta^*(\pi_b^e)$ and $\theta^*(\pi_w^e)$ will only be investigated if they are from group b).

This result is important because it implies that for any particular officer, even if this officer is not racially biased, maintaining statistical discriminatory behavior is strictly preferred to any other strategy. For example, if a particular officer chose not to behave in a discriminatory manner, and raised the signal threshold for members of race b (and/or lowered the signal threshold for members of race a), he would find that he was not investigating individuals over which he had a positive expected value of investigating (and/or would be investigating individuals over which he had a negative expected value of investigating). Hence, even if an officer gets a slight disutility from engaging in statistical discrimination, according to this model he would still engage in such behavior since it is strictly preferred to any other strategy. This is not true for the statistical discrimination occurring in KPT and the related papers, as racially unequal search rates are only weakly dominant strategies for unbiased officers in those papers.

Moreover, from equation (4) we can see that since $\theta^*(\pi_a^e) > \theta^*(\pi_b^e)$, we know that $p_{g,a}^e < p_{g,b}^e$. In words, the lower threshold guilt signal police use for group b causes the probability that officers investigate a guilty member of race b to be greater than the probability officers investigate a guilty member of race a. Furthermore, the probability a not guilty member of race j is investigated in equilibrium will be $\eta p_{n,j}^e$, where

$$p_{n,j}^e = 1 - F_n(\theta^*(\pi_j^e)) \tag{5}$$

Since $\theta^*(\pi_b^e) < \theta^*(\pi_a^e)$, the above expression implies that $\eta p_{n,b}^e > \eta p_{n,a}^e$, or that not guilty members of race b will be more likely to be subjected to an investigation than not guilty members of race a, even if officers are racially unbiased.

The next thing to note is that we can write out the overall equilibrium crime rate for the particular crime relevant to the jurisdiction in question as

$$\Pi^e = \beta\lambda_b[1 - G(\eta p_b^e c)] + (1-\beta)\lambda_a[1 - G(\eta p_a^e c)]. \tag{6}$$

Not surprisingly, the above expression shows that the overall crime rate is increasing in the proportion of *at risk* individuals from each race (i.e. λ_a and λ_b). Moreover, because

$\pi_b^e > \pi_a^e$ we know via the equilibrium requirements that $1 - G(\eta p_b^e c) > 1 - G(\eta p_a^e c)$. Combining this with the assumption that $\lambda_b > \lambda_a$, equation (6) implies that the overall crime rate is increasing in the proportion of the population from race b (i.e. β). Also, note that the overall crime rate is decreasing in the probability the individual will be observed by a law enforcement officer (i.e. η) and in the cost of being arrested (i.e. c). The intuition for all of these results is quite straightforward.

Furthermore, let $\rho(\pi_j^e)$ denote the equilibrium investigation rate for race j, meaning

$$\rho(\pi_j^e) = \eta[1 - F_g(\theta^*(\pi_j^e))]\pi_j^e + \eta[1 - F_n(\theta^*(\pi_j^e))](1 - \pi_j^e),$$

where the first term is the fraction of guilty race j individuals who are observed by police and emit a guilt signal high enough to warrant a search, and the second term is the fraction of not-guilty race j individuals are observed by police and emit a guilt signal high enough to warrant a search. Taking the derivative of the above expression gives

$$\begin{aligned}\frac{\partial\rho(\pi_j^e)}{\partial\pi_j^e} = \eta(-f_g(\theta^*(\pi_j^e))\frac{\partial\theta^*(\pi_j^e)}{\partial\pi}\pi_j^e + [1 - F_g(\theta^*(\pi_j^e))]\\ -f_n(\theta^*(\pi_j^e))\frac{\partial\theta^*(\pi_{kj}^e)}{\partial\pi}(1 - \pi_j^e) - [1 - F_n(\theta^*(\pi_j^e))]).\end{aligned}$$

Since $\frac{\partial\theta^*(\pi_j^e)}{\partial\pi_j^e} \leq 0$ and $[1 - F_g(\theta^*(\pi_j^e))] > [1 - F_n(\theta^*(\pi_j^e))]$, the above derivative will be strictly positive. In other words, if police are not racially biased, then this model implies that race b can be investigated at a higher rate than race a *if and only if* race b has a higher equilibrium guilt rate than race a (i.e. only if $\pi_b^e > \pi_a^e$). In principle, this implication means that we could "test" whether racially unequal investigation rates in a particular jurisdiction are due to racially biased officers or statistical discrimination, similar to what was done in Knowles, Persico, and Todd [2001]. Specifically, if there existed data on the overall guilt rates for each racial group in a jurisdiction, statistical discrimination could be a contributing factor for racially unequal investigation rates between races a and b only if race b has a higher overall guilt rate than race a. Alternatively, if race b is investigated at a higher rate than race a, but the overall guilt rate among race a is greater than or equal to the overall guilt rate among race b, it must be the case that v/q is smaller for group a than group b, implying officers are racially biased against group b.[16]

[16]To see why this is true, recall from equation (3) that a lower v/q means a higher $\theta^*(\widehat{\pi}_j)$, for any

However, the primary constraint inherent in this "test" is that accurate data regarding the underlying guilt rates for each race as a whole in a jurisdiction would be difficult, if not impossible to obtain. Rather, most available data can only reveal the equilibrium guilt rate among the investigated members of each race (since these are the only ones for whom police can identify whether they are guilty or not). This issue was not raised in KPT, since in their equilibrium police chose who to investigate at random, meaning the overall guilt rate for each race equals the guilt rate among the investigated for each race. However, this model reveals the importance of this issue. Specifically, the equilibrium guilt rate among the investigated can be denoted $\gamma(\pi_j^e)$ and will equal

$$\gamma(\pi_j^e) = \frac{1}{1 + \frac{(1-\pi_j^e)[1-F_n(\theta^*(\pi_j^e))]}{\pi_j^e[1-F_g(\theta^*(\pi_j^e))]}}. \tag{7}$$

It is straightforward to show that because $F_n(\theta) > F_g(\theta)$ for all θ, $\gamma(\pi_j^e)$ as defined in the above expression exceeds π_j^e. In other words, because police selectively choose which observed individuals to investigate on the basis of their guilt signals, the guilt rate among the investigated from each race will be greater than the guilt rate among that race as a whole.

Furthermore, equation (7) can be used to show that the guilt rate among the investigated from each race will generally not be sufficient for determining whether or not racially unequal investigation rates in a jurisdiction can be due to statistical rather than racial bias among law enforcement officers. Specifically, even with with $\pi_b^e > \pi_a^e$ (i.e. the necessary condition for unequal search rates to be due to statistical discrimination) $\gamma(\pi_b^e)$ can be greater than, equal to, or less than $\gamma(\pi_a^e)$. This result implies that Knowles, Persico, and Todd's [2001] "test" for determining whether unequal investigation rates are due to police officer racial bias or statistical discrimination is not necessarily valid. As discussed in the previously, they argue that if guilt rates among the investigated are not equal across races, then police officers must be biased against the race with the lower guilt rate. To put it another way, they argue that racially unequal investigation rates are consistent with unbiased police officers as long as the guilt rates among the investigated are equal across races. However, equation (7) reveals that guilt rates among

given beliefs $\widehat{\pi}_j$. Therefore, if v/q is lower for group a than group b, then even if $\pi_a^e \geq \pi_b^e$, it can still be the case that $\theta^*(\pi_a^e) > \theta^*(\pi_b^e)$, which makes it possible for $\rho(\pi_a^e) < \rho(\pi_b^e)$, or race b to be investigated at a higher rate even though they have a lower crime rate.

the investigated are not necessarily equalized across races, even with racially unbiased police officers.

Moreover, as will be shown below, rather stringent conditions must be met (although even these conditions are not sufficient) for racially unequal investigation rates to co-exist with racially equal guilt rates among the investigated, if law enforcement officers are truly racially unbiased. Specifically, from equation (7), the guilt rate among the investigated individuals from race b (i.e. $\gamma(\pi_b^e)$) will be equal to or less than $\gamma(\pi_a^e)$ if and only if

$$\frac{(1-\pi_b^e)[1-F_n(\theta^*(\pi_b^e))]}{\pi_b^e[1-F_g(\theta^*(\pi_b^e))]} \geq \frac{(1-\pi_a^e)[1-F_n(\theta^*(\pi_a^e))]}{\pi_b^e[1-F_g(\theta^*(\pi_a^e))]}.$$

Since $\pi_b^e > \pi_a^e$, then $\frac{1-\pi_b^e}{\pi_b^e} < \frac{1-\pi_a^e}{\pi_a^e}$, meaning a necessary condition for the above expression to be true is that

$$\frac{[1-F_n(\theta^*(\pi_b^e))]}{[1-F_g(\theta^*(\pi_b^e))]} > \frac{[1-F_n(\theta^*(\pi_a^e))]}{[1-F_g(\theta^*(\pi_a^e))]}. \tag{8}$$

Now, define $\tilde{\theta}$ to be the value of θ such that $f_g(\tilde{\theta}) = f_n(\tilde{\theta})$.[17] Given this definition and the fact that $\frac{f_n}{f_q}$ is decreasing on $(0,1)$, it will also be true that $\frac{[1-F_n(\theta)]}{[1-F_g(\theta)]}$ reaches a minimum at $\theta = \tilde{\theta}$, is decreasing in θ for $\theta < \tilde{\theta}$, and is increasing in θ for $\theta > \tilde{\theta}$. Therefore, since $\pi_a^e < \pi_b^e$, we know $\theta^*(\pi_a^e) > \theta^*(\pi_b^e)$, meaning equation (8) can only hold if $\theta^*(\pi_b^e) < \tilde{\theta}$. In other words, the guilt rate among investigated members of race a (i.e. $\gamma(\pi_a^e)$) can be greater than or equal to the guilt rate among the investigated members of race b (i.e. $\gamma(\pi_b^e)$) only if it is optimal for law enforcement officers to set a relatively low signal threshold individuals from race b must surpass in order to be investigated.

A relatively low signal threshold for race b not only implies that a relatively high proportion of race b individuals would be investigated, but also, recalling how $\theta^*(\pi)$ was defined in equation (3), $\theta^*(\pi_b^e)$ can be relatively low only if π_b^e is relatively large and/or q is small compared to v. Hence, for it to even be possible for the guilt rate for investigated members of race a to be equal to or greater than the guilt rate for investigated members of race b, a relatively high fraction of race b must be investigated, the actual guilt rate among race b must be relatively high, and/or the benefit officers

[17] We know such a $\tilde{\theta}$ exists and is unique since $\frac{f_n}{f_q}$ was assumed to be decreasing on $(0,1)$ and that both f_n and f_g are pdfs.

incur from a successful investigation must be relatively high in comparison to the cost of investigating an individual.

While it is difficult to determine the size of the benefit officers incur from a successful investigation relative to the cost of an investigation, the available data suggests the other two conditions are relatively unlikely to hold. For example, in Missouri, only 11.5 percent of non-white motorists who were stopped were actually searched (the relevant type of investigation in the context of motorists on highways) [Hernandez-Murillo and Knowles, 2003].[18] Similarly, only about 9.5 percent and 22 percent of non-white motorists who were stopped were searched by Rhode Island State Troopers [Farrell et al., 2003] and the Los Angeles Police Department [Los Angeles Police Department, 2002] respectively.[19] Also, the guilt rates among searched non-white motorists in Missouri, Rhode Island, and Los Angeles, were 15.6 percent, 13.9 percent, and 29.4 percent respectively.[20] Recalling that guilt rates among the investigated are likely to be higher than the guilt rate for the race as a whole, these findings suggest that guilt rates among non-white motorists are relatively low. Therefore, it is unlikely that in these localities where guilt rates among searched minority and white motorists are equal, the unequal search rates across races were due soley to statistical discrimination.

4 Analysis of the Theoretical Costs to Banning Profiling

As discussed in the introduction, while many people may ethically object to police using race as a factor in deciding who to search regardless of the underlying motivation, eradicating such behavior may have costs in terms of increasing overall crime rates. This section uses the model developed in Section 3 to analyze how these costs to an anti-profiling policy may differ by characteristics of the relevant jurisdiction and crime. Generally, the analysis presented below assumes officers to be racially unbiased. However, recall that in this model, if officers are racially biased, then v is bigger for group b than group a (or q is smaller for group a than group b). As discussed above, the only

[18]This compares to only 6.4 percent of white drivers who were stopped were searched.

[19]The analogous search rates for white drivers were 4.3 percent in Rhode Island and 6.6 percent in Los Angeles.

[20]The analogous guilt rates for white drivers who were searched were 23.7 percent in Missouri, 14.8 percent in Rhode Island, and 28.7 percent in Los Angeles.

thing that changes due to this bias is that $\theta^*(\pi_b^e)$ will be lower for any π_b^e or $\theta^*(\pi_a^e)$ will be higher for any π_a^e. As can be confirmed below, the basic implications that follow will not be changed by allowing officers to have such racial biases.

4.1 Implementing a Perfect Colorblind Policing Policy

Let us assume that it is theoretically possible to implement a perfect anti-profiling policy such that police officers could not use an individual's racial group in any part of their investigation decision. In other words, assume the policy could make police officers perfectly colorblind. With this policy, officers effectively only observe one race, and therefore cannot employ race specific beliefs. Rather, police must use only one belief concerning the average guilt rate among the whole population, $\widehat{\Pi}$. Given this belief, in the same manner as in Section 3, officers will choose a threshold level $\theta^*(\widehat{\Pi})$, such that they will investigate an observed individual only if they observe a guilt signal greater than $\theta^*(\widehat{\Pi})$, where $\theta^*(\widehat{\Pi})$ is derived analogously to before. Therefore, for individuals of both races, the probability of being investigated if guilty equals $\eta p_g(\widehat{\Pi})$, where

$$p_g(\widehat{\Pi}) = 1 - F_g(\theta^*(\widehat{\Pi})). \tag{9}$$

From the individual's perspective, the problem does not change. Specifically, an individual of race j chooses to commit the relevant crime only if $\epsilon \geq \eta \widehat{p}_{g,j} c$. This means the proportion of the total population choosing to commit the crime will equal

$$\Pi(\lambda_b, \lambda_a, \widehat{p}_{g,b}, \widehat{p}_{g,a},) = \beta\lambda_b[1 - G(\eta\widehat{p}_{g,b}c)] + (1-\beta)\lambda_a[1 - G(\eta\widehat{p}_{g,a}c)] \tag{10}$$

Equilibrium now consists of a set of beliefs $\{\widehat{p}_{g,a}, \widehat{p}_{g,b}, \widehat{\Pi}\}$ such that $\widehat{p}_{g,a} = \widehat{p}_{g,b} = p_g(\widehat{\Pi})$ (as described by equation 9) and $\widehat{\Pi} = \Pi(\lambda_b, \lambda_a, \widehat{p}_{g,b}, \widehat{p}_{g,a})$ (as described by equation 10). Using a similar argument as in the unconstrained case, we can prove such an equilibrium exists and is unique.[21] Denote this equilibrium set of beliefs in the colorblind environment as $\{\Pi^c, p_g^c\}$.

The first thing to note about this new equilibrium is that if $\lambda_a < \lambda_b$, causing $\pi_a^e < \pi_b^e$ in the unconstrained equilibrium, then this new equilibrium will be such that $\pi_a^e < \Pi^c <$

[21] See Appendix.

π_b^e.[22] A straightforward implication of this result is that $\theta^*(\pi_a^e) > \theta^*(\Pi^c) > \theta^*(\pi_b^e)$, since $\theta^*(\pi)$ was previously argued to be decreasing in π. This result in turn implies that $p_{g,a}^e < p_g^c < p_{g,b}^e$.[23] In words, the implementation of a policy that causes police to behave in a colorblind manner will increase the probability of search for guilty members from race a, while decreasing the probability of search for guilty members from race b.

4.2 The Effect of a Colorblind Policy on the Overall Crime Rate

By construction, this colorblind policy constrains police from using some available information relevant to maximizing the success rate of their investigations. If society incurs a cost for every individual who carries the relevant contraband, it is important to characterize how constraining police in this manner will affect the overall crime rate. In the context of this model, this question reduces to calculating $\Pi^c - \Pi^e$, where Π^c is described by equation (10) and Π^e is described by equation (6). Writing this expression out and re-arranging, we obtain

$$\Pi^c - \Pi^e = \beta\lambda_b[G(\eta p_{g,b}^e c) - G(\eta p_g^c c)] + (1-\beta)\lambda_a[G(\eta p_{g,a}^e c) - G(\eta p_g^c c)]. \qquad (11)$$

Since $p_{g,a}^e < p_g^c < p_{g,b}^e$, the term in the first set of brackets in the above expression will be positive, while the term in the second set of brackets will be negative. Therefore, the degree to which the policy increases the overall crime rate will depend on the relative magnitude of the first product versus the second product in equation (11). Intuitively, the first product is the degree to which the policy increases the number of group b individuals committing the relevant crime, while the second product is the degree to which the policy decreases the number of group a individuals committing the relevant crime.[24]

[22] Proof of this assertion can be found in Appendix. Note that this will be true even if law enforcement officers are racially biased, as racial bias just further exacerbates the discrimination that exists when officers are unbiased.

[23] This result is straightforward from the fact that $p_{g,a}^e = 1 - F_g(\theta^*(\pi_a^e))$, $p_g^e = 1 - F_g(\theta^*(\Pi^e))$, and $p_{g,b}^e = 1 - F_g(\theta^*(\pi_b^e))$

[24] Note that this condition is somewhat analogous to Persico's [2002] condition regarding when a racially "fair" (i.e. colorblind) policy can lead to less crime overall. Namely, the expression in equation (11) is simply comparing the relative elasticities of each race to more and less intensive policing.

The first thing to note about equation (11) is that it shows it is theoretically possible for the overall crime rate to actually decrease following the implementation of a colorblind policy, since the second product can theoretically be greater in absolute value than the first product. This result emphasizes an important point, namely that the policy that maximizes the number of criminals caught for any given number of investigations performed may be quite different than the policy which minimizes the overall number of guilty individuals.

The second thing to note about equation (11) is that it implies that the increase in the crime rate following the colorblind policy will be smaller the smaller the fraction of the overall population that is from race b (i.e. the smaller the β). Furthermore, equation (11) shows that the increase in the crime rate following the colorblind policy will be smaller the smaller λ_b is compared to λ_a. In other words, the crime cost to a colorblind policy will likely be smaller, the more similar the races are in terms of their proportions of *at risk* individuals. One implication of this is that the crime cost to a colorblind policy is likely to be smaller when the relevant racial groups have relatively similar income, age, and gender distributions, as such similarities would likely cause λ_a to approach λ_b.

The fourth thing revealed by equation (11) is that the increase in the crime rate following the colorblind policy will generally be smaller the smaller is η, or the smaller the fraction of the overall population that police can observe. Specifically, a smaller η means smaller differences between $\eta p^e_{g,b}c$ and $\eta p^c_g c$, and $\eta p^e_{g,a}c$ and $\eta p^c_g c$. Generally, this will lead to a smaller difference between $G(\eta p^c_g c) - G(\eta p^e_{g,b}c)$ and $G(\eta p^c_g c) - G(\eta p^e_{g,a}c)$.[25] Intuitively, if η is small, then law enforcement officers are only observing a small fraction of the overall population, meaning that they are not having much of an effect on individual criminal participation decisions in the first place. Therefore, constraining law enforcement officer behavior a little bit will not have a large effect on individual behavior. For analogous reasons, the increase in the crime rate following the colorblind policy will also generally be smaller the smaller is c, or the smaller the penalty for getting caught.

The final thing to note about equation (11) is that the cost of the colorblind policy

[25] However, this is not necessarily the case, as this result depends on the shape of the distribution of benefits to carrying contraband (i.e. the shape of G).

will depend on the distribution of the benefits to committing the relevant crime (i.e. the G function). Specifically, the crime rate cost to the policy will be smaller the more concave the G function is between $\eta p^e_{g,a}c$ and $\eta p^e_{g,b}c$ (see the lower graph in Figure 1(i) compared to the lower graph in Figure1 (ii)). This statement is equivalent to saying that the crime rate cost to a colorblind policy will be smaller if the elasticity of *at risk* individuals to greater search rates is decreasing in the search rate.

In the context of this model, the distribution of benefits will be relatively concave for crimes and/or jurisdictions where only a small fraction of *at risk* individuals obtain a large benefit from committing the crime.[26] This situation is depicted graphically in example (i) in Figure 1. Examples may include crimes such as driving with possession of small amounts of drugs or alcohol, or in possession of a firearm, where most individuals who would potentially engage in such behavior would only incur a small benefit from doing so, while only a very few individuals would incur really large benefits (e.g. drug addicts, alcoholics, or those who need guns for protection). Hence, the elasticity of potential violaters of these relatively minor crimes to greater search rates is likely to decrease as search rates increase, meaning a colorblind policy targeting these types of crimes will likely cause only small increases (or even a decreases) in these types of crimes, all else equal.

Alternatively, the cumulative distribution of benefits will generally not be concave for crimes and/or jurisdictions where a substantial fraction of the *at risk* individuals obtain a large benefit from committing the crime. This situation is depicted graphically in example (ii) in Figure 1. Possible crimes of this type include transportation of large amounts of drugs or firearms for distribution and supporting terrorist activities, where most individuals who benefit at all from these crimes incur very large benefits. Therefore, the elasticity of individuals at risk of committing these more serious crimes to greater search rates is not likely to decrease as search rates increase, meaning a colorblind policy targeting these types of crimes may cause large increases in these types of crimes, all else equal.

It is also worth noting that the benefits to this colorblind policy can be seen on

[26]This follows due to the fact that if only a few law-breakers obtain a large benefit from the relevant crime, while the rest obtain smaller benefits, the pdf of benefits will generally be downward sloping, thus implying the cdf will be generally concave.

two fronts. First, the probability that an innocent member of race b is investigated falls from $1 - F_n(\theta^*(\pi_b^e))$ to $1 - F_n(\theta^*(\Pi^e))$.[27] Second, the policy decreases the cross race difference in the probability of investigation (given observation by law enforcement officers) from $F_n(\theta^*(\pi_a^e)) - F_n(\theta^*(\pi_b^e))$ to zero for innocent members of each race, and $F_g(\theta^*(\pi_a^e)) - F_g(\theta^*(\pi_b^e))$ to zero for guilty members of each race. The benefit of this greater equality of treatment can certainly be argued to be very large for societies that place a high valuation on racial equity and individual rights.

5 Implications of the Model in the Context of Particular Examples

This section looks at how this model can help provide insights into how the effects of colorblind policing policies on the crime rate can differ across jurisdictions and/or law enforcement institutions. Because none of the parameters in the following examples are calibrated to equal any sort of estimated value, these examples are meant to be merely instructive for showing how the relevant parameters may differ across jurisdictions, and how this will alter the effect of colorblind policies on the crime rate. In other words, the results of this section are simply meant to suggest that, in some relevant types of jurisdictions, the crime rate costs to a colorblind policy can actually be expected to be very small, while in other types of jurisdictions, the magnitude of the crime rate costs may be less clear cut.

5.1 Colorblind Policies and the Highway Patrol

With respect to highway motor vehicle searches, the model suggests the crime rate costs to imposing a colorblind policy may be quite small. To see why, first note that since the racial composition of motorists on highways is likely to generally reflect the racial composition of the population at large, minorities should generally make up a relatively small proportion of the overall relevant population, meaning β should be relatively low. Second, the most relevant crimes for the majority of motor vehicle searches are possession of a small quantity of drugs, alcohol, or a firearm. Since the proportion of individuals

[27] However, the colorblind policy will have the offsetting effect of increasing the probability that an innocent member of race a is investigated from $1 - F_n(\theta^*(\pi_a^e))$ to $1 - F_n(\theta^*(\Pi^e))$.

who are *at risk* of carrying drugs, alcohol, or firearms is likely to be quite similar across races, λ_a will likely be quite close to λ_b (where group b is black or latino motorists). Third, highway troopers only observe a small fraction of drivers on a given highway, meaning η will be quite low in this context. Fourth, for these relatively minor crimes, the penalty c will also be quite small. Finally, as discussed above, for minor crimes such as drug possession, very few *at risk* individuals are likely to incur large benefits from possessing the drugs while driving. Therefore, as argued previously, the benefit distribution G is likely to be relatively concave for these crimes. From equation (11), we can see that all five of these points suggest that colorblind policies implemented on the highway patrol would likely have a relatively small effect on the overall crime rate on highways.

5.2 Colorblind Policies and the Border Patrol

Alternatively, for investigations of incoming foreigners by border guards at airports and highways, the magnitude of the effect of a colorblind policy on the relevant crime rate is less clear. On the one hand, β is likely to be quite small. For example, if the relevant population is all foreign travellers entering the United States, individuals from the middle-east (the race facing discrimination) make up only a small fraction of the relevant population.[28]

On the other hand, the relevant crime in this context may be supporting terrorist activity. While only a minuscule fraction of foreign nationals from any country are likely to be at risk of supporting terrorist activities against United States citizens, this fraction may be substantially larger for middle-easterners than foreigners from other regions (e.g. South Americans, Europeans, Asians), meaning λ_b (where group b consists of middle-easterners) will likely be substantially larger than λ_a (where group a consists of individuals from other parts of the world). Furthermore, since the penalty to supporting terrorist activity is likely to be large, c will also be large. Finally, in the group of foreign entrants (from any foreign country) who would potentially support terrorist activities, most would likely obtain very large benefits from such activity, meaning the benefit

[28]However, this is certainly not necessarily the case. For example, if the relevant jurisdiction is Israel, then the "race" facing discrimination is the arabs which likely make up a large fraction of incoming foreigners at Israeli border crossings, meaning β in this case would be quite large.

distribution in this case would likely not be very concave. These latter points imply that a colorblind policy can potentially have a large increased crime rate cost in the case of investigations of foreign nationals by border guards.

Finally, it is unclear what fraction of all foreign nationals entering the United States are observed by border guards. Hence, it is unclear whether η is relatively large or relatively small. This discussion shows that, unlike the previous example, the parameters in this context may have offsetting effects and unclear magnitudes. Therefore, it is not as straightforward to judge whether a colorblind policy in this context would also likely have only small effects on the relevant crime rate (i.e. the total number of foreign terrorist supporters in the U.S.).

6 Conclusion

Many people view racial profiling by law enforcement officers as a practice that fosters mistrust between the racial groups most affected by such profiling and law enforcement, as well as a fundamental violation of civil rights and ethical standards. For these reasons, many jurisdictions are discussing or implementing policies aimed at eliminating the practice of law enforcement officers using race as a factor in selecting whom to stop, search, or otherwise investigate more intensely.

While the ethical and constitutional benefits of these anti-profiling policies can be argued to be quite large, the magnitudes of the costs to these policies are less clear. If profiling is simply due to racial bias among officers, then the costs of implementing policies that eliminate profiling are likely to be small, as the only ones "hurt" by such policies are the biased officers. However, if profiling is a result of optimal behavior for unbiased officers under imperfect information, then there may be some substantial costs to banning such behavior, as such bans may lead to increases in the number of people committing the relevant crimes.

The model developed in this paper primarily looked at racial profiling of this latter form, where officers investigate one race at a higher rate than another because such behavior is optimal from an efficiency perspective. Analysis of the equilibrium of this model reveals several important points. First, for statistical discrimination to occur with unbiased law enforcement officers, one racial group must have a higher fraction of "at risk" individuals than the other, in the sense that a greater proportion of one group could be convinced to commit the relevant crime if the probability of getting caught were low enough. Such a difference across races in a particular jurisdiction could arise due to a variety of reasons, including racial differences in the income distribution, job opportunities, the age distribution, or political or religious beliefs.

The second point to come out of the model is that for unequal investigation rates to arise from racially unbiased officers, the overall guilt rate from the more frequently investigated group must be higher than the overall guilt rate from the less frequently investigated group. However, the guilt rate among the investigated members of the more frequently investigated group can be greater than, equal to, or less than the guilt rate among the investigated members of the less frequently investigated group. This result is important because it emphasizes that while there often exists data regarding guilt rates among those investigated in particular jurisdictions, such data will generally not be sufficient for identifying whether unequal investigation rates are due to statistical discrimination or racial bias. In order to reject the hypothesis that racially unequal investigation rates are due only to statistical discrimination rather than officer bias, it is necessary to have data regarding the overall guilt rates for each race, data that would be extremely difficult to collect.

The final results coming from the model discuss the extent to which a policy that eliminates racial profiling will increase (or possibly even decrease) the overall crime rate in a particular jurisdiction, given the characteristics the jurisdiction. In general, a colorblind policing policy should lead to smaller increases in the overall crime rate when the jurisdiction is relatively undiverse, when the racial groups have relatively similar socio-economic and demographic characteristics, when police can only observe a small fraction of the overall relevant population, when the penalty for being caught for the relevant crime is small, and when only a relatively small fraction of each group can be expected to incur very large benefits from committing the relevant crime.

Applying these results to real world examples of racial profiling suggests that colorblind policing policies will likely have only a small effect on the overall rate at which small amounts of drugs, guns, and alcohol are carried by motorists on highways. Given the cost to society for each instance where a small amount of drugs, guns, or alcohol is carried on a highway is likely quite small (although not necessarily negligible), the overall societal cost to implementing a colorblind policy on highway patrolmen can be argued to be quite low. Alternatively, it is not as clear how large the rate of terrorist supporters among foreigners entering the United States would change in response to a colorblind policy imposed on border guards. However, given the potentially tremendous societal cost associated with terrorism, the societal cost related to any increase in the relevant crime rate associated with implementing a colorblind policy on border guards is likely larger than with respect to the highway patrol.

It is worth emphasizing, however, that these results are only suggestive, and that more precise evidence on actual parameter values would be necessary to make any definitive statements concerning the potential increase in the crime rate associated with implementing a colorblind policing policy in any particular jurisdiction. Moreover, further analysis of the magnitudes of the benefits to implementing colorblind policies, and how such benefits may differ across jurisdictions, is necessary before any statements can be made concerning whether the societal benefits associated with eradicating racial profiling in particular jurisdictions exceed to societal costs. This study is simply meant to help describe why profiling may persist among law enforcement agencies, and how the costs to eradicating such behavior will depend on the particular characteristics of a jurisdiction.

Appendix

(a) *Proof of existence and uniqueness of equilibrium in unconstrained environment*

Recall that equilibrium is defined as a pair of beliefs $\{p_j^e, \pi_j^e\}$ such that

$$\pi_j^e = \lambda_j[1 - G(\eta p_j^e c)] \tag{12}$$

and

$$p_j^e = 1 - F_g(\theta^*(\pi_j^e)). \tag{13}$$

Hence, the equilibrium value for the proportion of race j carrying contraband, π_j^e, can be found by substituting equation (13) into equation (12) and solving for π_j^e. Doing this we get

$$\pi_j^e = \lambda_j[1 - G(\eta[1 - F_g(\theta^*(\pi_j^e))]c)] \tag{14}$$

Existence of such a π_j^e can be confirmed by first noting that equation (14) is continuous mapping of a non-empty, convex, compact set (i.e. [0,1]) onto itself. Hence, Brouwer's fixed point theorem can be directly applied to confirm existence of a solution to equation (14).

Moreover, this equilibrium value is unique. To see why this is true, first note that the derivative of the left-hand side of equation (14) with respect to π_j^e is strictly positive. Taking the derivative of the right-hand side of equation (14) with respect to π_j^e we get

$$\lambda_j g(\eta[1 - F_g(\theta^*(\pi_j^e))]c)\eta f_g(\theta^*(\pi_j^e))\frac{\partial\theta^*(\pi_j^e)}{\partial\pi_j^e}c \tag{15}$$

Since $\frac{\partial\theta^*(\pi_j^e)}{\partial\pi_j^e} \leq 0$, the above expression is weakly negative, meaning the right-hand side of equation (14) is weakly decreasing in π_j^e. Therefore, there can be only one value for π_j^e that equates the right-hand side with the left-hand side of equation (14).

Finally, equation (13) shows that p_j^e can then be directly calculated using the unique π_j^e.

(b) *Proof that if $\lambda_b > \lambda_a$, then $\pi_b^e > \pi_a^e$*

To prove this assertion, it is sufficient to prove that π_j^e is strictly increasing in λ_j. Taking the derivative of the right-hand side of equation (14) with respect to λ_j gives

$$[1 - G(\eta[1 - F_g(\theta^*(\pi_j^e))]c))]$$

Since the above expression is strictly positive, we know the right-hand side of equation (14) is increasing in λ_j. Therefore, as λ_j increases, in order maintain equality in equation (14), π_j^e must also increase, as by doing so the left-hand side of equation (14) will increase and right-hand side of equation (14) will decrease (by equation (15)).

(c) *Proof of existence and uniqueness of equilibrium in colorblind environment*

Note that equilibrium can be defined by a pair of beliefs $\{p^c, \Pi^c\}$ such that

$$\Pi^c = \beta\lambda_b[1 - G(\eta p^c c)] + (1 - \beta)\lambda_a[1 - G(\eta p^c c)] \tag{16}$$

and

$$p_b^c = p_a^c = p^c = 1 - F_g(\theta^*(\Pi^c)). \tag{17}$$

Now, note that equation (16) can be re-written as

$$\Pi^c = \lambda_c[1 - G(\eta p^c c)], \tag{18}$$

where $\lambda_c = [\beta\lambda_b + (1 - \beta)\lambda_a]$. Since these equations are of the same form as equations (13) and (12), the same argument as above can be used to prove existence and uniqueness of Π^c and p^c.

(d) *Proof that if $\lambda_b > \lambda_a$, then $\pi_b^e > \Pi^c > \pi_a^e$*

As above, define $\lambda_c = [\beta\lambda_b + (1-\beta)\lambda_a]$. Given this definition, then we know $\lambda_b > \lambda_c > \lambda_a$. From the proof in part (b) above, we then know that π_j^e is increasing in λ_j, proving this assertion.

References

[1] Bush, George W., Memorandum for the Attorney General on Racial Profiling, February 2001.

[2] Aiger, Dennis J. and Glen G. Cain, "Statistical Theories of Discrimination in the Labor Market," *Industrial and Labor Relations Review*, Jan. 1977, pp. 3-33.

[3] Arrow, Kenneth J., "The Theory of Discrimination" in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press, 1973, pp. 3-33.

[4] Coate, Stephen and Glenn Loury, "Will Affirmative Action Eliminate Negative Stereotypes?," *American Economic Review*, December 1993, *83*, 1220-1240.

[5] Cornell, Bradford and Ivo Welch, "Culture, Information, and Screening Discrimination," *Journal of Political Economy*, vol. 104, no. 3, 1996, pp. 542-571.

[6] Farmer, Amy and Dek Terrell, "Crime Versus Justice: Is There a Trade-off?", *Journal of Law and Economics*, October 2001.

[7] Farrell, Amy, Jack McDevitt, Shea Cronin, Erica Pierce, "Rhode Island Traffic Stop Statistics Act Final Report," Northeastern University Institute on Race and Justice, June 30, 2003.

[8] Hernandez-Murillo, Ruben, and John Knowles, "Racial Profiling or Rascist Policing?: Testing in Aggregate Data," Working Paper, University of Pennsylvania, April 2003.

[9] Knowles, John, Nicola Persico, and Petra Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, vol. 109(1), 2001.

[10] Lundberg, Shelly and Richard Startz, "Private Discrimination and Social Intervention in Competitive Markets," *American Economic Review*, Vol. 73-3, Jan. 1983, pgs. 340-347.

[11] Los Angeles Police Department, Citywide Driver Data, "Posting of Motor Vehicle and Pedestrian Stop Data,"

http://lapdonline.org/pdf_files/boi/consent_decree_fdr_02_07_11.pdf, data covering July 1, 2002 - November 30, 2002.

[12] Moro, Andrea, and Peter Norman, "Affirmative Action in a Competitive Economy," *Journal of Public Economics*, vol. 87(3), March 2003.

[13] Persico, Nicola, "Racial Profiling, Fairness, and Effectiveness of Policing," *The American Economic Review*, vol. 92(5), December 2002.

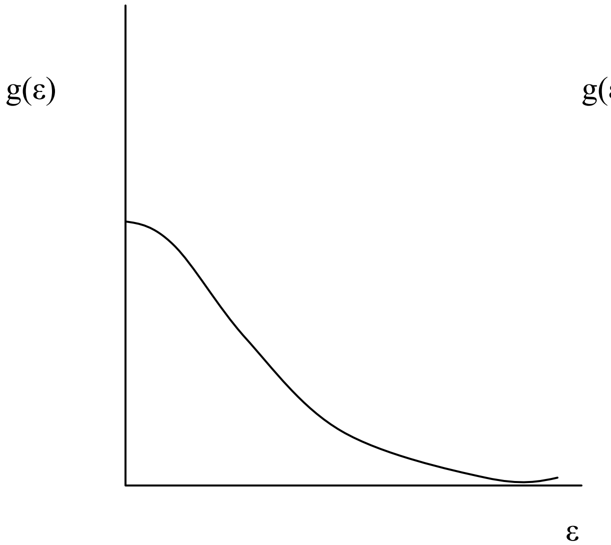
[14] Phelps, Edmund S., "The Statistical Theory of Racism and Sexism," *American Economic Review*, September 1972, 62, 659-61.

[15] Racial Profiling Data Collection Resource Center, "Legislation and Litigation," http://www.racialprofilinganalysis.neu.edu/legislation.php, accessed October 30, 2003.

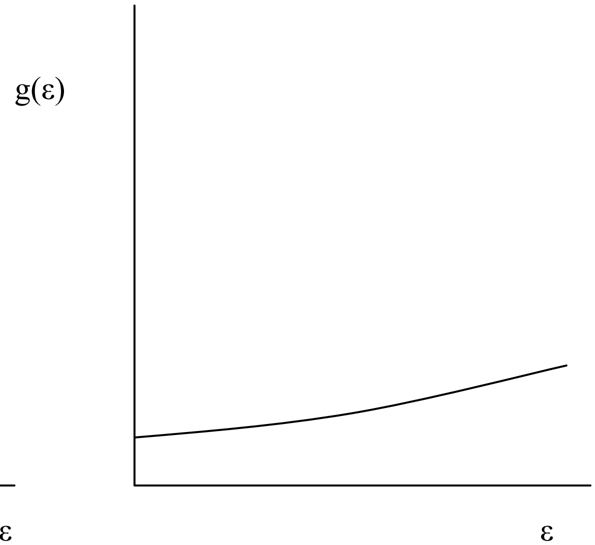
[16] Ramirez, Deborah, Jack McDevitt, and Amy Ferrell, "A Resource Guide on Racial Profiling Data Collection Systems: Promising Practices and Lessons Learned," Monograph, U.S. Department of Justice, November 2000.

[17] Verniero, Peter, and Paul H. Zoubek, "Interim Report of the State Police Review Team Regarding Allegations of Racial Profiling," New Jersey Office of the Attorney General, http://www.state.nj.us/lps/intm_419.pdf, April 20, 1999.

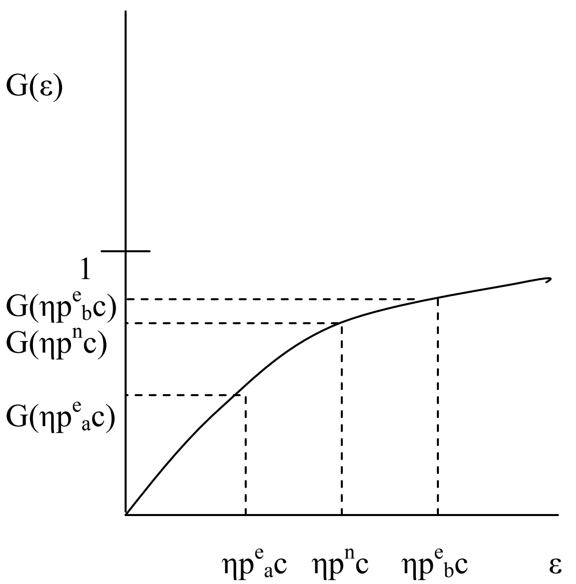
FIGURE 1



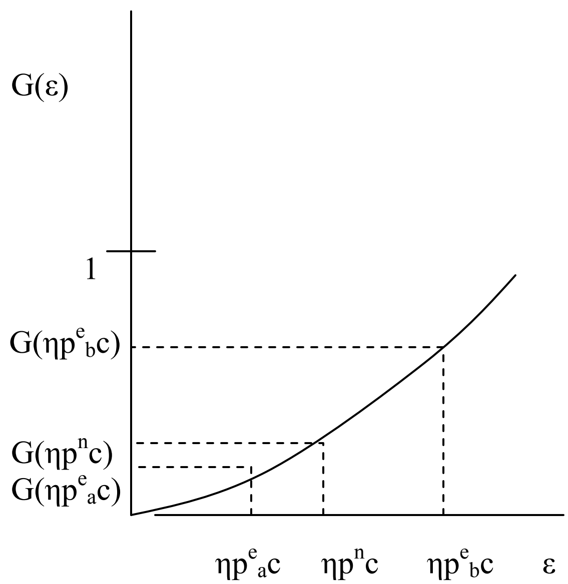
(*i*) pdf of benefits to carrying contraband, for contraband with generally small benefits



(*ii*) pdf of benefits to carrying contraband, for contraband with generally larger benefits



(*i*) cdf of benefits to carrying contraband, for contraband with generally small benefits



(*ii*) cdf of benefits to carrying contraband, for contraband with generally larger benefits