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

DAVID BJERK
Claremont McKenna College and RAND Corporation

Abstract

This paper develops a model of racial profiling by law enforcement officers when officers observe both an individual's race as well as a noisy signal of his or her guilt that depends on whether or not a crime has been committed. The model shows that given officers observe such a guilt signal, data regarding the guilt rate among those investigated from each race will not be sufficient for determining whether racially unequal investigation rates are due to statistical discrimination or racial bias on the part of officers. The model also reveals that when racially unequal investigation rates are due to statistical discrimination, imposing a colorblind policy on officers can increase, decrease, or have little effect on the crime rate, depending on specific characteristics of the jurisdiction and the crime in question.

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% of whites surveyed and 77% of blacks surveyed. The same survey also showed that 72% of black men between the ages of 18 and 34 believed they had been stopped because of

---

David Bjerk, Claremont McKenna College, Department of Economics, Bauer Center, 500 E. Ninth St., Claremont, CA 91711-6400.

Thanks to Seungjin Han, Lance Lochner, Stephen L. Ross, the seminar participants at the University of Western Ontario, the Canadian Public Economics Research Group, and the referees and editors at this journal for their helpful comments and suggestions on earlier drafts. Also, I would like to acknowledge the support of the economics department at McMaster University, as much of this work was done while I was on faculty there.

Received April 4, 2005; Accepted June 14, 2006.

Journal of Public Economic Theory, 9 (3), 2007, pp. 521–545.

race (Ramirez, McDevitt, and Farrell 2000). Issues of racial profiling have also been much discussed with respect to individuals of Arab or Middle-Eastern descent being targeted for search by airport security guards (Amnesty International 2004). 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 meant to document or eradicate racial profiling by law enforcement officers.[1] One component of many of these policies has been to require officers to note the race of the individuals they investigate/search, as well as whether or not such investigations revealed any illegal behavior.

A key question with respect to implementing policies meant to eradicate profiling is how will such policies likely impact the prevalence of crime? The answer to this question is not obvious. On the one hand, if profiling is simply the result of officers being racially biased against one racial group relative to others, then it is unlikely that the imposition of anti-profiling policies would lead to substantial increases in the overall crime rate, and therefore such policies could easily be justified on ethical or constitutional grounds. On the other hand, profiling could be the result of some type of statistical discrimination, where differential treatment of observationally similar individuals from different races arises and persists, not because of any racial bias on the part of officers, but because it is the most efficient way for officers to maximize the expected guilt rate among searched or investigated individuals. While the ethics and constitutionality of this type of discrimination are also debatable, eradicating such behavior through anti-profiling policies may come at the cost of increasing the overall crime rate or increasing the cost of crime control.

This discussion suggests the potential importance of addressing the following three questions. First, what factors might lead unbiased officers to engage in statistical discrimination in policing in the first place? Second, is it possible to identify whether racially unequal investigation rates are due to statistical discrimination as opposed to racial bias using the data most jurisdictions are collecting (i.e., data on only those investigated)? And third, even if racially unequal investigation rates in a particular jurisdiction were known to be the result of statistical discrimination, to what degree will implementing an effective anti-profiling policy likely increase (or decrease) the crime rate?

To attempt to address these questions, I begin by developing a model of policing and criminal participation. This model is somewhat similar to those used for examining statistical discrimination in the context of the labor market, most notably Coate and Loury (1993). In the model developed 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 20 states have passed legislation prohibiting racial profiling 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. Law enforcement officers, on the other hand, wish to maximize the guilt rate of those they investigate. To do so, each officer must decide on a case by case basis whether or not to investigate each individual he observes based on a guilt signal emitted by the individual that depends on whether or not the individual is guilty, 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 even unbiased officers is to treat some 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 than the other. While this result in and of itself is relatively unsurprising, one of the key contributions coming out from the first part of this paper is to show that if statistical discrimination is the cause of racially unequal investigation rates, then it must be the case that the more frequently investigated race must also have a higher overall guilt rate than the other race, but the guilt rate among the *investigated* members of the more frequently investigated race can be greater than, equal to, or less than the guilt rate among the *investigated* members of the other race. These results lie in stark contrast to some of the previous works in this area (e.g., Knowles et al. 2001, Persico 2002, Hernandez-Murillo and Knowles 2004), and more generally, reveal that it is not possible to identify whether racially unequal investigation rates are due to statistical discrimination or racial bias using data on the guilt rates among the *investigated* members of each racial group, which, as discussed above, is generally what is being collected in many jurisdictions where the perception of racial profiling has become an issue.[2]

As alluded to above, even if racially unequal investigation rates were known to be the result of statistical discrimination rather than racial bias, it is not clear how large the costs to eradicating such behavior would likely be. Therefore, in the second part of the paper, I adapt the model discussed above to examine how the crime rate for a particular type of crime may change if a theoretical policy were implemented such that officers who police this crime are constrained to behave in a perfectly colorblind manner. I show that the degree to which such a colorblind policy would alter the crime rate relative to the situation when officer behavior was unconstrained 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

[2]However, as argued more recently by Antonovics and Knight (2004) and by Anwar and Fang (2006), it may be possible to use data on race of the investigating officer to examine racial bias, when data on the overall guilt rate of each group is unavailable.

(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 at-risk individuals who incur a large benefit from committing the relevant crime relative to those who only incur a small benefit.

I argue that these results suggest that the effects anti-profiling policies on criminal activity may differ dramatically by the type of criminal activity and jurisdiction in question. For example, imposing a colorblind policing policy on U.S. Highway Patrolmen may quite plausibly lead to only very small increases (or even decreases) in the overall number of drivers who transport small amounts of contraband on highways. By contrast, however, the model also suggests that the crime rate costs to implementing a anti-profiling policy may be much larger in other contexts, such as security at an Israeli airport.

2. Previous Literature

Knowles et al. (2001), Persico (2002), and Hernandez-Murillo and Knowles (2004) (hereafter referred to as "KPT and related papers") were among the first studies to formally analyze the potential role of statistical discrimination within the context of racial profiling in police work. These papers essentially focused on 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. While a group generally refers to members of a particular race, a group can more generally refer to any set of individuals with similar observable characteristics. However, these characteristics must be determined prior to the individual choosing whether or not to commit the relevant crime or at least the realization of these characteristics must not depend on whether a crime was committed.

The equilibrium arising in KPT and the related papers is one in which police and individuals of all groups each play mixed strategies. In particular, 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 to the rate at which police are just indifferent between searching and not searching, and therefore also equal across all groups. Such behavior from police can be maintained in equilibrium since because the guilt rate for each group makes police just indifferent between searching and not searching any given individual, 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 focus all of their investigations on this higher crime group (since the expected probability of catching a law-breaker would be higher). However, this would then cause the other groups that were not being investigated to choose to break the law (for example carry contraband), and cause individuals in the group being focused on to cease criminal activity, which would cause police to want to switch their focus to others groups. 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.

In this equilibrium, 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 others. 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 2004) 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 assert 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 among the searched 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 a very important starting point for thinking rigorously about statistical discrimination in the context of policing and highlight many of the important forces inherent in thinking about racial profiling, the equilibria discussed in these papers lead to some unappealing results for this context. In particular, 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 not investigated. In practice, this implies that officers'

[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).

experiences do not make them any better at spotting law-breakers than a machine that picks individuals at random.

The key difference between the model developed in this paper and that developed in KPT and the related papers, is that I assume that in addition to race, officers observe a signal of guilt from each motorist they observe, where the actual realization of this signal depends on whether an individual committed the relevant crime. This signal is meant to capture observable characteristics that are more likely to arise when a crime has been or is being committed, but are such that, outside of deciding whether or not to commit a crime, an individual has little control over the actual realization of this characteristic. For example, erratic driving may be more likely to arise from motorists under the influence of drugs or alcohol. Similarly, individuals who have something to hide may often give vague or contradictory answers to simple questions from police or border officials. Finally, driving at odd hours or driving through a neighborhood far from home may often be necessary for committing crime, and therefore individuals doing such things may be relatively more likely to be criminals than individuals driving during rush hour or in their own neighborhoods.

Anwar and Fang (2006) also incorporate this notion that by committing a crime, individuals become more likely to emit certain traits or characteristics to police. However, unlike the model presented here, instead of examining the potential effects of anti-profiling policies, Anwar and Fang (2006) use their model to examine how race of the investigating officer may be used to distinguish between statistical discrimination and racial bias.[4] Dharmapala and Ross (2004) derive a model very similar to that in KPT and the related papers (i.e., without the presence of observable guilt signals), but show that the results of such a model can change dramatically if officers are not able to observe all relevant individuals. The reason being that if it is known that officers will not even observe some fraction of individuals, then even if officers have a policy of investigating all individuals who exhibit particular characteristics, individuals with these characteristics may still rationally commit crime. This breaks one of the crucial equilibrating forces integral to the results in KPT and the related papers. Therefore, for completeness, the model developed below also allows for the possibility that law enforcement officers are unable to observe all individuals.

As discussed above, after deriving the equilibrium of the basic model developed here, I then use the model to characterize how implementing a colorblind policy will affect the overall crime rate. In this way, this paper is most similar in spirit to Farmer and Terrell (2001) and Persico (2002). However,

[4]Moreover, also unlike this paper, in their main model, Anwar and Fang (2006) assume the fraction of each race that choose to commit crime is exogenous and not influenced by law enforcement behavior. However, in an appendix, they do show that their model can be altered to allow for crime to depend on officer's behavior.

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, 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 cumulative distribution function G, defined to be a continuous nondecreasing function over the support $(0, \bar{\epsilon}]$ (i.e., "at-risk" individuals are defined to be those for whom $\epsilon_i > 0$). However, assume that the two racial groups differ in that a fraction λ_a of race a are *at-risk* individuals, while a fraction λ_b of race b are *at-risk* individuals, with $\lambda_a < \lambda_b$. Note that this construction provides a particularly simple way of parameterizing differences in the *overall* distribution of ϵ_i across races.

---

[5]Specifically, 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, as discussed previously, allows for the possibility that police cannot observe all individuals (something not allowed for in either of the two 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, known before the crime commission decision is made. For example, age groups, gender groups, or race/gender/age group combinations can substitute for racial groups without loss of generality.

[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 crossing, the relevant crime may be supporting or planning terrorist activity.

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 to committing crime is prohibitively high for relatively high-income individuals, then one can interpret the "at-risk" population as being those individuals who are poor, with λ_j indicating the fraction of race j that is poor. The distribution G is then the distribution of benefits to committing crime (inclusive of any opportunity costs as well as disutility or utility associated with such acts) for the poor.[9] Alternatively, $\lambda_a < \lambda_b$ may arise if the utility benefit to committing the relevant crime is negligible for everyone after reaching a certain age, meaning young people constitute the at-risk population and race b has a higher fraction of young people than race a. 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.

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 individual who they observe, which is a fraction $\eta \in (0, 1]$ of the total population. 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 $\hat{p}_{g,j}$.

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 $\epsilon_i > \eta\hat{p}_{g,j}c$. Therefore, the proportion of *at-risk* individuals from race j who choose to commit the relevant crime equals $1 - G(\eta\hat{p}_{g,j}c)$, which implies that the overall fraction of race j that is guilty of committing the crime is given by

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

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

---

[9]This interpretation is essentially equivalent to the differences across races allowed in Persico's (2002) model.

[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.

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

When observing an individual, a law enforcement officer cannot directly observe whether an individual committed a crime, or even if the individual is a member of the "at-risk" population. Rather, officers can only observe an individual's race and a guilt signal denoted $\theta \in [0, 1]$, where this signal is some characteristic of the individual that is correlated with whether or not the individual is guilty. In particular, 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. Assume F_n and F_g are both twice continuously differentiable in θ. 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 $[0, 1]$, meaning the relative likelihood that a particular individual is guilty is greater the higher the signal.

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 underlying proportion of each race that is guilty. If we let these beliefs concerning race j be denoted $\hat{\pi}_j$, then 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 equals

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

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 $\frac{\partial \xi(\hat{\pi}_j, \theta)}{\partial \theta} > 0$ and $\frac{\partial \xi(\hat{\pi}_j, \theta)}{\partial \hat{\pi}_j} > 0$. In other words, the 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 these beliefs, for each individual they observe, law enforcement officers must decide whether or not to investigate, where an investigation is assumed to perfectly reveal whether a given individual is guilty or innocent. Let the cost to investigating an individual equal q, while the benefit to arresting a guilty individual is denoted v, where $v > q$. The benefit v can be interpreted as both the career advancement and any salary benefits officers receive for catching a law-breaker as well as the sense of satisfaction and accolades such an arrest incurs. The cost of investigation q can be interpreted as the opportunity cost of the time it takes to perform an investigation and the associated paperwork.

Similar to most other papers in this literature (e.g., KTP and the related papers, Dharmapala and Ross 2004, Anwar and Fang 2006, Antonovics and Knight 2004), let us consider an environment where law enforcement officers choose investigation strategies so as to maximize the expected net benefit of their time, meaning they will investigate an observed individual of race j if and only if the expected benefit to doing so exceeds the opportunity cost they must incur by doing so. Implicitly, this assumes officers attempt to maximize the guilt rate among those they investigate. Therefore, an officer will investigate an observed individual from race j if and only if the individual emits a signal such that $\xi(\hat{\pi}_j, \theta)v > q$, or only if

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

Given the above expression and the fact that $\frac{f_n(\theta)}{f_g(\theta)}$ is decreasing in θ, we know that there will exist a threshold level $\theta^*(\hat{\pi}_j)$ such that for any belief $\hat{\pi}_j$, Equation (3) holds if and only if $\theta \geq \theta^*(\hat{\pi}_j)$, where $\theta^*(\hat{\pi}_j) = 0$ if $\hat{\pi}_j$ is such that Equation (3) holds for all $\theta, \theta^*(\hat{\pi}_j) = 1$ if $\hat{\pi}_j$ is such that Equation (3) never holds for any θ, and $\theta^*(\hat{\pi}_j)$ equals the value of θ that causes (3) to hold at equality if possible. It can be easily confirmed that when $\theta^*(\hat{\pi}_j)$ is strictly between 0 and 1, it is strictly decreasing in $\hat{\pi}_j$, or the greater the prior belief of guilt, the lower the signal necessary for investigation.

Given this investigation 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

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

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 b than race a. Equation (3) shows that the effect of such a bias on police behavior would be to lower $\theta^*(\hat{\pi}_b)$ relative to $\theta^*(\hat{\pi}_a)$ for any given $\hat{\pi}_b$ and $\hat{\pi}_a$, which in turn would increase $p_g(\hat{\pi}_b)$ relative to $p_g(\hat{\pi}_a)$, all else equal.

3.3. Equilibrium

As shown above, given beliefs $\hat{p}_{g,j}$, optimal behavior for each individual i from race j is to participate in criminal activity if and only if $\epsilon_i > \eta \hat{p}_{g,j} c$. Similarly, given beliefs $\hat{\pi}_j$ over individuals from race j, optimal behavior by law enforcement officers is to investigate an individual from race j if and only if the individual emits a guilt signal greater than $\theta^*(\hat{\pi}_j)$. In equilibrium, individuals and officers must behave according to the optimal behavior rules specified above given their beliefs, and these beliefs must correspond to the truth. More specifically, for each race j, equilibrium beliefs must correspond to a pair $\{\hat{p}_{g,j}, \hat{\pi}_j\}$ that simultaneously satisfy $\hat{\pi}_j = \pi(\lambda_j, \hat{p}_{g,j})$ and $\hat{p}_{g,j} = p_g(\hat{\pi}_j)$, where $\pi(\lambda_j, \hat{p}_{g,j})$ is defined in Equation (1) and $p_g(\hat{\pi}_j)$ is defined in Equation (4). The Appendix proves that for any λ_j, there exists a unique equilibrium pair of such beliefs, hereafter denoted $\{p^e_{g,j}, \pi^e_j\}$. 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 $\pi^e_b > \pi^e_a$ (see the Appendix for proof). 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 lawbreakers than the other racial group. While this result is quite intuitive, it is worth noting that this result lies in 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/investigating any two individuals of different races. By contrast, in the equilibrium of the model presented here, police are only indifferent between investigating 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 $\theta^*(\pi^e_j)$ that arises given equilibrium beliefs.[11]

If we restrict our attention to parameterizations where $\theta^*(\pi^e_b) > 0$ or where $\theta^*(\pi^e_a) < 1$,[12] then because $\lambda_b > \lambda_a$ implies $\pi^e_b > \pi^e_a$, we also know

[11] Note that there is no reason for the marginal member of race j to have the same likelihood of being guilty as the average (i.e., randomly chosen) member of race j. This is in contrast to KPT and the related papers, where there was no marginal member of race j, since all members of the same race had the same likelihood of guilt and police decided who to investigate at random. Ayers (2002), and Anwar and Fang (2006) also emphasize this important point with respect to profiling policing. In a different but related context, Ross and Yinger (2002) refer to an issue similar to this as the "infra-marginality problem" in their discussion of discrimination in mortgage lending.

[12] 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.

that $\theta^*(\pi_b^e) < \theta^*(\pi_a^e)$. In 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 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).

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

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

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 in a particular jurisdiction as

$$\Pi^e = \beta\lambda_b\left[1 - G\left(\eta p_{g,b}^e c\right)\right] + (1-\beta)\lambda_a\left[1 - G\left(\eta p_{g,a}^e c\right)\right]. \tag{6}$$

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 $1 - G(\eta p_{g,b}^e c) > 1 - G(\eta p_{g,a}^e c)$. 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

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

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 who are observed by police and emit a guilt signal high enough to warrant a search. Taking the derivative of the above expression gives

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

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 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 $\pi_b^e > \pi_a^e$).[13] 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 et al. (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.[14]

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

[13] It is important to note that if F_n and F_g differ by race, then this results and many of the others discussed in this section would not necessarily hold. However, while it is certainly plausible that F_n and F_g differ by race in some instances, the results in this paper could still hold under appropriate restrictions placed on the differences between F_n and F_g across races. More generally though, this is an important issue to consider, and deriving the general implications of differences in these distributions across races provides a potentially interesting avenue for further research.

[14] To see why this is true, recall from Equation (3) that a lower v/q means a higher $\theta^*(\hat{\pi}_j)$ for any given beliefs $\hat{\pi}_j$. 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 for race b to be investigated at a higher rate even though they have a lower crime rate.

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

It is straightforward to show that because $F_n(\theta) > F_g(\theta)$ for all θ (a direct implication of the decreasing likelihood ratio assumption), $\gamma(\pi_j^e)$ 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 $\pi_b^e > \pi_a^e$ (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 the Knowles et al. (2001) "test" for determining whether unequal investigation rates are due to police officer racial bias or statistical discrimination is not necessarily valid. To put 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 when officers are able to observe signals of guilt beyond race, guilt rates among the investigated are not necessarily equalized across races, even with racially unbiased police officers.

In fact, as shown in the Appendix, the necessary (but not even sufficient) conditions that must be met for racially unequal investigation rates to coexist with racially equal guilt rates among the investigated, are very stringent. In general, 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 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% 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 2004).[15] Similarly, only about 9.5% and 22% of non-white motorists who were stopped were searched by Rhode Island State Troopers (Farrell et al. 2003) and the Los Angeles

[15]This compares to only 6.4% of white drivers who were stopped were searched.

Police Department (Los Angeles Police Department 2002), respectively.[16] Also, the guilt rates among searched non-white motorists in Missouri, Rhode Island, and Los Angeles, were 15.6%, 13.9%, and 29.4%, respectively.[17] 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 solely 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 investigate, eradicating such behavior may have costs in terms of increasing overall crime rates, especially if such behavior is not a result of racial bias or bigotry. This section uses the model developed in Section 3 to analyze these potential crime costs to an anti-profiling policy, and how these costs 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 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, $\hat{\Pi}$. Given this belief, in the same manner as in Section 3, optimal behavior for officers will be to choose a threshold level $\theta^*(\hat{\Pi})$, such that they will investigate an observed individual only if they observe a guilt signal greater than $\theta^*(\hat{\Pi})$, where $\theta^*(\hat{\Pi})$ is derived

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

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

analogously to before. Therefore, for individuals of both races, the probability of being investigated if guilty equals $\eta p_g(\hat{\Pi})$, where

$$p_g(\hat{\Pi}) = 1 - F_g(\theta^*(\hat{\Pi})). \tag{8}$$

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

$$\Pi(\lambda_b, \lambda_a, \hat{p}_{g,b}, \hat{p}_{g,a},) = \beta\lambda_b[1 - G(\eta\hat{p}_{g,b}c)] + (1-\beta)\lambda_a[1 - G(\eta\hat{p}_{g,a}c)]. \tag{9}$$

As before, in equilibrium, individuals and officers behave optimally given their beliefs, and beliefs must be such that $\hat{p}_{g,a} = \hat{p}_{g,b} = p_g(\hat{\Pi})$ (as described by Equation (8)) and $\hat{\Pi} = \Pi(\lambda_b, \lambda_a, \hat{p}_{g,b}, \hat{p}_{g,a})$ (as described by Equation (9)). Using a similar argument as in the unconstrained case, we can prove such equilibrium beliefs exist and are unique (see Appendix). Denote the 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 $\pi_a^e < \Pi^c < \pi_b^e$ (proof in Appendix). 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 $p_{g,a}^e < p_g^c < p_{g,b}^e$.[18] In words, the implementation of a policy that causes police to behave in a colorblind manner will increase the probability of investigation for guilty members from race a, while decreasing the probability of investigation 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 guilt rate among those they investigate. However, it is not clear how such a colorblind policy will affect the overall rate at which the relevant crime is being committed, which is arguably society's more primary concern. In the context of this model, this question reduces to calculating $\Pi^c - \Pi^e$, where Π^c is described by Equation (9) and Π^e is described by Equation (6). Writing this expression out and rearranging, we obtain

$$\Pi^c - \Pi^e = \beta\lambda_b\big[G\big(\eta p_{g,b}^e c\big) - G\big(\eta p_g^c c\big)\big] + (1-\beta)\lambda_a\big[G\big(\eta p_{g,a}^e c\big) - G\big(\eta p_g^c c\big)\big]. \tag{10}$$

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

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

be negative. Therefore, the degree to which a colorblind policy increases the overall crime rate simply reduces to the relative magnitude of the first product versus the second product in Equation (10). Intuitively, the first product is the degree to which the policy increases the relative number of group b individuals committing the relevant crime, while the second product is the degree to which the policy decreases the relative number of group a individuals committing the relevant crime.[19]

The first thing to note about Equation (10) is that it shows that even though the results of this model differ from Persico's (2002) on a number of key dimensions, one of his key results remains even in this more general environment where law enforcement officers observe other signals correlated with guilt beyond race. Namely, Equation (10) shows that 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. The intuition is quite simple—the policy which minimizes the overall number of individuals committing the relevant crime may be quite different than the policy that maximizes the guilt rate among the investigated.

More generally, Equation (10) implies that any 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) and the closer λ_b is to λ_a. In words, this later implication says that 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. This suggests 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.

Unlike Persico's analysis, Equation (10) also allows us to assess how a variety of other jurisdictional and relevant crime characteristics may influence the effect of a colorblind policy on the crime rate. For example, Equation (10) suggests 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 $G(\eta p^c_g c) - G(\eta p^e_{g,a}c)$.[20] 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

[19]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 (10) is simply comparing the relative elasticities of each race to more and less intensive policing.

[20]It should be said, however, that this is not necessarily the case. This result depends on the shape of the distribution of benefits to carrying contraband (i.e., the shape of G).

in the first place. Therefore, constraining law enforcement officer behavior a little bit will not have a large effect on each individual's criminal participation decision. For analogous reasons, any 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 (10) is that the crime rate cost to the policy will depend on the shape of the cumulative distribution function G, or how the benefits of committing the particular crime are distributed across at-risk individuals. In particular, the crime rate cost to a colorblind policy will generally be smaller for crimes in which most individuals at-risk of committing these crimes would incur only a relatively small benefit from doing so. Alternatively, the crime rate cost to the policy will generally be much larger for crimes where a relatively large fraction of at-risk individuals incur a substantial benefit from committing such crimes. This implication is illustrated in Figures 1(a) and 1(b). In Figure 1(a), the cumulative distribution function G is such that the bulk of at-risk individuals would incur only a small benefit from committing the relevant crime. As can be seen, this means that $|G(\eta p_g^c c) - G(\eta p_{g,a}^e c)|$ is relatively large, while $|G(\eta p_g^c c) - G(\eta p_{g,b}^e c)|$ is relatively small. In words, when most at-risk individuals only incur a small benefit from committing the relevant crime, a relatively large fraction of at-risk individuals from race a would be deterred by the colorblind policy, but only a relatively small fraction of at-risk individuals from race b would be induced into criminal activity by the colorblind policy. Alternatively, Figure 1(b) shows a cumulative distribution function G where the benefits to committing the relevant crime are more diverse across at-risk individuals. As shown in the figure, this means that $|G(\eta p_g^c c) - G(\eta p_{g,a}^e c)|$ is relatively small, while $|G(\eta p_g^c c) - G(\eta p_{g,b}^e c)|$ is relatively large, meaning relatively few at-risk individuals from race a would be deterred by the colorblind policy, but a relatively large fraction of at-risk individuals from race b would be induced into crime by the colorblind policy.

Given the above analysis, an important question to ask is to what extent can these implications help us think about the potential effects of implementing colorblind policies in different contexts? Clearly, it is difficult, if not impossible to know exactly how the relevant parameters differ across different types of crimes and jurisdictions. However, we can make some reasonable guesses. For example, consider crimes such as driving on the highway in possession of small amounts of drugs or alcohol, or in possession of a firearm. In such cases, it is reasonable to expect that 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). Moreover, given police generally only observe a small fraction of all drivers (small η), the costs of being arrested for possession of a small amount of contraband are likely small (i.e., small c), and black drivers are likely to be a relatively small fraction of all drivers on the highways (small β), the arguments above

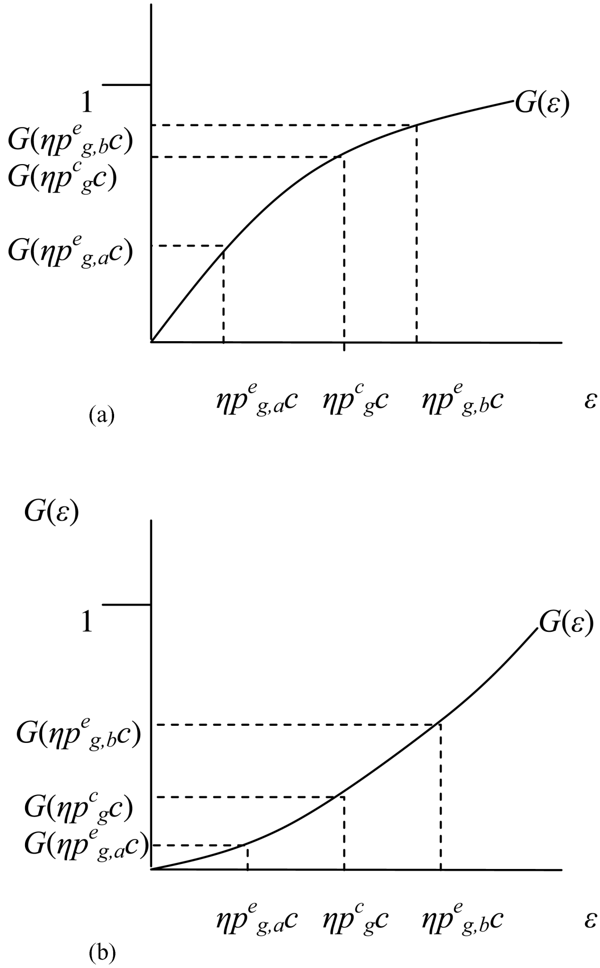

Figure 1: CDF of benefits to committing crime ($G(\varepsilon)$): (a) for crimes where most at-risk individuals incur small benefits; (b) for crimes where the benefit to crime is relatively diverse across at-risk individuals.

suggest that a colorblind policy on highway patrol officers will arguably lead to only a small increase (or possibly even a decrease) in the overall rate at which small amounts of contraband are transported in cars.

Alternatively, consider a crime such as sabotaging an airliner in a place such as Israel. One could certainly argue that a substantial fraction of those individuals at-risk of committing such a terrorist act would incur a large benefit from doing so. Moreover, Israeli airport security guards likely observe a high fraction of all individuals who board planes at Israeli airports (high η), the cost to being arrested for attempting to sabotage an airliner is large (high c), the overall population may contain a relatively large fraction of the potentially profiled race (in this case those of Arab descent) (high β), and even though the overall fraction of individuals at-risk of sabotaging an Israeli airliner may be very small for all racial groups, the relative frequency is likely quite a bit higher for those of Arab descent than Israelis (high λ_b relative to λ_a, where race b are Arabs and race a are Israelis). Therefore, unlike with respect to the highway patrol example above, these relative parameter magnitudes suggest that implementing a colorblind policy on Israeli airport security guards may have a potentially large impact on the relative frequency of airline sabotage.

Finally, it is also worth noting that the benefits to this colorblind policy can be seen on two fronts. First, the probability that an innocent member of race b is investigated falls from $1 - F_n(\theta^*(\pi_b^e))$ to $1 - F_n(\theta^*(\Pi^e))$.[21] 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 from $F_g(\theta^*(\pi_a^e)) - F_g(\theta^*(\pi_b^e))$ 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. 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 officers trying to maximize the expected guilt rates of their investigations 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.

[21] 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 $1 - F_n(\theta^*(\Pi^e))$.

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, but 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.

The final results coming from the model discuss the extent to which a policy that eliminates racial profiling will increase (or possibly decrease) the overall crime rate in a particular jurisdiction, given the characteristics of 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 socioeconomic 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 will generally incur large benefits from committing the relevant crime. As discussed above, these results suggest that implementing policies that eradicate racial profiling may have very differential effects on the crime rate depending on the type of crime and the jurisdiction in question.

Appendix

(a) *Proof of existence and uniqueness of equilibrium in unconstrained environment.* Recall that equilibrium is defined as a pair of beliefs $\{p^e_{g,j}, \pi^e_{g,j}\}$ such that

$$\pi^e_j = \lambda_j\left[1 - G\left(\eta p^e_{g,j} c\right)\right] \quad \text{(A1)}$$

and

$$p^e_{g,j} = 1 - F_g\left(\theta^*\left(\pi^e_j\right)\right). \quad \text{(A2)}$$

Hence, the equilibrium value for the proportion of race j carrying contraband, π_j^e, can be found by substituting Equation (A2) into Equation (A1) and solving for π_j^e. Doing this we get

$$\pi_j^e = \lambda_j[1 - G(\eta[1 - F_g(\theta^*(\pi_j^e))]c)]. \tag{A3}$$

Existence of such a π_j^e can be confirmed by first defining

$$h(\pi) = \pi - \lambda_j[1 - G(\eta[1 - F_g(\theta^*(\pi))]c)].$$

Now, note that there exists a π_j^e such that Equation (A3) holds if there exists a π_j^e such that $h(\pi_j^e) = 0$. To prove that such a π_j^e exists, first note that given our previous assumptions, we know $h(0) < 0$ and $h(1) \geq 0$. Therefore, by the intermediate value theorem, we know that if $h(\pi)$ is continuous, there exists a π_j^e such that $h(\pi_j^e) = 0$.

Clearly, since G and F_g were assumed to be continuous, $h(\pi)$ is continuous as long as $\theta^*(\pi)$ is a continuous function of π. To prove this is true, define $\psi(\theta) = \frac{f_n(\theta)}{f_q(\theta)}$. Given $\frac{f_n(\theta)}{f_q(\theta)}$ was assumed to be a continuous decreasing function of θ, we know $\psi(\theta)$ has a well-defined continuous inverse, meaning

$$\theta^*(\pi) = \psi^{-1}\left(\frac{\pi \frac{v}{q} - \pi}{1 - \pi}\right),$$

which is a continuous function of π. Therefore, there exists an equilibrium π_j^e for each race j.

Moreover, this equilibrium value is unique. To see why this is true, we can take the derivative of $h(\pi)$, which gives

$$\frac{\partial h(\pi)}{\partial \pi} = 1 - \lambda_j g(\eta[1 - F_g(\theta^*(\pi_j^e))]c)\eta f_g(\theta^*(\pi_j^e))\frac{\partial \theta^*(\pi_j^e)}{\partial \pi_j^e}c.$$

Since $\frac{\partial \theta^*(\pi_j^e)}{\partial \pi_j^e} \leq 0$ (as discussed in the text), the above expression is strictly positive, meaning $h(\pi)$ is strictly increasing. Therefore, there can be only one value for π_j^e such that $h(\pi_j^e) = 0$.

Finally, Equation (A2) shows that $p_{g,j}^e$ 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 (A3) 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 (A3) is increasing in λ_j. Therefore, as λ_j increases, in order maintain equality in Equation (A3), π_j^e must also increase, as by doing so the left-hand side of Equation (A3) will increase and right-hand side of Equation (A3) will decrease.

(c) *The necessary conditions for racially unequal investigation rates to coexist with racially equal guilt rates among the investigated when enforcement officers are truly racially unbiased.* 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

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

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

$$\frac{\left[1-F_n\left(\theta^*(\pi_b^e)\right)\right]}{\left[1-F_g\left(\theta^*(\pi_b^e)\right)\right]} > \frac{\left[1-F_n\left(\theta^*(\pi_a^e)\right)\right]}{\left[1-F_g\left(\theta^*(\pi_a^e)\right)\right]}. \quad \text{(A4)}$$

Now, define $\tilde{\theta}$ to be the value of θ such that $f_g(\tilde{\theta}) = f_n(\tilde{\theta})$.[22] 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 (A4) 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.

(d) *Proof of existence and uniqueness of equilibrium in colorblind environment.* Note that equilibrium can be defined by a pair of beliefs $\{p_g^c, \Pi^c\}$ such that

$$\Pi^c = \beta\lambda_b\left[1-G\left(\eta p_g^c c\right)\right] + (1-\beta)\lambda_a\left[1-G\left(\eta p_g^c c\right)\right] \quad \text{(A5)}$$

[22]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.

and

$$p_g^c = 1 - F_g(\theta^*(\Pi^c)). \tag{A6}$$

Now, note that Equation (A5) can be rewritten as

$$\Pi^c = \lambda_c\left[1 - G\left(\eta p_g^c c\right)\right], \tag{A7}$$

where $\lambda_c = [\beta\lambda_b + (1-\beta)\lambda_a]$. Since these equations are of the same form as Equations (A2) and (A1), the same argument as above can be used to prove existence and uniqueness of Π^c and p_g^c.

(e) *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

AMNESTY INTERNATIONAL (2004) *Threat and Humiliation: Racial Profiling, Domestic Security, and Human Rights in the United States*, http://www.amnestyusa.org/racial_profiling/report/rp_report.pdf.

ANTONOVICS, K., and B. KNIGHT (2004) A new look at racial profiling, NBER Working Paper #10634.

ANWAR, S., and H. FANG (2006) An alternative test of racial prejudice in motor vehicle searches: Theory and evidence, *American Economic Review* **96**, 127–151.

AYRES, I. (2002) Outcome tests of racial disparities in police practices, *Justice Research and Policy* **4**(Special Issue), 131–142.

BUSH, G. W. (2001) Memorandum for the Attorney General on racial profiling, February.

COATE, S., and G. LOURY (1993) Will affirmative action eliminate negative stereotypes? *American Economic Review* **83**, 1220–1240.

DHARMAPALA, D., and S. L. ROSS (2004) Racial bias in motor vehicle searches: Additional theory and evidence, *Contributions to Economic Analysis* **3**, 1–21.

FARMER, A., and D. TERRELL (2001) Crime versus justice: Is there a trade-off? *Journal of Law and Economics* **XLIV**, 345–366.

FARRELL, A., J. McDEVITT, S. CRONIN, E. PIERCE (2003) Rhode Island Traffic Stop Statistics Act final report, Northeastern University Institute on Race and Justice, June.

HERNANDEZ-MURILLO, R., and J. KNOWLES (2004) Racial profiling or rascist policing?: Testing in aggregate data, *International Economic Review* **45**, 959–989.

KNOWLES, J., N. PERSICO, and P. TODD (2001) Racial bias in motor vehicle searches: Theory and evidence, *Journal of Political Economy* **109**, 203–229.

LOS ANGELES POLICE DEPARTMENT, CITYWIDE DRIVER DATA (2002) 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.

PERSICO, N. (2002) Racial profiling, fairness, and effectiveness of policing, *American Economic Review* **92**(5), 1472–1497.

RACIAL PROFILING DATA COLLECTION RESOURCE CENTER (2003) Legislation and litigation, http://www.racialprofilinganalysis.neu.edu/legislation.php, accessed October 30.

RAMIREZ, D., J. McDEVITT, and A. FERRELL (2000) A resource guide on racial profiling data collection systems: Promising practices and lessons learned, *U.S. Department of Justice Monograph*, November.

ROSS, S. L., and J. YINGER (2002) *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair Lending Enforcement*. Boston: MIT Press.

VERNIERO, P., and P. H. ZOUBEK (1999) 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, accessed April 20.