Testing for Racial Bias in Law Enforcement[1]

Petra E. Todd
University of Pennsylvania

[1]This article was prepared for the Palgrave Dictionary of Economics. I thank Nicola Persico for helpful discussions.

1 Introduction

In recent years, numerous lawsuits have been brought against U.S. city police departments alleging racially biased law enforcement practices.[1] As a result, many police departments now routinely collect data on the characteristics of the individuals that they subject to stops and searches and on the outcomes of these encounters. The case for discrimination often rests on the observation that the proportion of African-Americans or Hispanics among drivers searched by police far exceeds the proportion in the general population.[2] For example, in the state of Maryland, African Americans represented 63% of motorists searched by state police on the I-95 highway between January 1995 and January, 1999, but only 18% of motorists on the road. A refined version of this benchmark test estimates the probability of being searched as a function of race and other observable characteristics thought to be related to criminal propensity. If race has explanatory power in the regression, then this is taken as evidence of discrimination.[3]

One drawback of benchmark tests is that they require data on the full set of characteristics that a police officer uses in deciding whether to search a motorist.[4] If some characteristics are missing from the data, then race could have explanatory power due to omitted-variable bias. Alternatively, if race is found to be insignificant, there is still the possibility that police target individuals with certain characteristics, because those characteristics are correlated with race and not because they are good predictors of criminality. Thus, the validity of benchmark tests for discrimination hinges crucially on judgments about what constitutes a set of admissible conditioning variables and on whether the analyst has access to the full

[1] Many of these lawsuits were initiated by the ACLU. Some investigations have also been initiated by the U.S. Department of Justice.

[2] In the 1993 case of *Maryland v. Wilkins*, for example, a statistician testified that "The disparities are sufficiently great that, taken as a whole, they are consistent with and strongly support the assertion that the Maryland State Police are targeting the community of black motorists for stop, detention and investigation...".

[3] See, e.g., Donohue (1999).

[4] A training manual issued by the Illinois State Police highlights some indicators of criminal activity. These include such characteristics as tinted windows, cellphones, leased vehicles, religious paraphernalia used to divert suspicion, and attorney's business cards. Source: John Donohue, Expert Witness Testimony in *ACLU v. State of Illinois*.

set of variables. Moreover, benchmark tests are only informative about whether a disparity exists and not about the motivation for the disparity. A key question of interest in many racial profiling investigations is whether higher rates of stop and search among certain groups can be justified as an optimal monitoring response to higher rates of criminality.

The judicial standpoint on racial profiling is not clear-cut. The dominant view seems to be that race or ethnicity can be used as a factor in determining the likelihood that a person is engaging in or has committed a crime, as long as its use is related to law enforcement and is not a pretext for racial harassment (Kennedy 1997).[5] However, there is also a significant dissenting tradition that argues that race should not be used as a criteria, except in very limited cases, as when the race of a perpetrator of a particular crime is known.[6] Whether racial disparities in law enforcement are deemed reasonable or not by the courts depends on assessments about the degree to which discrimination assists in apprehending criminals, the benefits of apprehending criminals and the costs imposed on people erroneously searched or detained. In evaluating the legality of racial disparities in law enforcement, the courts have clearly sought to determine the motivation for discriminating.[7]

This article reviews the recent theoretical and empirical literature in economics that aims to empirically distinguish whether police behavior is indicative of racial bias from data on stops and searches of motor vehicles. The early literature on crime (e.g. Becker (1968) and Stigler (1970)) examined citizens' incentives to misbehave under an exogenous probability of being monitored. The more recent literature assumes that police and citizens behave strategically, with police deciding on optimal search strategies and citizens deciding

[5]For example, in an early influential case *United States v Martinez-Fuerte*, the Supreme Court found it permissible for police at border crossings to detain motorists of Mexican ancestry and to search their vehicles, because these cars were more likely to be transporting illegal aliens. (*United States v Martinez-Fuerte,* 428 U.S. 543 (1976)). The U.S. Court of Appeals adopted a similar position in *United States v Weaver* in permitting airport searches of young black males prompted by suspicions that members of a black gang were trafficking in cocaine. (*United States v Weaver*, 966 F.2d at 394, n.2 (1992).)

[6]In *United States v Laymon*, for example, the court suppressed incriminating evidence found in a vehicle, arguing that the police officer did not have sufficient justification for searching the car and that race was a factor in the decision to search. *United States v Laymon*, 730 F.Supp. 332 (D. Colo. 1990)

[7]For detailed discussions of the legality of racial profiling practices, see for example Kennedy (1997), Harcourt (2003), Gross and Barnes, and Persico and Castleman (2005).

whether to break the law, given police search strategies. Unlike the criminology literature, where it is sometimes assumed that police can make citizens believe that the probabilities of being monitored are higher than they really are (e.g. Sherman, 1990), the economics literature typically assumes that individuals subject to police monitoring have rational expections about the probability of being monitored. [8] A major advance in the literature is a better understanding of the assumptions on police and motorist behavior required to justify alternative tests for discrimination.

We next describe the frameworks that have been developed in the recent economic literature and the tests that have been derived using these frameworks. Subsequently, we present a brief summary of some of the empirical evidence from police stop/search datasets.

2 Theoretical Models of Police-Motorist Behavior

There are two leading model paradigms put forth in the economic literature. One framework assumes that police officers operate in a fairly decentralized way, allocating their individual search activities so as to catch as many criminals as possible. In the context of motor vehicle searches, the goal is to maximize the number of successful searches given a cost of search, where a successful search is defined as one that uncovers some contraband. As noted in Persico (2002), an objective function that maximizes successful searches, or so-called hit rates, will, in general, not minimize the aggregate crime rate, because it does not give enough weight to deterrent effects of policing. Namely, it gives no reward to preventing a crime from being committed.

Nevertheless, a hit rate objective may still be a reasonable approximation to police behavior, in light of principal-agent considerations in policing. It is likely difficult for a police chief to verify that individual officers are engaging in search activities that deter crime, because the amount of crime they deter is not observed. How many criminals an officer apprehends is observed, providing a rationale for rewarding officers on the basis of their

[8] The recent literature on racial profiling is also related to the literature on optimal auditing, which mainly deals with income reporting and tax evasion (see Reinganum and Wilde 1986, Border and Sobel 1987, and Scotchmer 1987).

performance record in uncovering crime. A model where police act as independent agents trying to catch criminals can be viewed as a second best objective that a police chief might reasonably adopt.

An alternative modeling framework examined in the literature is one in which there is a centralized police chief who is able to allocate resources in a way that minimizes the overall crime rate. We describe the theoretical and empirical results derived using these two different modeling frameworks, with particular emphasis on how to devise tests for racial bias.

2.1 Models of Hit Rate Maximization

2.1.1 The Model of KPT (2001)

Knowles, Persico and Todd (2001) (KPT) develop a model of police-motorist behavior that they use to study the implications of racial bias for equilibrium search outcomes. In the model, police officers decide which vehicles to subject to searches and motorists decide whether to break the law by carrying contraband, such as drugs or illegal weapons. It is assumed that motorists take into account the probability of being searched when they make their decision about whether to carry.

In the absence of racial bias, each officer pursues a monitoring strategy that maximizes the number of successful search outcomes. Racial bias is introduced as a preference parameter that reduces the perceived cost of searching vehicles of black or Hispanic drivers relative to white drivers, which can lead to oversearching of these groups. An equilibrium implication of racially biased monitoring, shown in KPT and discussed further below, is that the hit rate (the rate at which contraband is seized) should be lower for the groups subject to bias.[9]

Let $r \in \{A, W\}$ denote the race of the motorist (African American or white), assumed to be observable by the police officer.[10] Let c denote all characteristics other than race that are potentially used by the officer in the decision to search cars. The variable c may be

[9] The general idea that tastes for discrimination lead to lower profits for discriminators originated with Becker (1957). For further discussion of such tests in policing contexts, see Ayres (2002).

[10] We assume two groups here for simplicity. The analysis extends straightforwardly to more groups.

unobserved or only partially observed by the econometrician. For expositional ease, treat c as a one-dimensional variable (results extend to the multidimensional case), and denote the distribution of c in the white and African-American populations by $F(c|W)$ and $F(c|A)$.

It is assumed that an individual police officer allocates search efforts so as to maximize the total number of convictions minus a cost of searching cars. Each officer can choose to search motorists of any type (c, r) at a marginal cost of t_r. Normalize the benefit of each arrest to equal 1, so that the cost is scaled as a fraction of the benefit (assume $t_W, t_A \in (0, 1)$).

In deciding whether to carry contraband, motorists consider the probability of being searched and the penalty if they were to be caught. If they do not carry, their payoff is zero whether or not the car is searched. If they do carry, their payoff is $v(c, r) + x$ if not searched and $-j(c, r)$ if they are searched, where both $j(c, r)$ and $v(c, r)$ are positive. x represents private information of the motorists about their own benefit from carrying contraband. $j(c, r)$ can be interpreted as the cost of being convicted.[11]

Denote by $\gamma(c, r)$ the probability that the police officer searches a motorist of type c, r. The expected payoff to a motorist from carrying contraband is

$$\gamma(c, r)[-j(c, r)] + [1 - \gamma(c, r)][v(c, r) + x]. \tag{1}$$

Given $\gamma(c, r)$, the motorist chooses to carry contraband if this expression is greater than zero. Given a certain probability of being searched, motorists with a high realized value of x strictly prefer to carry drugs and those with small values strictly prefer not to carry. However, search strategies can only be conditioned on c and r, since x is not directly observed by the police. Let G denote the event that the motorist searched is found guilty of carrying contraband, and denote the probability that a motorist of type c, r carries contraband by $P(G|c, r)$.[12]

Assume that police offers decide on the search probability $\gamma(c, r)$ (the probability of

[11] If there were discrimination in the court system leading to higher penalities for minority drivers found with contraband, this could be thought of as operating through $j(c, r)$. KPT do not test for this type of discrimination.

[12] We do not allow for the possibility of false accusation by police or planting of evidence, as considered in Donohue and Levitt (1998).

searching each motorist of type c,r) to maximize the number of successful searches, net of costs. He/she solves

$$\max_{\gamma(c,W),\gamma(c,A)} \sum_{r=W,A} \int \left[P\left(G|c,r\right) - t_r\right] \gamma\left(c,r\right) f\left(c|r\right) dc,$$

taking as given $P\left(G|c,r\right)$. The term $P\left(G|c,r\right) - t_r$ represents the expected profit from searching a motorist of type c,r. If $P\left(G|c,r\right) - t_r > 0$ then optimizing behavior implies $\gamma\left(c,r\right) = 1$, i.e., always search motorists of type c and r. If $P\left(G|c,r\right) = t_r$ then the police officer is willing to randomize over whether or not to search type c,r.

KPT introduce racial bias into this framework as a difference in the perceived cost of searching motorists of different races. That is, a police officer is said to be biased against race A (or equivalently, have a taste for discrimination) if $t_A < t_W$. If police officers have no taste for discrimination and yet the police officer chooses search probabilities that differ by race, then the equilibrium is said to exhibit statistical discrimination. Statistical discrimination is motivated out of efficiency considerations and not out of racial bias.[13]

KPT (2001) study the equilibrium implications of this model, for the case where officers are homogenous in their costs of search and motorists are heterogeneous in the benefits they derive from carrying contraband. The model implies the following equilibrium conditions, for all c

$$\begin{aligned} P^*\left(G|c,A\right) &= t_A \quad , \qquad P^*\left(G|c,W\right) = t_W \\ \gamma^*\left(c,A\right) &= \frac{v\left(c,A\right)}{\left[v\left(c,A\right) + j\left(c,A\right)\right]} \quad , \gamma^*\left(c,W\right) = \frac{v\left(c,W\right)}{\left[v\left(c,W\right) + j\left(c,W\right)\right]}, \end{aligned} \tag{2}$$

where * denotes equilibrium values.

Suppose now that $t_A = t_W = t$ (i.e. police officers are not biased). Then, for all c, guilt probabilities at equilibrium must be equal across races:

$$P^*\left(G|c,A\right) = t = P^*\left(G|c,W\right). \tag{3}$$

If guilt probabilities were not equalized, a police officer could do better by reallocating searches towards the group with the higher hit rate.

[13] Statistical discrimination is used in the same sense as in Arrow (1973).

An important observation is that equalization of guilt probabilities does not imply that equalization of search rates. The equilibrium search intensity may be higher for African Americans even in the absence of racial prejudice. This happens if $v(c,W)/[v(c,W)+j(c,W)] < v(c,A)/[v(c,A)+j(c,A)]$. That is, if the expected value of carrying drugs is higher or the cost of being convicted lower for black motorists, then the search rate on that group would have to be higher in order to equalize the guilt probability to that of whites.

Equation (3) is the basis for the outcomes based test proposed in KPT as a test for racial bias (a test for $t_W \neq t_A$). A key advantage of the test is that it is implementable even in the absence of complete data on c and on γ^*. It suffices to have data on the frequency of guilt by race conditional on being searched,

$$D(r) = \int P^*(G|c,r) \frac{\gamma^*(c,r) f(c|r)}{\int \gamma^*(s,r) f(s|r)\, ds} dc.$$

Using (3) to substitute for $P(G|c,r)$ we get the implication

$$D(W) = t = D(A), \tag{4}$$

which is one that KPT empirically test. In the model, there is nothing special about the characteristic "race." In fact, the analogue of (4) should hold for any observed characteristic on which police can condition their search decision. The model has the strong implication that the guilty rates should be equal for any set of observed conditioning variables.

The assumption that motorists respond to the probability of being searched is key to obtaining a test for prejudice that can be applied without data on all the characteristics that police may use in the search decision. If motorists did not react to the probability of being searched, testing for prejudice would require data on c. To see why, consider a model where the probability that a motorist with characteristic c and race r carries drugs does not depend on the actions of police, and the only optimizing agents are the police. Let $\pi(c,r)$ denote the probability (now, exogenously given) that a type c,r carries drugs and suppose that $\pi(c,r)$ is increasing in c. Then, it is optimal for police officers to choose two cutoffs k_W and k_A and to search any motorist of race r with a c greater than k_r. In the absence of prejudice, police will choose k_W and k_A so that the probability that types k_W, W and k_A, A are guilty equals the marginal cost t of searching motorists, as shown in Figure 1.

Without data on c, one cannot empirically identify the marginal motorists and so cannot test the equilibrium implications of this model, in the absence of strong assumptions on the shape of $\pi(c, r)$ and on the distribution of the unobservables.

How does endogenous response of motorists change Figure 1? All motorists above the cutoff k_r, knowing that they are going to be searched for sure, would respond by not carrying drugs. Motorists below the cutoff would similarly react by choosing to carry drugs.[14] Then, $\pi(c, r)$ would become *decreasing* in c, which would make suboptimal the policy of searching motorists above k_r. When $\pi(c, r)$ is determined endogenously, the only equilibrium is for $\pi(c, r)$ to equal t_r for all c. Thus, allowing for endogenous response of motorists to the probability of being searched eliminates the problem of having to identify the marginal motorists.

Extensions of the KPT Model A number of papers have explored extensions or variations of the KPT model. Antonovics and Knight (2004) raise concern that police heterogeneity is a potential threat to the validity of the outcomes based tests. They present evidence that police are more likely to search the vehicles of drivers of a different race. Persico and Todd (2004) generalize the KPT model to allow for police heterogeneity in costs of search and to allow for the possibility that drivers' characteristics are mutable, in the sense that drivers can adapt some of their characteristics to reduce the probability of being monitored by the police.[15] They find that the hit rate test is still valid under these extensions. Persico and Todd (2005) extend the basic KPT model to allow for imperfections in the monitoring technology, in particular that searches do not always uncover contraband.

[14] In this stylized model, we restrict attention to the population that has some positive value from carrying drugs. This is without loss of generality. We could easily add a fraction of motorists who police know would never want to carry drugs (even if they are never searched), and the police would therefore never search these motorists in equilibrium.

[15] For example, if drivers with sports cars are subject to high monitoring rates, an individual might choose to drive a different type of car.

Even with detectability rates that vary across groups, they find that the hit rate test can still be justified as a test for racial bias.[16]

Hernandez-Murillo and Knowles (2004) consider the question of how to test for racial bias with aggregated data that is contaminated by observations on nondiscretionary stops. In particular, the KPT model assumes that police searches are discretionary, whereas Hernandez-Murillo and Knowles analyze a dataset from Missouri that mixes discretionary and nondiscretionary stops. They derive tests for racial bias (inspired by the nonparametric bounding approach of Manski and Horowitz (1995)) that are robust to the contamination.

Dharmapala and Ross (2004) extend the KPT model by relaxing the assumption that police can search any motorist. That is, they impose a technological limitation on the search capacities of police, whereby police only observe motorists with some probability less than one. Because of this constraint, there can be some groups of motorists for whom the unobservability constraint leads them to carry contraband all the time. Police would like to search this group harder in order to equalize guilty rates across groups, but cannot due to the observability constraint. In equilibrium, such motorists will be searched whenever police observe them. Dhamapala and Ross (2004) demonstrate that if this type of motorist is distributed differently among different racial/ethnic groups, then the KPT hit rate test breaks down. Another contribution of Dharmapala and Ross (2004) is to consider the set of equilibria in a modified version of the KPT model in which there are offenses of varying levels severity and motorists sort over the level of severity.[17]

2.1.2 The model of Anwar and Fang (2004)

A limitation of the KPT model is that it assumes that police officers first see some characteristics of the motorist and then decide whether to search them. A more realistic set-up would assume that police see some information prior to the stop decision and then acquire information from interacting with the motorist. Anwar and Fang (2004) develop a frame-

[16] The extensions are developed in an application of the model to monitoring of passengers at airports.

[17] The analysis in KPT assumed that, conditional on having decided to break the law, the severity of a motorist's offense was exogenous.

work in which the officers' search decisions can depend on information that they acquire additional information after the stop. Their framework also allows for the possibility that police behavior varies by the race of the police officer. For example, white police may be biased against minority drivers and minority police biased against white drivers.[18]

Anwar and Fang's (2004) model is in the spirit of statistical discrimination models (e.g. Coate and Loury, 1993). It assumes that during the stop and prior to deciding whether to search the vehicle, the police officer observes a noisy but informative signal about whether the driver carries contraband. The signal is informative in the sense that guilty drivers (those carrying illegal contraband) are more likely than innocent drivers to generate suspicious signals, such as nervousness in answering the police officer's questions.[19] As in the KPT model, police officers are considered to be racially biased if their cost of search depends on the race of the motorist and the objective of officers is to maximize the number of successful searches.

Anwar and Fang (2004) develop two tests. The first is a test for whether police officers of different races use different search criteria when dealing with motorists of the same race, which would indicate heterogeneity in police search behavior. The test is based on the observation that if officers do not differ in search costs, then the search rates and success rates of different groups of officers should on average be the same. The second test they develop is a test for racial predjudice. They note that if officers' search costs depend only on the officer's race but not on the race of the driver (i.e. are nondiscriminatory), then the ranking of search rates for white and minority police officers should not depend on the race of the driver. Moreover, the ranking of the success rates (guilty rates) should also not depend on the race of the driver. For example, suppose minority officers have lower search costs than white officers, but that search costs of all officers do not vary depending on the race of the driver. In equilibrium, we would expect minority officers to have higher search

[18]Persico and Todd (2004) also allow for police heterogeneity in the degree of bias, but do not allow for the sign of the bias to differ for individual officers. That is, they do not allow for the possibility that some police may be biased while others may exhibit favortism, the case considered in Anwar and Fang (2004).

[19]It is assumed that the drivers themselves do not know at the time of deciding whether to carry contraband whether they will generate a suspicious signal, only the probability that they will generate one.

rates and lower average success rates than white officers, regardless of the type of driver they are searching. An implication of racially unbiased policing is, therefore, that the ranking of officers' search rates by race of the officer be invariant to the race of the driver. Anwar and Fang's (2004) test can uncover whether at least one of the groups of officers (e.g. white or minority officers) is searching in a racially biased manner, although it cannot distinguish which group is biased.[20]

2.2 Models of Crime Minimization

The previous class of models assumed that individual officers choose search strategies that maximize succesful search rates, or so-called hit rates. As noted above, a hit rate objective function would be a reasonable approximation to police behavior if officers are rewarded on the basis of criminals apprehended, something that is easily observed. As demonstrated in Persico (2002), however, the hit rate objective function does not minimize the aggregate crime rate.

An alternative modeling framework assumes that there is a centralized authority, a police chief, say, who can direct officers to focus their searches on particular subgroups. In such a model, the hit rate test fails as a test of the unbiasedness of the police chief, because, in the equilibrium of such a model, an unbiased police chief will allocate searches to equate the deterrence effect, and not the hit rates, across groups. Crime deterred is unobserved, making it difficult to devise a test for whether the deterrence effect is being equalized across groups. While one could conceivably introduce racial bias in a crime minimization model in the same way as in the hit rate maximization models–as a difference in the costs of searching different types of motorists–there is currently no empirically implementable test for racial bias in such a model.[21]

[20] Anwar and Fang's (2004) testing procedure requires that officers of different races face the same population of motorists. In datasets where such an assumption is not satisfied, they propose a resampling procedure to implement their tests.

[21] Eeckhout, Persico and Todd (2003) study optimal monitoring strategies for police assuming that the objective is to minimize crime. They show that in some cases it can be optimal to randomly subject even identical motorists to different levels of monitoring. This could be considered random profiling, in that motorists are randomly divided into different groups and are subjected to different levels of monitoring.

2.3 Imposing a Race-Blind Constraint on Police Behavior

Persico (2002) studies the effects of constraints on police behaviors, within a model where police maximize hit rates, but the assumed socially efficient objective is to minimize the aggregate crime rate. He shows that imposing a "race-blind" constraint on police search behavior does not necessarily entail any loss in efficiency and can in some cases increase efficiency. That is, forcing the police to be racially fair, by not allowing search probabilities to depend on race, can sometimes lead to a lower cost way of achieving a given crime rate. This somewhat surprising result follows because search strategies that aim to maximize arrests do not take into account the deterrence value that arrests have on different groups. The incentive scheme that minimizes the crime rate would place a higher value on arresting motorists of the race that is more likely to be deterred by the prospect of being arrested.

To see how a fairness constraint can increase efficiency, suppose, for example, that whites were more numerous in the population and that they were less likely to carry drugs than blacks, at a given rate of search. In the absence of any constraints on search behavior, police would search blacks at a higher rate so as to equalize the hit rates across groups. Under a fairness constraint, however, the two groups are pooled and everyone experiences the same probability of being searched. In equilibrium, the overall carrying rate will remain the same as in the unconstrained equilibrium and will equal the cost of search. However, equalizing search rates by race leads to an increase in the black carrying rate, and an offsetting decrease in the white rate. If whites are deterred by a relatively small increase in the probability of search and they are relatively numerous in the population, then it is possible to achieve the same overall carrying rate at a lower search cost in the constrained equilibrium than in the unconstrainted equilibrium. Persico (2002) finds that whether imposing a fairness constraint leads to an increase in efficiency (defined as a decrease in the crime rate for the same cost) crucially depends on the proportion of blacks in the population relative to the cost of search.

For further consideration of how to incorporate efficiency and equity considerations into an assessment of racial profiling as a public policy, see Durlauf (2004) and Risse and Zeckhauser (2004). Also, see Dominitz (2003) for discussion of the statistical relationship between various outcomes that could be considered when formulating public policy, such as search rates, find rates, thoroughness of search, rates of detention of the innocent, and rates of apprehension of the guilty.

3 Empirical Evidence

As noted above, models that assume that police maximize hit rates and that motorists take into account the probability of being caught when deciding whether to break the law lead to simple ways of testing for racial bias. The KPT hit rate test is based on a comparison of hit rates across different groups of motorists. The Anwar-Fang rank test evaluates whether ranks of search rates by type of officer are invariant to the race of the driver. The simple hit rate test can be performed using the type of data that is conventionally available. Implementing the rank test requires information on characteristics of the police officer making the stop, which is only rarely available.

Table 1 summarizes findings from 16 different city-level and state-level racial profiling studies/reports, in which the hit rates by race/ethinicity are reported. The table displays what appears to be an empirical regularity: there is not a large disparity in hit rates for black and white drivers, especially when compared with the disparity in search/stop rates. This regularity is puzzling in the context of a crime-minimizing police chief, but not in light of a simple hit rate maximization model, which offers a rationale for the equalization of hit rates across races, namely (a) that police are allocating searches in a way that maximizes efficiency in catching criminals and (b) that police departments, on average, are not afflicted by widespread bias against African Americans. In Table 1, the hit rates for Hispanics are in many cases notably lower than that of whites or blacks, which is suggestive of bias against Hispanics. Whether in fact this is really the case can only be ascertained with more work on the police data sets that are recently becoming available.

Anwar and Fang (2004) apply their proposed rank tests to a dataset on highway stops and searches collected by the Florida Highway Patrol, which records the race of the police officer making the stop. The data reveal search patterns that differ significantly by race of the officer, with white officers searching the highest percentage of motorists stopped. Despite the differences in police search behavior, however, Anwar and Fang's (2004) rank test for racial bias does not reject the null hypothesis of no relative racial predjudice between black and white officers.[22]

4 Summary and Conclusions

Recent advances in the economic literature have led to a better appreciation of the assumptions that underlie different approaches to testing for racial/ethnic bias in policing. Simple benchmark tests for discrimination only uncover whether a disparity exists but do not reveal the motivation for the disparity, which plays a key role in racial profiling invsetigations. Assuming a hit rate objective function and strategic behavior on the part of both police officers and motorists, leads to a simple, empirically implementable outcomes-based test for whether the disparity is due to racial bias. Such a model can potentially explain an observed empirical regularity in many police datasets that there is little disparity in hit rates for black and white drivers, despite large disparities in search/stop rates.

An alternative modeling framework is one in which a police chief allocates resources so as to minimize the aggregate crime rate. An implication of unbiased policing in that type of model is that the deterrence effect be equated across different groups of motorists. This implication is difficult to test empirically.

Even if an outcomes-based test concludes that racial disparities in search rates do not reflect racial bias, there is still the question of whether statistical discrimination is justified. Statistical discrimination may be considered unfair, because drivers experience different probabilities of being searched, depending on their race. An intriquing aspect of Persico's (2002) findings is that when police are maximizing a second-best objective (hit rates), imposing a race-blind constraint can come closer to a first-best objective (minimizing overall crime). In practice, though, it may be difficult to ask police officers to simply ignore race in their

[22] They advise caution in interpreting these results, as the test is only informative about relative racial bias and cannot rule out the possibility that all police (of both races) might be biased.

decision-making. More race-neutral policing might instead be achieved by giving police differential rewards for hit rates on white verses minority drivers. Designing optimal incentive schemes for achieving a particular objective is a current area of research.

Finally, the economic literature on racial profiling is relatively nascent and there are many ways of extending existing models to make them more realistic. For example, none of the existing models specify how a police chief allocates police officers to patrol particular areas. For this reason, existing tests are usually applied to data on highway searches where selective allocation of officers to monitor certain populations is less of an issue. Moreover, existing theory has mainly been developed for discretionary stops, but a major proportion of police stops and searches are triggered by events that make a search of the vehicle mandatory.

References

[1] Antonovics, K. L. and Knight, Brian G. (2004): "New Look at Racial Profiling: Evidence from the Boston Police Department," National Bureau of Economic Research Working Paper # W10634.

[2] Anwar, Shamena and Hanming Fang (2004) "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," manuscript forthcoming in *American Economic Review*.

[3] Arrow, Kenneth "The Theory of Discrimination." In *Discrimination in Labor Markets*, edited by O. Ashenfelter and A. Rees, Princeton University Press, Princeton NJ, 1973.

[4] Ayres, Ian (2002): "Outcomes Tests of Racial Disparities in Police Practices," http://islandia.law.yale.edu/ayers/jrp.pdf.

[5] Becker, Gary *The Economics of Discrimination*, Chicago University Press 1957.

[6] Becker, Gary "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (March-April 1968): 169-217 .

[7] Border, Kim C; Sobel, Joel. "Samurai Accountant: A Theory of Auditing and Plunder." *Review of Economic Studies*. Vol. 54 (4) (October 1987): 525-40.

[8] Carter, D., Katz-Bannister, A., Schafter, J. (2002): "Analysis of the Lansing Police Department MATS Data: An Eighteen Month Status Report," Report submitted to Chief Mark Alley, Lansing Police Department.

[9] Cohen-Vogel, Dan and Doss, Brian (2002): "Vehicle Stops and Race: A Study and Report in Response to Public Chapter 910 of 2000" Report to Comptroller of the Treasury.

[10] Cordner, G., Williams, B. and Velasco, A. (2002): "Vehicle Stops in San Diego: 2001," San Diego State University (Velasco).

[11] Dharmapalam Dhammika and Stephen L. Ross (2004) "Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence", Contributions to Economic Analysis & Policy: Vol. 3: No. 1, Article 12. http://www.bepress.com/bejeap/contributions/vol3/iss1/art12

[12] Dominitz, Jeffrey (2003): "How Do the Laws of Probability Constrain Legislative and Judicial Efforts to Stop Racial Profiling," in *American Law and Economics Review*, Vol. 5, No. 2, p. 412-432.

[13] Donohue, John. Expert witness testimony in the case *Chavez v. Illinois State Police* 1999.

[14] Donohue, John and Levitt, Steven. "The Impact of Race on Policing, Arrest Patterns, and Crime," *Journal of Law and Economics*, University of Chicago Press, Vol. 4412, pp. 367-394.

[15] Donohue, John (2005): "The Law and Economics of Antidiscrimination Law," Public Law & Legal Theory

[16] Durlauf, Steven (2004): "Racial Profiling as a Public Policy Question: Efficiency, Equity, and Ambiguity," working paper, University of Wisconsin.

[17] Eeckhout, Jan, Persico, Nicola and Petra Todd (2003): "A Theory of Random Crackdowns," manuscript, University of Pennsylvania.

[18] Epstein, Richard. *Forbidden Grounds: The Case Against Employment Discrimination Laws*. Cambridge, Ma.: Harvard University Press 1992.

[19] Farrell, A., McDevitt, D. J., Cronin, Shea, Pierce, Erica (2003): "Rhode Island Traffic Stop Statistics Act Final Report," Northeastern University, Institute on Race and Justice.

[20] Greenwald, H. (2003): "Police Vehicle Stops in Sacramento, CA: Second of three reports," School of Policy, Planning and Development, University of Southern California.

[21] Gross, S. R. and K. Barnes (2002): "Road Work: Racial Profiling and Drug Interdiction on the Highway," *Michigan Law Review*, 101 (3): 653-754.

[22] Harcourt, Bernard (2003): "Rethinking Racial Profiling: A Critique of the Economics, Civil Liberties, and Constitutional Literature and of Criminal Profiling More Generally," Public Law and Legal Theory Working Paper No. 51, University of Chicago.

[23] Kennedy, Randall *Race, Crime and the Law*. New York: Pantheon Books 1997.

[24] Knowles, J., Persico, N. and P. Todd (2001): "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 109 (1).

[25] Lovrich, N., Gaffney, M., Mosher, C., Pickerill, M. and Smith, M. (2003): "WSP Traffic Stop Data Analysis Project: Data Analysis Project Report," Washington State University.

[26] Nixon, J. (2003): Executive Summary on 2003 Missouri traffic stops, http://www.ago.mo.gov/racialprofiling/2003/racialprofiling2003.htm, copy on file with the authors.

[27] Persico, Nicola (2002): “Racial Profiling, Fairness and Effectiveness of Policing," in *American Economic Review*, 92(5), 1472-97.

[28] Persico, Nicola and Todd, Petra. “Using Hit Rates Tests to Test for Racial Bias in Law Enforcement: Vehicle Searches in Wichita.” *National Bureau of Economic Research* (Cambridge, MA) Working Paper No. 10947, December 2004.

[29] Persico, Nicola and Petra Todd (2005): “Passenger Profiling, Imperfect Screening and Airport Security,” *American Economic Review*, May.

[30] Persico, Nicola and David A. Castleman (2004): "Detecting Bias: Using Statistical Evidence to Establish Intentional Discrimination in Racial Profiling Cases," manuscript, University of Pennsylvania.

[31] Reinganum, Jennifer and Louis Wilde “Equilibrium Verification and Reporting Policies in a Model of Tax Compliance.” *International Economic Review* 27 (October 1986): 739-760.

[32] Risse, Mathias and Richard Zeckhauser (2004): "Racial Profiling," *Philosophy and Public Affairs*, 32(2), pp. 31-70.

[33] Ross, Stephen L. “Mortgage Lending Discrimination and Racial Differences in Loan Default.” *Journal of Housing Research* 7(1) (1996): 117-26.

[34] Scotchmer, Suzanne. “Audit Classes and Tax Enforcement Policy.” *American Economic Review* 77 (2) (May 1987): 229-33.

[35] Sherman, Lawrence W. (1990): "Police Crackdowns: Initial and Residual Deterrence," in M. Tonry and N. Morris (eds.), *Crime and Justice: A Review of Research*, vol. 12 (1990), Chicago: University of Chicago Press.

[36] Smith, W., Davison, E., Zingraff, M., Rice, K. and Bissler, D. (2004): “An Empirical Investigation of the Possible Presence and Extent of Arbitrary Profiling in the Charlotte-

Mecklenburg Police Department," Final Report to Charlotte-Mecklenburg Police Department.

[37] Spitzer, Eliot (1999): *A Report to the People of the State of New York from the Office of the Attorney General*. Civil Rights Bureau, New York, New York.

[38] Stigler, George "The Optimum Enforcement of Laws." *Journal of Political Economy*, 78 (May-June 1970): 526-536.

[39] Thomas, D. and Hansen, R. (2004): "Second Annual Report, Denver Police Department, Contact Card Data Analysis."

[40] U.S. General Accounting Office. *Customs Service: better targeting of airline passengers for personal searches could produce better results* GAO/GGD-00-38U.S.Washington DC: U.S. Government Printing Office, March 2000.

[41] Verniero, P. and Zoubek, P. (1999): *Interim Report of the state police review team regarding allegations of racial profiling*, Trenton: NJ, Attorney Generals Office.