

Generalizing the Hit Rates Test for Racial
Bias in Law Enforcement, With an
Application to Vehicle Searches in Wichita¹

Nicola Persico and Petra Todd

University of Pennsylvania²

June 28, 2006

¹[To the attention of the printer. Pagehead: Generalizing the Hit Rates Test.]
The authors are grateful to the NSF for financial support. The first author also gratefully acknowledges support in the form of a Sloan Research Fellowship. We thank Charles Manski, Steven Machin, and an anonymous referee for helpful comments.

²Department of Economics, 3718 Locust Walk, Philadelphia, PA 19104. Emails: persico@ssc.upenn.edu, and petra@athena.sas.upenn.edu.

Abstract

This paper considers the use of outcomes-based tests for detecting racial bias in the context of police searches of motor vehicles. We characterize the police and motorist decision problems in a game theoretic framework, where police encounter motorists and decide whether to search them and motorists decide whether to carry contraband. Our modeling framework generalizes that of Knowles, Persico and Todd (2001) (KPT) by allowing for police heterogeneity in costs of search and in tastes for discrimination and for motorist heterogeneity in the costs and benefits from crime. We also consider the possibility that drivers' characteristics are endogenously determined in that drivers can alter their characteristics to reduce the probability of being monitored. We establish the properties of the equilibrium in these more general settings and show that the outcomes-based test proposed in KPT can still be applied. After developing the theory, we apply the tests to data on police searches of motor vehicles gathered by the Wichita police department. The empirical findings are consistent with the notion that police in Wichita choose their search strategies to maximize successful searches. We also summarize evidence on these tests when applied to other police datasets.

JEL Numbers: J70, K42

Keywords: Racial Profiling, Crime, Police, Wichita

In his first address to Congress, President George W. Bush reported directing his Attorney General "to develop specific recommendations to end racial profiling. It's wrong, and we will end it in America."¹ By the term "racial profiling," he was referring to a presumed unlawful use of race or ethnicity in police interdiction. Over the last ten years, numerous lawsuits have been brought against US city police departments alleging racially biased law enforcement practices.² Partly in response to this litigation, many police departments now routinely collect information on the demographic characteristics of the individuals that they subject to stops and searches and on the outcomes of these encounters.

A common pattern found in police datasets is that blacks and Hispanics tend to be overrepresented in police traffic stops and searches. This pattern raises concerns as to whether the disparities reflect police bias or whether they are the byproduct of goal-oriented enforcement by unbiased police. Various tests have been proposed in the literature to assess whether police behavior is racially biased. The simplest tests are so-called *benchmarking* tests, which compare the racial/ethnic composition of the population monitored by the police against population benchmarks, for example, the racial composition of the general population in a given area. A more sophisticated version of the benchmark test examines whether race/ethnicity predicts whether an individual is subject to monitoring, after taking account other characteristics that the police are permitted to use as potential indicators of criminality. One drawback of this type of test is that the result usually depends on the

particular set of characteristics used. Another drawback is that the test cannot distinguish whether police subject certain groups to higher rates of monitoring because of bias or because the groups are known to have higher levels of criminality. Despite these shortcomings, benchmarking tests are commonly used in practice.

In previous work (Knowles, Persico and Todd, 2001, henceforth KPT), an alternative *outcomes-based* test was proposed for distinguishing the motivation for differential monitoring rates. The test was derived from a rational choice model of police and motorist behavior. The model yielded the prediction that unbiased monitoring should result in the equalization of the expected *hit rate*, the rate at which contraband is seized, across all observable categories of drivers. KPT applied outcomes-based tests to a dataset gathered by the Maryland Police Department and found that the hit rates were equal across black and white drivers, which they interpreted as an indication that the racial disparity in monitoring rates was not due to police bias. Moreover, the hit rates were found to be equal across most distinguishable categories of drivers, lending additional credibility to the model as a descriptive model of police and motorist behavior.

The KPT test has had considerable impact at the policy level (see Fridell 2004, ch. 11). At the academic level, a number of papers have explored extensions or variations of the KPT model, some examining the degree to which modifying the theoretical model in various ways (discussed in detail in the next section) would preserve the test for racial bias.³ Anwar and

Fang (2006), in particular, develop a theoretical model in which the KPT test would not be a valid test for police bias. In their empirical analysis of Florida Highway Patrol stops data, the equalization of hit rates across racial and other characteristics is not observed. In another study, Antonovics and Knight (2004) emphasize police heterogeneity in tastes for discrimination as a potential threat to the validity of outcomes-based tests. This recent literature seems to call into question whether the hit rate equalization observed in the Maryland data is atypical and also the usefulness of outcomes based tests in more general settings.

Motivated by this recent research, this paper contributes to the debate along three dimensions. Firstly, in Section 1, it generalizes the police-motorist model to incorporate police heterogeneity in the degree of racial bias and in costs of search as well as motorist heterogeneity in the costs and benefits from committing crimes. In the Appendix, we also extend the analysis to the case where driver's characteristics are endogenously determined in the sense that drivers can adapt some of their characteristics to reduce the probability of being subjected to police monitoring.⁴ We establish the properties of the equilibrium in these more general settings, drawing on recent results in the game theory literature on so-called large crowding games. Outcomes-based tests for discrimination are shown to still be applicable in these more general environments.

The second contribution of this paper is to present new evidence on the application of outcomes-based tests to police datasets. In Section 3, we

describe results based on a dataset obtained from the Wichita Police Department containing information on all vehicle stops and searches that took place in Wichita, Kansas during the first 9 months of 2001. An initial examination of these data by Withrow (2004) found a statistically significant disparity between the proportions of black drivers stopped (21%) and their representation in the Wichita population (11.4%). There is less of a disparity for Hispanics, who represent 9.2% of the stops and 9.6% of the Wichita population.⁵ Applying benchmarking testing criteria, these disparities would be taken as indicative of police bias. When we apply an outcomes-based test, however, we find that the data do not reject the hypothesis that hit rates are equal for drivers of all races/ethnicities. According to the KPT model, this result is broadly consistent with the disparity in search rates being attributable to statistical discrimination rather than bias. Moreover, we find that the hit rates do not differ by gender, by age, and differ only slightly by time of day of the search. The KPT model offers a simple rationale for the widespread equalization of hit rates, and thus we view this empirical finding as supportive of the descriptive validity of the model.

The third contribution relates to the broad contours of the phenomenon of racial disparities in policing. In Section 4, we summarize related evidence from many different police dataset pertaining to cities and states throughout the United States. We observe that the equalization of hit rates for whites and African Americans, that was found in both the Wichita and Maryland datasets, seems to be a general feature of other traffic stop data sets. We

discuss the extent to which this observation can be seen as a validation of the KPT model. Our discussion highlights the importance of the posited objective function of police officers in justifying particular tests for discrimination. Specifically, the assumption that police officers maximize hit rates is crucial for our outcomes-based test but may lead to socially suboptimal police behavior. To the extent that the assumption is valid, our summary suggests that police departments are not, on average, afflicted by strong and widespread bias, at least with regard to their vehicle search activities. Section 5 concludes.

This paper is complementary to several recent papers that also aim to relax the modeling assumptions of KPT. Three particularly relevant papers are Dharmapala and Ross (2004), Anwar and Fang (2006), and Antonovics and Knight (2004). Although each paper studies a somewhat different model, the first two papers share a basic distinguishing feature. They both assume that it is infeasible for the police to perfectly deter crime in a given subgroup of the population and show that, under this assumption, the hit rates test is not necessarily valid. In addition, Anwar and Fang (2006) provide a test for "differential bias" within different subgroups of the police. They use their theory to test whether there is a differential bias between black and white Florida police officers, and cannot reject the hypothesis of no differential bias. Antonovics and Knight (2004) is similar in spirit to Anwar and Fang (2006) in that they also look at differential bias between black and white officers. Of relevance to this paper is the concern they raise that the hit

rate test might break down in the presence of heterogeneity in police search costs. This paper demonstrates that the test is valid even in the presence of this type of heterogeneity. The reason for the difference in the findings is that Antonovics and Knight (2004) implicitly assume that motorists are randomly matched with a small set of police officers. We will return to this point in Section 4.

1 The Model of Motorist and Police Behavior

We next describe our model of police and motorist behavior, which incorporates potential police heterogeneity in intensity of racial bias and in costs of searching as well as motorist heterogeneity in the benefits and costs from committing a crime. In the Appendix, we further generalize the model to allow motorists to delegate the crime to others, or to disguise their appearance by posing as a member of another group that is monitored at a lower rate. Our main theoretical finding is that outcome based tests are still valid in these more general environments.

To describe the model, we first define some notation. Let r denote the race of the motorist, which is assumed to be observable by the police. Without loss of generality, assume that there are motorists of two races, either African American (A) or white (W). Other characteristics that are observable by the police are represented by the number $c \in \{1, \dots, C\}$. These characteristics might represent such things as type of car and age or gender of motorist.

From the police viewpoint, a motorist is characterized by two variables, r and c . Let $N_{r,c}$ denote the number of motorists belong to group (r, c) .

We assume that the police can distinguish between motorist groups (r, c) , but cannot detect motorist heterogeneity within (r, c) groups. Two sources of unobserved motorist heterogeneity within groups are the values to an individual of committing a crime and the costs of being detected, which may vary, for example, due to foregone earnings costs. Let v represent the value of committing a crime. If the crime is detected, the payoff to the motorist is $v - j$, where j captures the cost of being detected. We allow v and j to vary across individuals within a (r, c) group and denote the joint conditional distribution of v and j by cdf $F_{r,c}(v, j)$. We assume that this cdf's has no atoms, a property that is important in establishing uniqueness of the equilibrium as described below.

A motorist in group (r, c) makes a binary decision: commit a crime or not. In the Appendix we extend the analysis to include the case in which a motorist can delegate the crime to a member of another group at a cost, which is a case that has not been treated in the existing literature.

Just as motorists may differ in their costs and benefits, we allow police officers to be heterogeneous in three respects: their search capacity, their per-search cost, and their racial bias. We assume that there is a mass P of police officers, indexed by $p \in [0, P]$. Each police officer p is endowed with a search capacity of S_p and a per-search cost t_p . If a search of a motor vehicle does not yield any evidence of crime (such as contraband, illegal drugs, or

weapons), then we term the search unsuccessful and assume that the police officer incurred the cost of search without any benefit. We introduce the potential for police bias by allowing the benefit that the police derives from a successful search to depend on the race of the motorist. Suppose the benefit to a police officer p of finding a criminal of race W is y_p^W and the benefit of finding criminal of race A is $y_p^A = y_p^W + B(p)$. We say that police are biased against African American motorists if $B(p) > 0$ for all p , against whites if $B(p) < 0$ for all p , and unbiased if $B(p) = 0$ for all p . If no search is conducted, there is a zero payoff. As described, this setup accommodates police heterogeneity in intensity of bias. However, we rule out environments in which $B(p)$ changes sign as p varies, i.e., where some policemen are biased against whites and some are biased against African Americans. Below, we propose a test for inferring the sign of $B(p)$.

2 Equilibrium Analysis

A member of group (r, c) with given v, j , who commits a crime and expects σ members of his group to be searched receives an expected payoff

$$u_{r,c}(v, j, \sigma) = v - j \cdot \frac{\sigma}{N_{r,c}}.$$

When this payoff exceeds zero, the individual will choose to commit a crime. Let $K_{r,c}(v, j, \sigma)$ be an indicator function that equals 1 if the individual

chooses to commit a crime. The fraction of motorists within each group (r, c) who commit a crime is given by

$$K^{r,c}(\sigma) = \int K_{r,c}(v, j, \sigma) dF_{r,c}(v, j).$$

The function $K^{r,c}(\sigma)$ summarizes the crime rate in group (r, c) when the police search that group with intensity σ . One can think of this function as a response function or as the supply of crime.

Denote by $S_p(r, c)$ the number of searches that officer p decides to devote to group (r, c) . The total number of searches of members of group (r, c) is obtained by aggregating the behavior of all police officers:

$$S(r, c) = \int_0^P S_p(r, c) dp.$$

Officer p 's expected payoff is the sum of the expected payoffs of all his searches, given by

$$(1) \quad \sum_{r,c} S_p(r, c) [y_p^r \cdot K^{r,c}(S(r, c)) - t_p],$$

which depends on the officers perceived benefit from apprehending someone of race r (y_p^r) as well as the officer's costs of search (t_p).

2.1 *Existence and Uniqueness of Equilibrium*

Equation (1) represents the payoff function for police officers. We can think of this expression as a payoff function for a game that is played among officers. The game has a continuum of players and finite action sets. Moreover, the game is *anonymous*, in the technical sense that each player's payoff only depends on his own strategy (the vector $[S_p(r, c)]_{r,c}$) and on the aggregate response of the other players (the vector $[S(r, c)]_{r,c}$). Schmeidler (1973) established the existence of an equilibrium for games of this type. For Schmeidler's theorem to apply, two conditions must be met. First, the payoff function in equation (1) must be continuous in $[S(r, c)]_{r,c}$, which means that each function $K^{r,c}(\cdot)$ must be continuous. This is the case since by assumption the cdf's $F_{r,c}(v, j)$ have no atoms.

The second condition requires that for each pair of actions, the set of types p that strictly prefer one action to the other is measurable. In our model, this means that the set of police that prefer to search group (r, c) rather than group (r', c') must be measurable. Formally, for every κ, κ', r, r' the set of p 's such that

$$y_p^r \kappa - t_p > y_p^{r'} \kappa' - t_p$$

must be measurable. Because, in our model, "not search" is also an action, we also need that for every κ, r the set of p 's such that

$$y_p^r \kappa - t_p > 0$$

is measurable. These conditions are equivalent to requiring that the benefit-ratio and benefit-cost ratio functions $y_p^r/y_p^{r'}$ and y_p^r/t_p be measurable, which they are because it is assumed that for each r , y_p^r and t_p are measurable. Then, we can apply the results from Schmeidler (1973) to obtain existence of equilibrium.

Uniqueness of equilibrium follows from existing results in the literature on large games. Because for every p the payoff function (1) is linear in $S_p(r, c)$, we can restrict attention, without loss of generality, to the set of strategies for officer p that allocate all of the available searches on just one group. Then, $S(r, c)$ represents the mass of officers that decide to devote their searches to group (r, c) . We note that each officer's payoff depends on the actions of the other officers only through the total number of officers that choose to search group (r, c) . This dependence is negative—the more officers search group (r, c) the less profitable such a search is, because the function $K^{r,c}(\cdot)$ is decreasing. This game is a *large crowding game* in the precise sense of Milchtaich (2000). Milchtaich (2000) shows that “generically” large crowding games have a unique equilibrium.⁶ Thus, Schmeidler (1973) guarantees existence of the equilibrium and Milchtaich (2000) gives uniqueness, resulting in the following theorem:

Theorem 1 *A Nash equilibrium exists. The equilibrium is generically unique.*

2.2 Characterization of the Equilibrium

Let $[S^*(r, c)]_{r,c}$ be a vector denoting the search intensities at the Nash equilibrium. Suppose groups (r, c) and (r', c') are searched in equilibrium. Then, there must be a p and a p' such that

$$\begin{aligned} y_p^r K^{r,c}(S^*(r, c)) - t_p &\geq y_{p'}^{r'} K^{r',c'}(S^*(r', c')) - t_{p'}, \\ y_{p'}^r K^{r,c}(S^*(r, c)) - t_{p'} &\leq y_p^{r'} K^{r',c'}(S^*(r', c')) - t_p. \end{aligned}$$

If $r = r'$, or if the police are unbiased then $y_p^r = y_{p'}^{r'}$ for all p 's, and so the two inequalities can only be simultaneously satisfied if

$$K^{r,c}(S^*(r, c)) = K^{r',c'}(S^*(r', c')).$$

If the police are biased against race r then $y_p^r > y_{p'}^{r'}$, and so the second inequality can only be satisfied if the crime rates are such that

$$K^{r,c}(S^*(r, c)) < K^{r',c'}(S^*(r', c')).$$

Note that in our model the hit rate, i.e., the likelihood that a search of group (r, c) yields a find, coincides with that group's crime rate $K^{r,c}$. Thus, the implications on the crime rate translate into testable implications on the hit rates. This observation yields the following theorem:

Theorem 2 *In the equilibrium, the hit rate is the same across all subgroups*

within a race that are distinguishable by police. Also, if the police are unbiased, then the hit rate is the same across races. If the police are biased against race r , the hit rate is lower in race r than in the other race.

This theorem provides the justification for the outcomes based test applied in the next section.

3 Empirical Results

3.1 Data Description

We now analyze data that were collected by the Wichita police department for the purpose of assessing whether officers engage in racially biased policing practices. The data contain information on every police/citizen encounter from January, 2001 to September, 2001, including vehicle, bicycle and pedestrian stops as well as traffic accident investigations. The data include demographic information on the race, ethnicity, gender and age of the person who has the contact with the police. In addition, there is information on time of day, on whether a search of the vehicle was conducted, on the rationale the officer gives for stopping/searching, on whether any contraband was found, and on the duration of the stop. There is also some limited information on the characteristics of the officer (rank and type of officer), and on the number of officers present at the incident.⁷

A key assumption of the model developed in the previous section is that

police choose whom to search so as to maximize successful searches. Police presumably have little discretion in cases where they pull over a driver because they have a warrant for the driver's arrest or when the search is incident to an arrest.⁸ Therefore, we limit our analysis sample to observations on police-motorist encounters where police have discretion over whether to initiate the search. Our analysis sample contains information on 2,396 vehicle searches.

Table 1a shows the racial/ethnic distribution of drivers involved in stops and searches and, for comparison, the percentage of each group in the Wichita population. The percentage of stops and searches involving blacks (21.45% and 32.65%, respectively) is significantly higher than the black fraction of the population (11.4%). The percentage of whites in stops (64.37%) and searches (50.81%) is lower than in the population (65.6%). Hispanics are stopped at a rate roughly the same as their percentage in the population, but searched at a slightly higher rate. Asians, Native Americans and other races constitute a small percentage of the stops and searches and of the population. To ensure samples of adequate size, we focus our empirical analysis on blacks, whites and Hispanics.

Table 1b gives the age distribution of individuals subject to stops and searches. Most stops and searches are for persons age 18-24. Additionally, most involve male motorists; in 66% of all stops and 80% of all searches, the driver is a male. Another pattern is that most searches (76%) are carried out at night, between the hours of 7pm to 8am.

[Insert Tables 1a and 1b about here]

Table 2 reveals the type of contraband found during these searches, by the race/ethnicity of driver. For each type, the table shows the percentage of drivers found with that particular type of the total drivers searched and found with any type.⁹

[Insert Table 2 about here]

The most common type of contraband seized is drugs/drug paraphernalia, followed by alcohol/tobacco, stolen property and firearms. Black and Hispanic drivers are significantly more likely to be found with drugs/drug paraphernalia, while white drivers are more likely to be found with alcohol/tobacco and with firearms. Table 3 summarizes the types of rationales that police officers give for conducting the search.¹⁰

[Insert Table 3 about here]

3.2 *Empirical Findings*

Our test for racial bias compares the probability of finding contraband across groups with different observed characteristic. The model described in the previous section has the strong implication that the hit rates should be equal

across all observable groups. Because all the characteristics in our dataset are discrete variables, we can test the hypothesis of equal guilt rates across groups nonparametrically using Pearson χ^2 tests. These tests compare the proportion of drivers found carrying contraband within cells defined by the conditioning variables with the proportion that would be expected under the null hypothesis of no association between the hit rate and the set of conditioning characteristics. For example, the test statistic for testing the hypothesis of no association between hit rate and race is

$$\sum_{r \in \mathcal{R}} \frac{(\hat{p}_r - p)^2}{p} \sim \chi^2(R - 1),$$

where R is the cardinality of the set of race categories, \mathcal{R} , and \hat{p}_r is the conditional on r estimated guilt proportion and p is the expected proportion under the null.

Table 4a-4g show the percentage of motorists found to be carrying contraband (the hit rates) for groups of motorists defined by their characteristics. As seen in Table 4a, the percentages are nearly equal for blacks and whites (22.03% and 22.69%) and slightly lower for Hispanics (18.87%). The Pearson chi-square test does not reject the hypothesis that the percentages are equal for all the race groups (p-value is 0.365), even though the sample sizes are relatively large. According to the model, this finding is consistent with no racial bias in police search behavior.

[Insert Tables 4a-g about here]

Table 4b breaks down the percentages by the age of the driver. We do not reject the hypothesis that hit rates are equal across all age groups. As shown in Table 4c, the tests also do not reject equality across race groups when the test is performed within age groups. In Table 4d, we examine the hit rates according to gender of the driver. As noted previously, the search rates are much higher for male than female drivers; however, the hit rates are roughly the same.

Tables 4f and 4g show the hit rates by time of day of the search. In this case, the hit rates are statistically significantly lower at daytime than at nighttime. Most of the searches are conducted at nighttime, so it seems that police search efforts are being concentrated at the time when hit rates are higher. The disparity in hit rates for nighttime versus daytime searches could plausibly be due to a higher cost of conducting search activities at night (for example, if officers who work at night need to be paid more). In Table 4g, we examine whether the hit rates differ by race after conditioning on time of day, and we cannot reject the hypothesis that they are equal.

Equality of hit rates is a key prediction of the theoretical model when the police are unbiased. Overall, the empirical results show that the hit rates are very similar across groups of motorists no matter how these groups are defined. The evidence is thus consistent with the notion that police in Wichita are searching blacks and Hispanics at higher rates relative to their proportion in the population in order to maximize the probability of finding contraband.

4 Discussion

The model described in Section 1 is one where individual officers choose search strategies that maximize the hit rates. Implicitly, it is assumed that individual officers can focus their searches on whatever subgroup (r, c) they choose. Alternatively, we might have written a model with a centralized authority, a police chief, say, whose goal is to minimize the aggregate crime rate and who can direct officers to focus their searches on particular subgroups. The goal of minimizing the crime rate is different from allowing individual police to maximize hit rates.¹¹ Intuitively, in order to catch criminals there has to be some crime. An objective function that maximizes hit rates does not give enough weight to deterrent effects of policing, because it gives no reward to the police officer from preventing a crime from being committed.

Crucially, in a model where the police chief can allocate interdiction without any constraints, the hit rate test is no longer necessarily an appropriate test of the unbiasedness of the police chief. 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. This argument suggests some boundaries for the applicability of the KPT model. For example, it may not apply well to city policing situations where the police chief can influence the search activities of the individual officers by allocating them to specific beats. On the other hand, allocating officers to specific beats may be ineffective if criminals are mobile, and can easily shift their activity to other

beats.

Another consideration in deciding whether the KPT model is a reasonable approximation to police behavior is that it is likely difficult for a police chief to verify that individual officers are engaging in search activities that deter crime. The amount crime deterred by the activities of individual officers is never observed but how many criminals they catch is observed. It may therefore be easier to reward officers on the basis of their performance record in catching criminals, which the hit rate objective function implicitly assumes. For this reason, we believe the model where police act as noncentralized, independent agents trying to catch criminals could be viewed as a second best objective that even a police chief might reasonably adopt.¹²

Table 5 summarizes findings from 16 different city-level and state-level racial profiling studies/reports, in which the hit rates by race/ethnicity are reported. The table displays what appears to be an empirical regularity, that there is not a large disparity in hit rates for black and white drivers, especially when compared with the disparity in search/stop rates (which is not reported in the table but is generally large).¹³ The combination of disparate search rates and similar hit rates would be hard to explain within a model of a crime-minimizing police chief. The KPT model provides a simple rationale for it, namely that individual police officers are allocating searches in a way that maximizes successful searches (and that police departments, on average, are not afflicted by widespread bias). Of course, the evidence in Table 5 is rough, and whether in fact this is really the case can only be ascertained with more

detailed examination of the police data sets that are recently being made available to researchers.¹⁴

[Insert Table 5 about here]

Another point worth emphasizing is that different assumptions about matching between officers and motorists might yield different conclusions. Antonovics and Knight (2004) implicitly assume that motorists are randomly matched with a small set of police officers. This type of friction in the “arbitraging process” that the police follows means that not all drivers in a high-risk category need be searched. Under these circumstances, differences in hit rates might persist even in the absence of racial bias. Whether this type of friction is sufficiently important empirically to materially affect the conclusion of the KPT test is unknown and warrants further exploration.

5 Conclusions

This paper considers the use of an outcomes-based test for detecting racial bias in the context of police searches of motor vehicles. It shows that the hit rates test for racial bias can be applied in a more general environment where police officers are heterogenous in tastes for discrimination and in costs of search and motorists are heterogeneous in the benefits and costs from criminal behavior. We used recent results from the game theoretical

literature on large crowding games to establish that an equilibrium exists in such a model and that it is unique and we characterized the properties of the equilibrium. The appendix to this paper considers the case where drivers' characteristics can be altered to reduce the probability of being monitored. Our key finding is that outcomes-based (or hit-rate) tests can be applied in these more environments that are more general than those considered in the existing racial profiling literature.

We apply the hit rates test to a dataset gathered by the Wichita Police department on all police-citizen encounters in 2001. In this dataset, the stop rates and the search rates clearly differ by driver characteristics. For example, blacks and Hispanics are stopped and searched at higher rates than would be expected given their representation in the Wichita population. Also, males are searched four times as often as females. When we examine hit rates, however, we find that the hit rates do not differ by race/ethnicity, by gender, or by age. Remarkably, the hit rates are stable across various groups of drivers, which is a key prediction of the theoretical model when police are not motivated by racial bias. Thus, our empirical findings are broadly consistent with the notion that individual police officers in Wichita choose their search strategies mainly so as to maximize efficiency in finding contraband.

The empirical results described in this study are in many ways similar to empirical results that have been documented in other studies and reports, as discussed in Section 4. It appears to be an emerging empirical regularity that there is not a large disparity in hit rates for black and white drivers, es-

pecially when compared with the disparity in search/stop rates.¹⁵ The KPT model offers a simple rationale for the widespread equalization of hit rates, and that is that police departments are, on average, not afflicted by widespread bias in this dimension of their enforcement activities. We are mindful, however, that other models may be consistent with the observed regularity and may deliver different implications concerning police bias. Further explorations with alternative theoretical frameworks and with new datasets would be useful to obtain a more comprehensive view of what type of model best explains the outcomes of police-motorist encounters.

Notes

¹Address delivered on February 27, 2001.

²Most of these lawsuits were initiated by either the American Civil Liberties Union (ACLU) or the US Department of Justice. See Durlauf (2004) for a discussion of some normative considerations that arise in the context of racially disparate police interdiction.

³See e.g. Antonovics and Knight (2004), Anwar and Fang (2006), Dharmapala and Ross (2004).

⁴For example, if drivers with sports cars are subject to high monitoring rates, an individual might choose to drive a different type of car, or may hire a “mule” to carry drugs.

⁵Population figures are based on US Census 2000 data.

⁶Here, generically means that the set of functions $K^{r,c}(\cdot)$ that give rise to a game with a unique equilibrium is a dense $G - \delta$ in the space of all matrices of decreasing functions. A set is $G - \delta$ if it is the intersection of countably many dense open sets. This is a topological notion of a “large” set.

⁷The original dataset also includes additional demographic information on the police officer, such as the gender and race of the officer, years of experience, and information on the location of the stop (the beat). The Wichita Police Department would not release this additional information to us. When Withrow (2004) analyzed the data with respect to these variables, he concluded that "enforcement patterns do not differ substantially or illogically with respect to any of the variables internal to the department (officer age, officer gender, officer race, officer experience, day, time or beat). Importantly, this suggests that the pattern of [racial] disparity may better be explained by variables external to the department."

⁸The Wichita Police Department requires that officers conduct a search pursuant to an arrest (Withrow, 2004).

⁹Columns do not necessarily sum to 100% because drivers can be found with more than

one type of contraband.

¹⁰In any given search, there may be multiple rationales, so the categories are not mutually exclusive.

¹¹This point is made in several papers, including Alexeev and Leitzel (2002), Harcourt (2004), Eeckhout, Persico and Todd (2003), Manski (forthcoming 2006), Dominitz and Knowles (2004), and Persico (2002).

¹²See Blume (2004) for a model of learning about unobserved racial characteristics in the context of labor market discrimination.

¹³This paper also found the hit rates for Hispanics to be statistically equal to the hit rates for whites and Blacks. A common finding in the literature, though, is that the hit rates tend to be lower for Hispanics.

¹⁴Further evidence would be provided if the equalization of hit rates were found to extend to characteristics other than race, especially characteristics for which police bias would be less plausible. Recent empirical work that makes use of the race of the police officers has the potential of advancing the debate on this front. See Antonovics and Knight (2004), Anwar and Fang (2006).

¹⁵A common finding, though, is that the hit rates tend to be lower for Hispanics.

References

- [1] Alexeev, M. and Leitzel J. (2002). ‘Racial Profiling’, Manuscript, Indiana University in Bloomington.
- [2] Antonovics, K. L. and Knight, B. G. (2004). ‘New Look at Racial Profiling: Evidence from the Boston Police Department’, NBER Working Paper No. w10634.
- [3] Anwar, S. and Fang, H. (2006). ‘An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence’, *American Economic Review*, vol. 96(1), pp. 127-151.
- [4] Ayres, I. (2002). ‘Outcomes Tests of Racial Disparities in Police Practices’, *Justice Research and Policy*, vol. 4. Available at <http://islandia.law.yale.edu/ayers/jrp.pdf>.
- [5] Blume, L. (2004). ‘The Dynamics of Statistical Discrimination’, Manuscript, Cornell University. IN THIS ISSUE OF EJ
- [6] Bunzel, H. and Marcoul, Ph. (2003). ‘Can Racially Unbiased Police Perpetuate Long Run Discrimination?’, Manuscript, Iowa State University.
- [7] Carter, D., Katz-Bannister, A. and Schafter, J. (2002). ‘Analysis of the Lansing Police Department MATS Data: An Eighteen Month Status Report’, Report to Chief Mark Alley, Lansing Police Department.

- [8] Cohen-Vogel, D. and Doss, B. (2002). ‘Vehicle Stops and Race: A Study and Report in Response to Public Chapter 910 of 2000’, Report to the Comptroller of the Treasury, TN.
- [9] Cordner, G., Williams, B. and Velasco, A. (2002). ‘Vehicle Stops in San Diego: 2001’, Manuscript, San Diego State University.
- [10] Dominitz, J. and Knowles, J. (2004). ‘Crime Minimization and Racial Bias: What Can We Learn From Police Search Data?’, Manuscript, University of Pennsylvania. IN THIS ISSUE OF EJ
- [11] Durlauf, S. (2004). ‘Assessing Racial Profiling’, Manuscript, University of Wisconsin. IN THIS ISSUE OF EJ
- [12] Dharmapalam D. and Ross, S. L. (2004). ‘Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence’, *Contributions to Economic Analysis & Policy*, vol. 3(1). Available at <http://www.bepress.com/bejeap/contributions/vol3/iss1/art12>.
- [13] Eeckhout, J., Persico, N. and Todd, P. (2005). ‘A Rational Theory of Random Crackdowns’, Manuscript, University of Pennsylvania.
- [14] Farrell, A., McDevitt, D. J., Cronin, S., Pierce, E. (2003). ‘Rhode Island Traffic Stop Statistics Act Final Report’, Northeastern University, Institute on Race and Justice.
- [15] Fridell, L. (2004). *By the Numbers: A Guide for Analyzing Race Data from Vehicle Stops*, Police Executive Research Forum, Washington, DC

- [16] Greenwald, H. (2003). 'Police Vehicle Stops in Sacramento, CA: Second of Three Reports', School of Policy, Planning and Development, University of Southern California.
- [17] Harcourt, B. (2004). 'Rethinking Racial Profiling: A Critique of the Economics, Civil Liberties, and Constitutional Literature, and of Criminal Profiling More Generally', *University of Chicago Law Review*, vol. 71(4), pp. 1275-1381.
- [18] Knowles, J., Persico, N. and Todd, P. (2001). 'Racial Bias in Motor Vehicle Searches: Theory and Evidence', *Journal of Political Economy*, vol. 109(1), pp. 203-229.
- [19] Lamberth, J. (2003). 'Racial Profiling Data Analysis Study: Final Report for the San Antonio Police Department', Final Report for the San Antonio Police Department. Available at http://www.sanantonio.gov/sapd/pdf/LamberthSanAntonioRpt_2003.pdf.
- [20] Lovrich, N., Gaffney, M., Mosher, C., Pickerill, M. and Smith, M. (2003). 'WSP Traffic Stop Data Analysis Project: Data Analysis Project Report', Manuscript, Washington State University.
- [21] Manski, C. F. (forthcoming 2006). 'Search Profiling with Partial Knowledge of Deterrence', IN THIS ISSUE OF EJ.

- [22] Milchtaich, I. (2000). ‘Generic Uniqueness of Equilibrium in Large Crowding Games.’ *Mathematics of Operations Research*, vol. 25(3), pp. 349-64.
- [23] Nixon, J. (2003). ‘Executive Summary on 2003 Missouri Traffic Stops’, Available at <http://www.ago.mo.gov/racialprofiling/2003/racialprofiling2003.htm>.
- [24] Persico, N. (2002). ‘Racial Profiling, Fairness and Effectiveness of Policing’, *American Economic Review*, vol. 92(5), pp. 1472-97.
- [25] Schmeidler, D. (1973). ‘Equilibrium Points of Nonatomic Games,’ *Journal of Statistical Physics*, vol. 7(4), pp. 295-300.
- [26] 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.
- [27] Spitzer, E. (1999). ‘A Report to the People of the State of New York from the Office of the Attorney General’, Civil Rights Bureau, New York, NY.
- [28] Thomas, D. and Hansen, R. (2004). ‘Second Annual Report, Denver Police Department, Contact Card Data Analysis’, Report for the Denver Police Department.

- [29] Verniero, P. and Zoubek, P. (1999). 'Interim Report of the State Police Review Team Regarding Allegations of Racial Profiling', Attorney Generals Office, Trenton, NJ.
- [30] Withrow, B. L. (2002). 'The Wichita Stop Study', Available at <http://www.wichita.gov/NR/rdonlyres/445AC311-6154-49DF-B67F-A76026616793/0/StopStudydocument.pdf>.
- [31] Withrow, B. L. (2004). 'Driving While Different: A Potential Theoretical Explanation for Race-Based Policing', *Criminal Justice Policy Review*, vol. 15(3), pp. 344-364.

A Delegating Crime

We now generalize the model by allowing motorists to delegate the crime to others, or to disguise their appearance in order to pose as a member of a group that is less prone to crime. Our main theoretical finding is that the “hit rates” test of Theorem 2 still holds in this more general environments.

Let us assume that, in addition to (a) committing a crime and (b) not committing a crime, a motorist in group (r, c) can (c) delegate the crime to a member of another group (r', c') at a cost $d_{r,c}^{r',c'}$. Delegating the crime to a member of a different group is expedient if one’s own group is at a high risk of interdiction. The cost $d_{r,c}^{r',c'}$ represents the amount of money that a member of group (r, c) needs to pay for a member of group (r', c') to be willing to commit the crime on his behalf. The cost of delegation is allowed to depend on (r', c') to capture the notion that it might, for example, be more costly to convince an old lady to carry drugs than a young male. An alternative interpretation of our model is that the motorist who delegates the crime is essentially a criminal who disguises himself as a member of a different group. We assume that the benefit of committing a crime, as well as the cost if detected, accrue to the delegator.¹⁶ For expositional convenience, we adopt the convention that $d_{r,c}^{r,c} = 0$, i.e., committing a crime is equivalent to delegating to someone in one’s own group. The case treated in the main body of the paper corresponds to $d_{r,c}^{r',c'} = \infty$ for every $(r', c') \neq (r, c)$, that is, the cost of hiring someone else is prohibitively high. In this case delegating

to someone else (or disguising oneself) is not possible.

We now allow for within-group heterogeneity among motorists with respect to the benefits from crime, the costs of being detected, and the costs motorists face to delegate the crime to someone else (or to disguise themselves). Let $\mathbf{d}_{r,c}$ denote the vector $[d_{r,c}^{r',c'}]_{r',c'}$. Within each motorist group (r, c) , heterogeneity is captured by a joint distribution of v, j and $\mathbf{d}_{r,c}$, denoted $F_{r,c}(v, j, \mathbf{d}_{r,c})$.¹⁷

Let $\sigma(r, c)$ denote the number of searches of members of group (r, c) , and let σ denote the vector $[\sigma(r, c)]_{r,c}$. A member of group (r, c) with given $v, j, \mathbf{d}_{r,c}$ who hires someone in group (r', c') to commit a crime receives an expected payoff

$$u_{r,c}(v, j, \mathbf{d}_{r,c}, r', c', \boldsymbol{\sigma}) = v - d_{r,c}^{r',c'} - j \cdot \frac{\sigma(r', c')}{N_{r',c'}}.$$

Let $\mathbf{K}_{r,c}(v, j, \mathbf{d}_{r,c}, \boldsymbol{\sigma})$ denote the $2 \times C$ matrix that represents the crime delegation decision by an individual in group (r, c) with characteristics $v, j, \mathbf{d}_{r,c}$ when the police searches according to $\boldsymbol{\sigma}$. All entries in this matrix are zero except the one corresponding to the group from which this individual hires an agent. Formally, all entries are zero except the one corresponding to $\arg \max_{r',c'} u_{r,c}(v, j, \mathbf{d}_{r,c}, r', c', \boldsymbol{\sigma})$, which is equal to one if $\max_{r',c'} u_{r,c}(v, j, \mathbf{d}_{r,c}, r', c', \boldsymbol{\sigma}) \geq 0$ and zero otherwise. (If the argmax is a set then we select one of its elements at random and call it the argmax.) The matrix $\mathbf{K}_{r,c}(v, j, \mathbf{d}_{r,c}, \boldsymbol{\sigma})$ represents the optimal choice of the motorist as to which group to hire from, if any.

The choice of not committing a crime is represented by the null matrix.

The aggregate crime generated by members of group (r, c) is obtained by aggregating the choices of every member of group (r, c) , and it is captured by the $2 \times C$ matrix

$$\mathbf{K}_{r,c}(\boldsymbol{\sigma}) = N_{r,c} \int \mathbf{K}_{r,c}(v, j, \mathbf{d}_{r,c}, \boldsymbol{\sigma}) dF_{r,c}(v, j, \mathbf{d}_{r,c}).$$

The matrix denoting total crime committed within each group is obtained by adding up all the crime generated by all groups, and it is given by the $2 \times C$ matrix

$$\mathbf{K}(\boldsymbol{\sigma}) = \sum_{r,c} \mathbf{K}_{r,c}(\boldsymbol{\sigma}).$$

Let $K^{r,c}(\boldsymbol{\sigma})$ denote the r, c element in the matrix $\mathbf{K}(\boldsymbol{\sigma})$ divided by $N_{r,c}$. This element represents the amount of crime *committed by* members of group (r, c) , and is therefore denoted by a superscript, whereas a subscript would denote the crime *commissioned by* that group.

Let \mathbf{S} denote the vector $[S(r, c)]_{r,c}$. Officer p 's expected payoff is

$$(2) \quad \sum_{r,c} S_p(r, c) [y_p^r K^{r,c}(\mathbf{S}) - t_p]$$

Just as before, the results from Schmeidler (1973) yield existence of equilibrium. This game is not, however, a large crowding game in the sense of Milchtaich (2000). Uniqueness of equilibrium, therefore, is not guaranteed in this case from any theorem we know.

Go through the same steps followed on page 12 (except replacing \mathbf{S}^* for $S^*(\cdot, \cdot)$) to obtain the following counterpart to Theorem 2.

Theorem 3 *In any equilibrium, the hit rate is the same across all subgroups within a race. If the police are unbiased, the hit rate is the same across races, too. If the police are biased against race r , the hit rate is lower in race r than in the other race.*