Using Hit Rates to Test for Racial Bias in Law Enforcement: Vehicle Searches in Wichita*

Nicola Persico and Petra Todd
University of Pennsylvania†

November 18, 2004

Abstract

This paper considers the use of outcomes-based tests for detecting racial bias in the context of police searches of motor vehicles. It shows that the test proposed in Knowles, Persico and Todd (2001) can also be applied in a more general environment where police officers are heterogenous in their tastes for discrimination and in their costs of search and motorists are heterogeneous in their benefits and costs from criminal behavior. We characterize the police and motorist decision problems in a game theoretic framework and establish properties of the equilibrium. We also extend of the model to the case where drivers' characteristics are mutable in the sense that drivers can adapt some of their characteristics to reduce the probability of being monitored. After developing the theory that justifies the application of outcomes-based tests, we apply the tests to data on police searches of motor vehicles gathered by the Wichita Police deparment. The empirical findings are consistent with the notion that police in Wichita choose their search strategies to maximize successful searches, and not out of racial bias.

JEL Numbers: J70, K42

*The authors are grateful to the NSF for financial support.

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

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 are now collecting data on the demographic characteristics of the individuals that they subject to stops and searches and on the outcomes of these encounters. A common finding from the data is that blacks and Hispanics tend to be overrepresented in police stops and searches. The question arises as to whether higher rates of stop and search among these groups reflect a police bias or whether the higher rates can be justified as an optimal monitoring response to higher rates of criminality.

Various kinds of tests have been proposed in the literature to assess whether police behavior is racially biased. The simplest tests compare police monitoring rates against population benchmarks. These test are referred to as *benchmarking tests*. For example, let $D = 1$ if an individual is stopped and/or searched and let r denote the race/ethnicity of the individual. The simplest benchmarking test evaluates whether $\Pr(D = 1|r) = \Pr(D = 1)$ for all r. A more refined version of the test asks whether race/ethnicity predicts whether an individual is subject to monitoring, after taking account other characteristics (c) that the police are permitted to use as potential indicators of criminality, i.e. a test for whether $\Pr(D = 1|r, c) = \Pr(D = 1|c)$. Obviously, one drawback of this test is that the result can vary depending on the particular set of characteristics c used, and there is no concensus over what constitute valid conditioning characteristics. Another drawback of benchmark tests, in general, is that they are only informative on whether a racial disparity in monitoring rates exists and not on the motivation for the disparity. Notably, the test cannot distinguish whether police subject certain

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

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 applied work in this area and are the approach of choice for practitioners.

In previous work (Knowles, Persico and Todd, 2001, henceforth KPT), we proposed a different type of test, an *outcomes-based* one, for distinguishing the motivation for differential monitoring rates. We developed a model of police and motorist behavior in the context of police deciding which vehicles to subject to searches and motorists deciding whether to carry contraband. The model assumes that in the absence of racial bias, officers pursue a monitoring strategy that maximizes the number of successful outcomes, where a successful search outcome is defined as uncovering some contraband, such as drugs or illegal weapons. Racial bias is introduced into the model as a preference parameter that reduces the perceived cost of searching vehicles of black or Hispanic drivers, which can lead to oversearching of these groups. An equilibrium implication of racially biased monitoring that was shown in KPT is that the expected hit rate, the rate at which contraband is seized, should be lower for the groups subject to bias.

KPT tested the implication of the model that, in the absence of biased preferences, hit rates should be equal across all observable categories of drivers. This outcomes-based test was applied to a dataset gathered by the Maryland Police Department pertaining to all vehicles searches conducted on a stretch of highway 95 over a four year period. The dataset included information on the driver's race/ethnicity, gender, type of car, time of day the search took place, on whether any contraband was found and on the type of contraband. KPT found that the data do not reject the hypothesis that the hit rates are equal across black and white drivers and most other distinguishable categories of drivers.

Since the publication of KPT, a number of papers have explored extensions or variations of the model, some arguing that extending the KPT model in particular ways would invalidate the test for racial bias.[2] Partly motivated by this recent research, this paper extends the analysis proposed in KPT in several ways. In Section 2, we generalize the model of police and motorist behavior to allow for the possibility that individual police officers may differ in their degree of racial bias and motorists can vary in their benefits and costs from committing crimes. We also 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.[3] We show that the outcomes-based test for discrimination can still be applied in this more general environment.

In Section 5 we apply the outcomes-based test to a new dataset obtained from the Wichita Police Department. These data contain information on all vehicle stops and searches that took place in Wichita, Kansas during the first 9 months of 2001. Withrow (2004) initially analyzed these data and found a substantial disparity between the proportions of black drivers stopped (21%) and their representation in the Wichita population (11.4%). The disparity is much less for Hispanics, who represent 9.2% of the stops and 9.6% of the Wichita population.[4] In benchmarking-type tests, these disparities are taken to indicate some degree of racial animus on the part of the police. When we compare the hit rates for different groups of drivers who were subjected to discretionary vehicle searches, however, we find that the data do not reject the hypothesis that hit rates are equal for drivers of all races/ethnicities. Thus, the outcomes-based test suggests that the disparity is due to statistical discrimination, not to racial bias on the part of the police. Moreover, we find

[2]See e.g. Antonovics and Knight (2004), Anwar and Fang (2004), Dharmapala and Ross (2004).

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

[4]Population figures are based on US Census 2000 data.

that the hit rates do not differ by gender, by age, and differ only slightly by time of day.

In Section 6 we discuss additional implications of our analysis in light of alternative models of policing. We observe that the equalization of hit rates found in this paper coincides not only with results reported in KPT but also seems to be a general feature of traffic stop data sets. We discuss the extent to which this observation can be seen as a validation of the proposed model. Our discussion pays special attention to the objective function of police officers posited in this model, which is key for our outcomes-based test but which may lead to socially suboptimal police behavior. Section 7 concludes.

1.1 Related Literature

This paper extends the model originally introduced in Knowles, Persico, and Todd (2001). Some other papers have taken issue with the modeling assumptions of KPT. Three in particular deserve mention; they are Dharmapala and Ross (2004), Anwar and Fang (2004), and Antonovics and Knight (2004). Although each paper studies a different model, the first two papers share one basic modeling feature, that it is infeasible for the police to perfectly deter crime in a given subgroup of the population. Both papers conclude that under this assumption the hit rates test is not necessarily valid. Anwar and Fang (2004), moreover, 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 (2004) in that they also look at differential bias between black and white officers. In their paper, the police have heterogenous costs of search and they claim that this causes the hit rate test to break down. The results presented in this paper imply that, contrary to that claim, the hit rate test is valid even in their model. Nevertheless, we believe that Antonovics and Knight's broader point is correct, and that there is something to be learned from the rate at which officers of different races search citizens.

2 The Model of Motorist and Police Behavior

We next describe a model of police and motorist behavior that generalizes the model of KPT. Let r denote the race of the motorist, which is assumed to be distinguishable by the police. Without loss of generality, in this section we let $r \in \{A, W\}$. We denote other characteristics that are observable by the police by $c \in \{1, ..., C\}$. A group (r, c) has a mass of motorists $N_{r,c}$. We assume that police can distinguish motorist groups (r, c) and but cannot detect motorist heterogeneity within (r, c) groups. Let v represents the value to a motorist of committing a crime. If the crime is detected, the payoff to the motorist is $v - j$.

A driver in group (r, c) can (a) commit a crime, (b) not commit a crime, or (c) delegate the crime to a member of another group (r', c') at a cost $d_{r,c}^{r',c'}$. We assume that the benefit of committing a crime, as well as the cost if caught, accrue to the delegator. $\mathbf{d}_{r,c}$ denotes the vector $\left[d_{r,c}^{r',c'}\right]_{r',c'}$. We adopt the convention that committing a crime is equivalent to hiring someone in one's own group. Thus, $d_{r,c}^{r,c} = 0$, i.e., by doing the crime himself the motorist avoids the cost of hiring another motorist. Hiring a member of a different group to commit the crime is expedient if one's own group is at a high risk of interdiction. A motorist who delegates the crime to a member of a different group is essentially a criminal who disguises himself as a member of a different group. A special case arises when $d_{r,c}^{r',c'} = \infty$ for $(r, c) \neq (r', c')$, where hiring someone else (or disguising oneself) is impossible. This case, which we refer to as the fixed criminal characteristics case, is treated in the existing literature on racial profiling.

Within each motorist group (r, c), heterogeneity is captured by a joint distribution of v, j and $\mathbf{d}_{r,c}$, denoted $F_{r,c}(v, j, \mathbf{d}_{r,c})$. Thus, we allow for heterogeneity with respect to benefits from crime, the costs from crime, and the cost they face to hire someone else to commit the crime or disguise themselves. It is assumed that the cdf's $F_{r,c}$ have no atoms and their support is a rectangle.

We assume 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. Antonovics and Knight (2004) argue that the heterogeneity in search cost across police officers (in our notation, the fact that t_p may vary with p) would invalidate the hit rates test. This article clarifies that this is not the case.

If a search of a motor vehicle does not yield any contraband (such as illegal drugs or weapons), then we term the search unsuccessful and assume there is no benefit. We introduce the potential for police bias by allowing the benefit that police derive 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 Americans if $B(p) > 0$ for all p, biased against whites if $B(p) < 0$ for all p, and unbiased if $B(p) = 0$ for all p.[5] If no search is conducted, there is a zero payoff.

As described, this setup can accommodate police heterogeneity in intensity of the bias. However, we rule out environments in which $B(p)$ changes sign, i.e., where some policemen are biased against whites and some are biased against African Americans. Below, we propose a test for infering the sign of $B(p)$.

[5] To ensure that integration with respect to p is well defined, we assume that the functions S_p, t_p, y_p^A, and $B(p)$ are Lebesgue measurable.

3 Examples

3.1 Unbiased police

First, consider the case where police are unbiased i.e., $B(p) \equiv 0$. Also, for simplicity assume that the only observable characteristic is r (there are no c). We next establish that if both groups are searched in equilibrium, then both must have the same crime rate. Indeed, suppose not, and suppose the crime rate $\kappa_A > \kappa_W$, where the crime rate corresponds to the fraction of the group choosing to carry contraband. When $B(p) = 0$, the police's expected utility from searching either group is

$$\kappa_r y_p^W - t_p.$$

If $\kappa_A > \kappa_W$, then for every p this expression is larger when $r = A$. Thus, every policeman will concentrate his/her searches solely on group A, which contradicts the assumption that both groups are searched. Thus, if in equilibrium both groups are searched, the crime rate must be equal across the groups ($\kappa_A = \kappa_W$). Also, note that only policemen for which $t_p \leq \kappa_A y_p^W$ will conduct searches. Every police who searches is indifferent between searching either group, so the police will randomize their searches in such a way as to achieve $\kappa_A = \kappa_W$.

3.2 Biased police

Suppose now that the police are biased against African Americans, that is, $B(p) > 0$ for all p. If in equilibrium both groups are searched, then it cannot be that $\kappa_A \geq \kappa_W$. Indeed, suppose $\kappa_A \geq \kappa_W$. Then a policeman's payoff from searching group A is

$$\begin{aligned} & \kappa_A y_p^A - t_p \\ = \; & \kappa_A \left(y_p^W + B\left(p\right)\right) - t_p \\ \geq \; & \kappa_W \left(y_p^W + B\left(p\right)\right) - t_p \\ > \; & \kappa_W y_p^W - t_p, \end{aligned}$$

where the weak inequality follows from the assumption that $\kappa_A \geq \kappa_W$ and the strict inequality follows from $B\left(p\right) > 0$. This chain of inequalities shows that no police would search group W, which contradicts the assumption that both groups are searched. Thus, if both groups are searched and police officers are biased, then $\kappa_A < \kappa_W$.

Consider equilibria in which both groups are searched. As we have seen, whenever the police are unbiased we have equal crime rates across groups ($\kappa_A = \kappa_W$), and whenever the police are biased against group r, in equilibrium $\kappa_r < \kappa_{r'}$. Thus, a comparison of the crime rates in the two groups reveals whether there is a bias as well as the direction of the bias; namely, the crime rate will be lower for the group subject to bias.

4 Equilibrium

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

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

Let $\mathbf{K}_{r,c}\left(v, j, \mathbf{d}_{r,c}, \boldsymbol{\sigma}\right)$ denote the 2xC matrix in which all entries are zero except the one corresponding to $\arg\max_{r',c'} u_{r,c}\left(v, j, \mathbf{d}_{r,c}, r', c', \boldsymbol{\sigma}\right)$, which is equal to one if $\max_{r',c'} u_{r,c}\left(v, j, \mathbf{d}_{r,c}, r', c', \boldsymbol{\sigma}\right) \geq 0$ 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 a motorist as to which group to hire from, if any. The choice of not committing a crime is represented by a null matrix. The crime generated by members of group (r, c) is captured by the 2xC matrix

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

and total crime is given by the 2xC 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 $N_{r,c}$. The function $K^{r,c}(\boldsymbol{\sigma})$ summarizes the crime rate in group (r, c) when the search strategy of the police is $\boldsymbol{\sigma}$. One can think of this function as the supply of crime.

Denote by $S_p(r, c)$ the number of searches that officer p devotes to group (r, c). The total number of searches of members of group (r, c) is

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

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

$$\sum_{r,c} S_p(r, c) \left[y_p^r K^{r,c}(\mathbf{S}) - t_p \right] \tag{1}$$

4.1 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. This game has a continuum of players and finite action sets. Moreover, this 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 $\mathbf{S}$). Schmeidler (1973) establishes the existence of an equililbrium 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 $\mathbf{S}$, which means that each function $K^{r,c}(\mathbf{S})$ 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 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 require 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 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. Thus, the results from Schmeidler (1973) yield existence of equilibrium.

Uniqueness of equilibrium can also be obtained in an important special case. If criminal characteristics are fixed, and so criminals cannot hire from different groups, the return to the police from searching group (r,c) is not affected by the distribution of searches across other groups. Thus, officer p's expected payoff simplifies to

$$\sum_{r,c} S_p(r,c)\left[y_p^r K^{r,c}(S(r,c)) - t_p\right],$$

where each function $K^{r,c}(S(r,c))$ is decreasing in $S(r,c)$. Because for every p this payoff function 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). Because each officer's payoff is affected by the actions of the other officers only through the total number of officers that choose to search group (r,c), this game is a *large crowding game* in the sense of Milchtaich (2000). Milchtaich (2000) shows that "generically" large crowding games have a unique equilibrium. Generically means that the set of matrices $\mathbf{K}(\boldsymbol{\sigma})$ that give rise to a game with a unique equilibrium is the intersection of countably many dense open sets in the space of all matrices of decreasing functions. The following theorem obtains.

Theorem 1 *A Nash equilibrium exists. If the criminals' characteristics are fixed, then the equilibrium is generically unique.*

4.2 Characterization of equilibrium

Suppose in equilibrium groups (r,c) and (r',c') are searched. Then, there must be a p and a p' such that

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

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

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

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

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

Note that the implications on the crime rate translate into testable implications on the hit rates, i.e., on the likelihood that a search yields a find. These observations motivate the following theorem:

Theorem 2 *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.*

5 Empirical Results

5.1 Data description

We now apply the outcomes-based test to data that were collected by the Wichita police department for the purpose of assessing whether officers engage in racially biased policing practices. The dataset contains information on every police/citizen contact 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 search was conducted, on the rationale 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 involved in the incident.[6]

[6] The 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). Unfortunately, the Wichita Police Department would not release this information to us. When Withrow (2004) analyzed the data with respect to this variables, he concluded that "enforcement patterns do not differ substantially

A key assumption of the model developed in the previous section is that police have a choice over whom to monitor. The model therefore does not apply to cases where police have no discretion, as when they are called to investigate the scene of an accident. Also, 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.[7] We therefore limit our analysis sample to observations on police-motorist encounters where police have discretion over whether to initiate the encounter. Also, to make plausible the assumption that the monitoring cost is constant across stops, we restrict ourselves to a relatively homogenous set of observations on stops that involve vehicle searches. Approximately 13% of the stops in our dataset included a vehicle search. After applying these sample restrictions, our final analysis sample contains information on 2,288 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 blacks involved in stops (21.45%) and searches (32.65%) is significantly higher than their representation in the population (11.4%). The percentage of whites in seaches (63.61%) is lower than in the population (65.2%). 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

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."

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

searches, which shows that most 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 are carried out at night. 76% are conducted at night, between the hours of 7pm to 8am.

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. The most common type of contraband siezed is drugs/drug paraphenalia, followed by alcohol/tobacco, stolen property and firearms. Among persons searched, black and Hispanic drivers are significantly more likely to be found with drugs/drug paraphenalia, 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.[8]

5.2 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 and in KPT (2001) 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 with contraband within cells defined by the conditioning variables against the proportion that would be expected under the null hypothesis of no association between the hit rate and the set of conditioning characteristics. The test statistic for testing the

[8]In any given search, they may be multiple rationales, so the categories are not mutually exclusive.

hypothesis of no association between hit rate and race is

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

where R is the cardinality of the set of race categories, $\mathcal{R}$, and $\widehat{p}_r$ and $\widehat{p}$ are conditional and unconditional estimated guilt proportions.

Table 4a-4g show the percentage of motorists found to be carrying contraband for groups of motorists defined for different sets of 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 hit rates are equal for all the race groups (p-value is 0.365), even though the sample sizes are relatively large. According to our model, this finding is consistent with no racial bias in police search behavior.

Table 4b breaks the percentages down by the age of the driver. We find that the test does 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 by gender.

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 nighttime verses daytime searches could plausibly be due to a higher cost of conducting search activities are night (for example, if police who work at night are 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.

Overall, these results show that the hit rates are very similar across groups of motorists no matter how these groups of motorists are defined. Equality of hit rates if a key prediction of the theoretical model when the police are unbiased. In light of the model, this evidence is consistent with the notion that police in Wichita are searching blacks and Hispanics at higher rates relative to their population in order to maximize the probability of finding contraband and not out of racial bias.

6 Discussion

The model described in Section 2 is one where individual officers choose search strategies that maximize the hit rates. Implicitly, it is assumed that officers can focus their searches on whatever subgroup (r, c) they choose. In particular, the model does not account for the presence of a central authority, a police chief, say, who can direct his officers to focus their searches on particular subgroups. This is relevant because a police chief whose goal was to minimize crime in his jurisdiction might want to focus interdiction on specific subgroups (r, c). The goal of minimizing the crime rate in a jurisdiction is different than allowing individual police to pursue a policy that maximizes hit rates.[9] Intuitively, in order to catch criminals there has to be 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 KPT test fails as a test of the unbiasedness of the police chief. Indeed, in the equilibrium of such a model, an unbiased police chief will allocate searches to equate the deterrence effect, and not the hit

[9] This point is made in several papers, including Alexeev and Leitzel (2002), Harcourt (2004), Eeckhout, Persico and Todd (2003), Manski (2004), and Persico (2002).

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 would be ineffective if criminals were mobile, and could 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 individuals officers is never observed but how many criminals they catch is observed. It is therefore easier to reward officers on the basis of their performance record in catching criminals. For this reason, we believe the model where police act as independent agents trying to catch criminals, could be viewed as a second best objective that a police chief might reasonably adopt.

Table 5 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.[10] This regularity is puzzling in the context of a crime-minimizing police chief but not in light of the KPT model, which offers a simple rationale for the equalization of hit rates across races, namely (a) that individual police officers are allocating searches in a way that maximizes successful searches and (b) that police departments, on average, are not afflicted by widespread bias against African Americans. Whether in fact this is really the case can only be ascertained

[10]This 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.

with more work on the data sets that are recently becoming available.[11]

7 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 test for racial bias proposed in KPT 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. This paper also extends the KPT model to the case where drivers' characteristics can be altered to reduce the probability of being monitored.

The goal of the modeling part of the present work is not to argue that the KPT test is robust to any change in the modeling assumptions, but rather to generalize the conditions under which the test is valid. Whether any modeling approach, including the KPT model, is a useful tool is ultimately an empirical question.

Working towards this goal, in this paper we apply the KPT test to a dataset gathered by the Wichita Police deparment 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,

[11] Further evidence on this front 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 (2004).

the hit rates are stable across various groups of drivers. Equality of hit rates is a key prediction of the theoretical model when police are not motivated by racial bias. Thus, our empirical findings are consistent with the notion that individual police officers in Wichita choose their search strategies to maximize efficiency in finding contraband and not out of racial bias.

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 6. It 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.[12] The KPT model offers a simple rationale for the widespread equalization of hit rates across races, and that is that police departments are, on average, not afflicted by widespread bias against African Americans. We are mindful, however, that other models may be consistent with the observed regularity and may deliver different implications concerning police bias. Clearly, more research is needed and more data sets need to be examined in order to obtain a comprehensive view of what type of model best explains the outcomes of police-motorist encounters.

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

References

[1] Alexeev,Michael, and Jim Leitzel (2002) "Racial Profiling". Manuscript, Indiana University in Bloomington.

[2] Antonovics, K. L. and Knight, Brian G. (2004): "New Look at Racial Profiling: Evidence from the Boston Police Department,"NBER Working Paper No. w10634.

[3] Anwar, Shamena and Hanming Fang (2004) "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," working paper.

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

[5] Bunzel, Helle and Marcoul, Phillipe (2003): "Can Racially Unbiased Police Perpetuate Long Run Discrimination?," working paper, Iowa State University.

[6] 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.

[7] 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.

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

[9] 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

[10] Eeckhout, Jan, Persico, Nicola and Petra Todd (2003): "Crackdowns," working paper.

[11] 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.

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

[13] Harcourt, Bernard.(2004) "Rethinking Racial Profiling: A Critique of the Economics, Civil Liberties, and Constitutional Literature, and of Criminal Profiling More Generally," 71 *University of Chicago Law Review* (forthcoming Fall 2004).

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

[15] Lamberth, John (2003): "Racial Profiling Data Analysis Study: Final Report for the San Antonio Police Department."

[16] 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.

[17] Manski, Charles F. (2004) "SEARCH PROFILING WITH PARTIAL KNOWLEDGE OF DETERRENCE." Manuscript, Northwestern University, October 2004.

[18] Milchtaich, Igal "Generic Uniqueness of Equilibrium in Large Crowding Games." *Mathematics of Operations Research* 25(3), August 2000, pp. 349-64.

[19] 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.

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

[21] Schmeidler, David (1973), "Equilibrium points of nonatomic games," *Journal of Statistical Physics* 7, pp. 295-300.

[22] 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.

[23] 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.

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

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

[26] Withrow, Brian L. "Driving While Different: A Potential Theoretical Explanation for Race-Based Policing" in *Criminal Justice Policy Review,* Vol. 15, No. 3, 344-364 (2004)

[27] Withrow, B. L. (2002). The Wichita stop study. Retrieved from http://www.wichitapolice.com/

Table 1a
Comparison of Stop and Search Percentages Against Population Benchmark

	Percentage in Population*	Percentage of stops	Percentage of searches+
Black	11.4	21.45	32.65
Asian	4.0	2.81	2.09
White (incl Hispanic)	75.2	73.90	63.61
White - NonHispanic	65.6	64.37	50.81
White -Hispanic	9.6	9.53	12.80
Native American	1.2	0.17	0.48
Other	8.2	1.68	49.19

*** Based on Withrow (2002b), tabulated from Census 2000 data for Wichita.**
+ Excludes searches that were incident to an arrest (where officers are required to search) and searches for which there was a warrant for the arrest of the driver.

Table 1b
Age Distribution of Persons subject to Stops and Searches+

	Percentage of stops	Percentage of searches+
Age < 18	6.68	8.63
Age 18-24	31.13	36.28
Age 25-34	26.20	25.22
Age 35-50	26.84	25.60
Age over 50	9.14	4.27

+ Excludes searches that were incident to an arrest (where officers are required to search) and searches for which there was a warrant for the arrest of the driver.

Table 2
Percentage Found with Contraband of Given Type
by Race/Ethnicity

	Black	White	Hispanic
Currency	2.72	1.67	2.16
Firearm	7.07	11.67	3.96
Other weapon	2.72	3.33	3.60
Drugs, Paraphenalia	53.26	38.33	58.99
Alcohol, Tobacco	23.37	40.00	25.90
Stolen Property	9.78	13.33	14.75
Other	20.65	5.00	5.04

Table 3
Search Rationale by Race/Ethnicity

	Black	Hispanic	White
Vehicle Indicators	8.75	9.12	6.66
Verbal Indicators	10.6	12.26	10.86
Physical Indicators	30.33	24.53	32.25
Document Indicators	3.82	3.14	1.43
Incident to Arrest	27.13	27.36	25.12
Other	26.02	21.07	18.46
Not app.	10.36	16.67	20.05

*P-value from Pearson Chi-Squared Tests of Hypothesis that Proportion Equal Across Race Groups

Table 4a
Hit Rates by Race/Ethnicity
(Total Number of Observations in parenthesis)

	Percentage
Black	22.69
	(811)
Hispanic	18.87
	(318)
White	22.03
	(1262)
P-values from Pearson Chi-Squared Tests of Hypothesis that Proportion Equal	0.365

Table 4b
Hit Rates by Age
(Total Number of Observations in parenthesis)

	Percentage
Age < 18	28.64
	(206)
Age 18-24	21.71
	(866)
Age 25-34	19.93
	(602)
Age 35-50	21.77
	(611)
Age over 50	20.59
	(102)
P-value from Pearson Chi-Squared Tests of Hypothesis that Proportion Equal	0.137

Table 4c
Hit Rates by Age and Race
(Total Number of Observations in parenthesis)

	Black	Hispanic	White	p-value*
Age < 18	30.77	26.09	27.97	0.885
	(65)	(23)	(118)	
Age 18-24	20.00	17.12	24.58	0.113
	(300)	(146)	(419)	
Age 25-34	23.12	16.00	19.14	0.309
	(199)	(100)	(303)	
Age 35-50	24.09	20.09	19.77	0.367
	(220)	(46)	(344)	
Age over 50	20.0	50.0	20.0	0.5830
	(25)	(2)	(75)	

*P-value from Pearson Chi-Squared Tests of Hypothesis that Proportion Equal Across Race Groups

Table 4d
Hit Rates by Gender
(number of observations in category in parenthesis)

	Percentage
Male	21.61
	(1916)
Female	22.64
	(477)
P-value from Pearson Chi-Squared Tests of Hypothesis that Proportion Equal	0.625

Table 4e
Hit Rates by Gender and Race
(number of observations in category in parenthesis)

	Black	**White**	**Hispanic**	**P-value**
Male	22.71	18.86	21.67	0.418
	(687)	(281)	(946)	
Female	22.58	18.92	23.10	0.847
	(124)	(37)	(316)	

Table 4f
Hit Rates by Time of Day
(number of observations in category in parenthesis)

	Proportion
Daytime	17.11
	(561)
Nighttime	23.38
	(1835)
P-value from Pearson Chi-Squared Tests of Hypothesis that Proportion Equal	0.002

* Daytime is in between the hours of 8am and 7pm.

Table 4g
Hit Rates by Time of Day and Race
(number of observations in category in parenthesis)

	Black	**Hispanic**	**White**	**P-value***
Daytime	19.34	9.09	17.69	0.208
	(181)	(55)	(294)	
Nighttime	23.65	16.77	23.35	0.165
	(630)	(155)	(968)	

*P-value from Pearson Chi-Squared Tests of Hypothesis that Proportion Equal Across Race Categories

Table 5
Summary of Hit Rate Findings for Racial Profiling Studies

	Hit Rates for Whites	Hit Rates for Blacks	Hit Rates For Hispanics
Wichita, KS (this study)	22.7	22.03	18.9
Maryland++	32	34	11
Florida§§	25.1	20.9	11.5
Tennessee§	20.1	19.2	10.3
New Jersey**	10.5	13.5	nr
Rhode Island+	23.5	17.8‡	17.8‡
New York (pedestrian)*	13	11	nr
Charlotte, NC¶	30.9	24.2	nr
Lansing, MI ¶¶	6.8	8.7	nr
Missouri †	23.2	17.5	14.7
San Antonio, TX††	17.2	14.6	14.9
Denver, CO#	16.5	19.7	11.3
Denver, CO (pedestrian)#	18.7	20.6	14.6
Los Angeles, CA ##	23.8	18.2	17.2
Sacramento, CA***	26.5	22.4	28
San Diego, CA§§§	11	12	5
Washington State†††	32	21	nr

nr = not reported
‡The hit rate is reported for minorities.
*These searches pertain to pedestrians. Spitzer (1999)
+Farrel et. al. (2003)
++KPT (2001)
**Verniero and Zoubak (1999)
§§Anwar and Fang (2004)
§Cohen-Vogel and Doss (2002)
¶ Smith et. al. (2004)
¶¶ Carter et. al. (2002)
† Nixon (2003)
†† Lamberth (2003)
# Thomas and Hansen (2004)
Tabluations provided by LAPD on file with the authors, Jan-Jun, 2001
*** Greenwald (2003)
§§§ Cordner, G. et al. (2001)
††† Lovrich et. al. (2003)