AN ALTERNATIVE TEST OF RACIAL PREJUDICE IN MOTOR VEHICLE SEARCHES: THEORY AND EVIDENCE

By

Shamena Anwar and Hanming Fang

July 2004

COWLES FOUNDATION DISCUSSION PAPER NO. 1464

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence*

Shamena Anwar[†] Hanming Fang[‡]

June 28, 2004

Abstract

We exploit a simple but realistic model of trooper behavior to design empirical tests that address the following two questions. Are police monolithic in their search behavior? Is racial profiling in motor vehicle searches motivated by troopers' desire for effective policing (statistical discrimination) or by their racial prejudice (racism)? Our tests require data sets with race information about both the motorists and troopers. When applied to vehicle stop and search data from Florida, our tests can soundly reject the null hypothesis that troopers of different races are monolithic in their search behavior, but fail to reject the null hypothesis that none of the racial groups of troopers are racially prejudiced.

Keywords: Statistical Discrimination, Racial Prejudice, Racial Profiling.

JEL Classification Number: J7.

---

*We are grateful to Joseph Altonji for his constant support and many suggestions, and to Kate Antonovics and Nicola Persico for helpful discussions. We thank Richard Zeller at Florida Highway Patrol for providing us with the data, and Jessica Duan for able research assistance in the beginning stage of this project. Fang is grateful to financial support from a Yale University Junior Faculty Fellowship. All remaining errors are our own.

†Department of Economics, Yale University, P.O. Box 208268, New Haven, CT 06520-8268. Email: shamena.anwar@yale.edu

‡Department of Economics, Yale University, P.O. Box 208264, New Haven, CT 06510-8264. Email: hanming.fang@yale.edu

1 Introduction

Black motorists in the United States are much more likely than white motorists to be searched by highway troopers. Several recent lawsuits against state governments have used this racial disparity in treatment as evidence of "racial profiling," a term that refers to the police practice of using a motorist's race as one of the criteria in their motor vehicle search decisions. Racial profiling originated with the attempt to interdict the flow of drugs from Miami up Interstate 95 to the cities of the Northeast. For example, in 1985 the Florida Department of Highway Safety and Motor Vehicles issued guidelines for police on "The Common Characteristics of Drug Couriers," in which race/ethnicity was explicitly mentioned as one characteristic (Engel, Calnon and Bernard, 2002). While the initial motivation for such guidelines may have been to increase the troopers' effectiveness in interdicting drugs, it also unfortunately opened up the possibility for troopers to engage in racist practices against minority motorists.

Following the public backlash generated by several cases in the 1990s such as *Wilkins v. Maryland State Police* [1996] and *Chavez v. Illinois State Police* [1999], almost all highway patrol departments have denounced using race as a criterion in stop and search decisions. But many citizens, especially minorities, are skeptical of this claim: motor vehicle search decisions, by their very nature, are made in the midst of face-to-face interactions, and thus it is simply hard to imagine that troopers can block the race and ethnicity information that a motorist presents. Moreover, data on trooper searches continue to show that they tend to search a higher proportion of minority motorists than white motorists. As is now well known, however, racial disparities in the aggregate rates of stops and searches do not necessarily imply *racial prejudice* (see, for example, Knowles, Persico and Todd 2001, Engel, Calnon and Bernard 2002). If, for example, black drivers are more likely than white drivers to carry contraband, then the aggregate rate of stops and searches would be higher for black drivers even when race was hypothetically invisible to troopers. Moreover, racial profiling may also arise if police attempt to maximize successful searches and race helps predict whether a driver carries contraband. This situation is called *statistical discrimination* in the terminology of Arrow (1973).

How can we empirically distinguish racism from statistical discrimination? This question has garnered enormous public and academic interest (see, for example, National Research Council 2004), but it is also challenging, partly as a result of data limitations. For example, unless truly random searches are conducted, researchers typically will not observe the *true* proportion of drivers who carry contraband. Ethnographic studies such as Sherman (1980) and Riksheim and Chermak (1993) have shown that many situational factors, including suspects' demeanor in the police-citizen encounter, influence police behavior. Such data are also typically unavailable. Because we have no way of controlling for all of the legitimate factors that might cause minority drivers to be searched with higher probability than white motorists, it becomes very difficult to determine the true motivation behind racial profiling with the available data.

A seminal paper by Knowles, Persico and Todd (2001, KPT hereafter) developed a simple but elegant theoretical model about motorist and police behavior that suggests an empirical test using data on search outcomes (i.e., the percent of searches in which contraband was found) for each racial group – a statistic typically available to researchers.[1] The primary idea of KPT's empirical test comes from the outcome test originated by Becker (1957).[2] It is based on the following intuitive notion. If troopers are profiling minority motorists due to racial prejudice, they will search minorities even when the returns from searching them, i.e., the probabilities of successful searches against minorities, are smaller than those from searching whites. More precisely, if racial prejudice is the reason for racial profiling, then the success rate of the *marginal* minority motorist (i.e., the last minority motorist deemed suspicious enough to be searched) will be lower than the success rate of the *marginal* white motorist. In contrast, if racial profiling results from statistical discrimination (i.e., if the troopers are profiling to maximize the number of successful searches), then the optimality condition would require that the search success rate for the marginal minority motorist be equal to that of the marginal white motorist. While this idea has been well understood, it is problematic in empirical applications because researchers will never be able to directly observe the search success rate of the *marginal* motorist. Instead we can only observe the *average* success rate of white and minority searches. Precisely for this reason, KPT proposed a simple model of motorist and police behavior to cleverly circumvent this problem. In their model, motorists differ in their characteristics, including race and possibly other characteristics (that are observable to troopers but may or may not be available to researchers). Troopers decide whether to search motorists and motorists decide whether to carry contraband. In this "matching pennies"-like model they show that, if troopers are not racially prejudiced, then all motorists, regardless of their race

[1] The ideas in our paper are inspired from reading KPT, from which we learned a great amount.

[2] See, also, Ayres (2001) for other applications of the outcome test idea.

and other characteristics, would in equilibrium carry contraband with equal probability, and thus there is *no difference* between the marginal and the average search success rates.[3] Their model thus suggests a simple test based on the comparison of the *average* search success rate by the race of the *motorists*. A lower average search success rate implies racial prejudice against that group. Applying their test to a data set of 1,590 searches on a stretch of the I-95 in Maryland from January 1995 through January 1999, they find no evidence of racial prejudice against African-American motorists, but do find evidence of racial prejudice against Hispanics.

KPT's model of motorist and police behavior provides a theoretical rationale for using the average search success rates as the basis of an empirical test for racial prejudice. Therefore, the validity of the test also hinges on the realism of the model that justifies it. We now present two weaknesses of their model.[4]

First, KPT's model predicts that all motorists for a given race, regardless of their other characteristics that may be observed by the police, will carry contraband with equal probability. This is the vital prediction that allows them to equate the average search success rate in a given racial group of motorists to the marginal search success rate. This, however, also implies that a motorist's characteristics other than race should provide no information when a trooper decides whether to search. This implication of police behavior goes against trooper guidelines which require them to base their search decisions on the information the motorist presents to the trooper at the time of the stop, including the motorist's personal characteristics, their demeanor, and the contents of their vehicle that are in plain view, etc. (see, e.g., Sherman 1980 and Riksheim and Chermak 1993). KPT's basic model assumes that motorists' characteristics are exogenous, thus ruling out the plausible scenario that a motorist's demeanor when stopped is intimately related to whether or

[3]They of course allow the motorists with different characteristics to have different costs and benefits from carrying contrand. These differences, however, only imply that in equilibrium troopers will search motorists with different characteristics at different rates. In fact, these different search rates provide the necessary deterrence to ensure that all motorists will carry contrand with equal probabilities.

[4]Dharmapala and Ross (2003) also point out that KPT's test does not generalize if potential drug carriers may not be *observed* by the police or if there are different levels of drug offense severity. In the first case, the equilibrium of the model may involve a group of motorists carrying drugs with probability one even when they are searched with probability one whenever the troopers observe them (KPT recognized this issue in their footnote 16). If the probability of being a "dealer" is higher for minorities, then the average success rate for minorities should be greater than that for whites under statistical discrimination, and equal average success rates would actually indicate taste discrimination, contrary to KPT's conclusion. In the second case, KPT's test has to be modified.

not he or she is carrying contraband.

Second, KPT (and this field of research in general) assume that all troopers' behavior is *monolithic*. Due to lack of data on the characteristics, including race information, of troopers, it is assumed that all troopers have the same racial prejudice against motorists, regardless of their race. While there is no direct evidence on this assumption in the context of highway searches, Donohue and Levitt (2001), in their study on arrest patterns and crime, find that the racial composition of a city's police force has an important impact on the racial patterns of arrests, suggesting that police behavior (or information they possess) is not monolithic. The consequence of an invalid monolithic trooper behavior assumption is serious. Imagine a world in which minority troopers are racially prejudiced against white motorists, while white troopers are prejudiced against minority motorists. It is possible that when examining the aggregate search outcomes of white and minority troopers, we would reach a conclusion that the police as a whole are not racially prejudiced. But this seriously underestimates the harassment experienced by both white and minority motorists.

In this paper, we develop an alternative model of motorist and police behavior in which troopers are allowed to behave differently depending on their own race and the race of the motorists they interact with.[5] Our model does not yield the convenient, but in our view unrealistic, implication that all drivers of the same race carry contraband with the same probability. As a result, the distinction between average and marginal search success rates becomes, yet again, the central issue in the empirical determination of racial prejudice versus statistical discrimination. Our model follows the spirit of labor market statistical discrimination models (see, e.g., Coate and Loury 1993). Police officers observe noisy but informative signals about whether or not a driver carries contraband when they decide if a search is warranted. Guilty drivers, i.e., drivers who actually carry contraband, are more likely than innocent drivers to generate suspicious signals. A police officer incurs a cost of search $t(r_m; r_p)$ that depends on both his/her own race r_p and the race of the motorist r_m. Troopers of a particular race, say r_p, are said to be racially prejudiced if their cost of searching motorists depend on the race of the motorist. The police force exhibits non-monolithic behavior if the cost of searching motorists of a given race r_m depend on the race of the

[5] We assume that race is the only characteristic of troopers that is likely to affect their search behavior. This is a plausible assumption because we are examining if troopers search white and minority motorists differently, so the race of the trooper is the most likely characteristic to affect their search patterns. We assume that within a trooper racial group, all troopers are monolithic.

trooper. Troopers are assumed to make their search decisions to maximize the number of successful searches (or arrests). The optimal decision of a race-r_p police officer in deciding whether a race-r_m motorist should be searched satisfies a threshold property: motorists should be searched if and only if their posterior probability of being guilty exceeds the search cost of race-r_p officers against race-r_m motorists, $t(r_m; r_p)$. We show that the police officers exhibit monolithic behavior if and only if both the search rate and average search success rate of any given race of motorists are independent of the race of the troopers conducting the search. Moreover, if none of the racial groups of troopers are racially prejudiced, then the ranking across the race of troopers of search rates and average search success rates for a given race of motorists should not depend on the race of the motorists. That is, if troopers of race r_p have a higher search rate against race-r_m motorists than troopers of race r'_p, then race-r_p troopers should also have a higher search rate against race-r'_m motorists than race-r'_p troopers. We use these theoretical predictions of the model to design empirical tests for both monolithic behavior and racial prejudice. Another nice feature of our model is that it could potentially be refuted by the data we have available.

The implementation of our empirical tests relies on data sets that have race information on both the troopers and motorists. While such data has not been available for use in earlier empirical studies on racial profiling, we were able to obtain a data set from the Florida Highway Patrol which contains information on all vehicle stops and searches conducted on Florida highways between January 2000 and November 2001, together with the demographics of the trooper that conducted each stop and search. In implementing our empirical tests, we find strong evidence that the Florida Highway Patrol troopers do not exhibit monolithic behavior, but we fail to reject the null hypothesis that none of the racial groups of troopers are racially prejudiced.

There is now a growing economics literature on the issue of empirical distinction between statistical discrimination and racial prejudice in motor vehicle searches. Hernandez-Murillo and Knowles (2004) extend KPT's test to the analysis of racial bias using only aggregate statistics, and apply this new test to Missouri's annual traffic-stop report for the year 2001. Antonovics and Knight (2004) generalize KPT's model to allow for trooper heterogeneity, and show that KPT's tests are not robust to such a generalization. As we do in our paper, they show that if officers of different races have the same search cost against motorists of a given race, then the search rate against these motorists should be independent of the officers' race. They run a Probit regression using data from the Boston Police Department where the dependent variable is an indicator for whether a search took place for a given stop, and the explanatory variables include some observable characteristics of the driver and officer and a dummy variable indicating whether there is a racial mismatch between the officer and the driver. In their baseline regression, they find a positive coefficient on the "racial mismatch" variable, indicating that officers are more likely to conduct a search against motorists of races different from their own. They interpret this finding as evidence of racial prejudice. We argue in subsection 2.1.2 that their interpretation of the evidence may be misleading. It is also useful to point out that their data is from the Boston Police Department and consists mainly of stops and searches in *local* neighborhoods. There are two potential problems with such data. First, as Hernandez-Murillo and Knowles (2004) argued, many stops and searches conducted in local streets are in response to specific crime reports. In these situations, officers tend to have less discretions over who they search. Second, as argued by Donohue and Levitt (2001), for stops and searches conducted in local neighborhoods, it is much more likely that officers of different races may possess different amounts of information regarding a motorist, as residents in the neighborhood may be more willing to share information with officers with the same race as theirs. In contrast, our data consists only of stops and searches conducted on highways, and as a result the above two issues are less concerning.

The remainder of the paper is structured as follows. Section 2 presents and analyzes our model of trooper search behavior, and proposes empirical tests based on the theoretical predictions of the model; Section 3 describes the data set from the Florida Highway Patrol, presents our test results, and contrasts our results with those using KPT's test; Section 4 concludes. In Appendix A we present a simple equilibrium model of drug carrying behavior to show that our focus on trooper behavior in Section 2 is not problematic.

2 The Model

We now present a simple model of trooper search behavior that underlines the empirical work in Section 3.2.[6] There is a continuum of troopers (interchangeably, police officers) and motorists (interchangeably, drivers). Let r_m and $r_p \in \{M, W\}$ denote the race of the motorists and the

[6]Borooah (2001) develops a somewhat related model of policing behavior.

troopers respectively, where M stands for minorities and W for whites.[7] Suppose that among motorists of race $r_m \in \{M, W\}$, a fraction $\pi^{r_m} \in (0,1)$ of them carry contraband.[8,9]

The information that is available to an officer when he or she makes the search decision consists of the motorist's race and many other characteristics pertaining to the motorist. Such characteristics may include, for example, the gender, age and residential address of the driver, the interior of the vehicle that is in the trooper's view, the smell from the driver or the vehicle, whether the driver is intoxicated, the demeanor of the driver in answering the trooper's questions, the make of the car, whether the car has an out-of-state plate, whether the car is rented or owned, location and time of the stop, as well as the seriousness of the reason for the stop, etc.[10] Note that while the police officer observes all the characteristics in the decision to search, a researcher will typically have access to only a small subset of them. We assume, however, that the police officer will use a *single-dimensional index* $\theta \in [0,1]$ that summarizes all of the information that these characteristics indicate about the likelihood that a driver may be carrying contraband. We assume that, if a driver of race $r_m \in \{M, W\}$ actually carries contraband, then the index θ is randomly drawn from a continuous probability density distribution $f_g^{r_m}(\cdot)$; if a race r_m driver does not carry contraband, θ would be randomly drawn from $f_n^{r_m}(\cdot)$.[11] Without loss of generality, we can assume that the two densities $f_g^{r_m}$ and $f_n^{r_m}$ satisfy the strict monotone likelihood ratio property (MLRP), i.e., for $r_m \in \{M, W\}$,

MLRP: $f_g^{r_m}(\theta)/f_n^{r_m}(\theta)$ is strictly increasing in θ.

The MLRP property on the signal distributions essentially means that a higher index θ is a

[7] In the empirical part of the paper, we will examine three racial or ethnic groups: whites, blacks, and Hispanics. For now, though, we group blacks and Hispanics together as minorities for ease of exposition.

[8] For the purpose of deriving our empirical test, we will assume that π^{r_m} is exogenous. For an equilibrium model in which π^{r_m} is endogenously determined, see Appendix A.

[9] A trooper must first stop the motorist prior to a search. Examining the possibility of racial prejudice in highway stops is beyond the scope of this paper. In our analysis, we will take the sample of cars that are stopped as our population and focus solely on determining racial prejudice in troopers' search decisions. The presence, or lack therof, of racial prejudice at the stop level should not affect our conclusions.

[10] The questions the trooper will ask the motorist are typically focused on where the motorist is headed and the purpose of their visit. In listening to the response the trooper will try to discern how nervous or defensive the motorist is, and how logical the motorist's response is.

[11] The subscripts g and n stand for "guilty" and "not guilty," respectively.

signal that a driver is more likely to be guilty.[12] To the extent that there may be obviously guilty drivers (for example, if illicit drugs are in plain view), we assume that:

Unbounded Likelihood Ratio: $f_g^{r_m}(\theta)/f_n^{r_m}(\theta) \to +\infty$ as $\theta \to 1$.

The MLRP also implies that the cumulative distribution function $F_g^{r_m}(\cdot)$ first order stochastically dominates $F_n^{r_m}(\cdot)$, which implies that drivers who actually carry contraband are more likely to generate higher and thus more suspicious signals. We think this single dimensional index formulation summarizes the information that is available to troopers when they make their search decisions on the highway in a simple but realistic manner.

Each police officer can choose to search a vehicle after observing the driver's vector (r_m, θ), where r_m is the driver's race and θ is the single-dimensional index that summarizes all other characteristics observed during the stop. We assume that a trooper wants to maximize the total number of convictions (or the number of drivers found carrying illicit contraband) minus a cost of searching cars.[13] This is an important assumption because it requires that police officers always use any statistical information contained in the race of the motorist in their search decisions.[14]

Let $t(r_m; r_p)$ be the cost of a police officer with race r_p searching a motorist with race r_m, where $r_p, r_m \in \{M, W\}$. We normalize the benefit of each arrest (or successful drug find) to equal one, and scale the search cost to be a fraction of the benefit, so that $t(r_m; r_p) \in (0,1)$ for all r_m, r_p. It is worth emphasizing that, different from KPT, we allow the troopers' cost of searching a vehicle to depend on the races of both the motorist and the officer, and thus we can directly confront the possibility that police officers may not be monolithic in their search behavior.

Let G denote the event that the motorist searched is found with illicit drugs in the vehicle. When a police officer observes a motorist of race r_m and signal θ, the posterior probability that such a motorist may be guilty of carrying contraband, $\Pr(G|r_m, \theta)$, is obtained via Bayes' rule:

$$\Pr(G|r_m, \theta) = \frac{\pi^{r_m} f_g^{r_m}(\theta)}{\pi^{r_m} f_g^{r_m}(\theta) + (1 - \pi^{r_m}) f_n^{r_m}(\theta)}.$$

[12]For any one dimensional index θ, we can always reorder them according to their likelihood ratio $f_g^{r_m}(\theta)/f_n^{r_m}(\theta)$ in an ascending order. Thus the MLRP assumption is with no loss of generality.

[13]This is also the police objective postulated in KPT. It is a plausible assumption because awards (such as Trooper of the Month honors) and/or promotion decisions are partly based on troopers' success in catching motorists with contraband.

[14]This assumption rules out the possibility that some officers ignore the race of a motorist even when it provides useful information.

It immediately follows from the MLRP that $\Pr(G|r_m,\theta)$ is monotonically increasing in θ. From the unbounded likelihood ratio assumption, we know that $\Pr(G|r_m,\theta)\to 1$ as $\theta\to 1$.

The decision problem faced by a police officer of race r_p when facing a motorist with race r_m and signal θ is thus as follows:

$$\max\left\{\Pr(G|r_m,\theta)-t(r_m;r_p);0\right\}$$

where the first term is the expected benefit from searching such a motorist and the second term is the benefit from not searching, which is normalized to zero. Thus the optimal decision for a trooper of race r_p is to search a race-r_m motorist with signal θ if and only if

$$\Pr(G|r_m,\theta)\geq t(r_m;r_p).$$

From the monotonicity of $\Pr(G|r_m,\theta)$ in θ, we thus conclude:

Proposition 1 *A race-r_p police officer will search a race-r_m motorist if and only if*

$$\theta\geq\theta^*(r_m;r_p)$$

where $\theta^*(r_m;r_p)$ is uniquely determined by

$$\Pr(G|r_m,\theta^*(r_m;r_p))=t(r_m;r_p).$$

Moreover, the search threshold $\theta^*(r_m;r_p)$ is monotonically increasing in $t(r_m;r_p)$.

Proposition 1 says that the probability of a successful search for the marginal motorist is equal to the cost of search. Any infra-marginal motorist will have a higher search success probability. In what follows, we will refer to $\theta^*(r_m;r_p)$ as the *equilibrium search criterion* of race-r_p police officers against race-r_m motorists. We define the *equilibrium search rate* of race-r_p police officers against race-r_m motorists as $\gamma(r_m;r_p)$, which is given by

$$\gamma(r_m;r_p)=\pi^{r_m}\left[1-F_g^{r_m}(\theta^*(r_m;r_p))\right]+(1-\pi^{r_m})\left[1-F_n^{r_m}(\theta^*(r_m;r_p))\right]. \tag{1}$$

The *equilibrium average search success rate* of race-r_p police officers against race-r_m motorists, denoted by $S(r_m;r_p)$, is given by

$$S(r_m;r_p)=\frac{\pi^{r_m}\left[1-F_g^{r_m}(\theta^*(r_m;r_p))\right]}{\pi^{r_m}\left[1-F_g^{r_m}(\theta^*(r_m;r_p))\right]+(1-\pi^{r_m})\left[1-F_n^{r_m}(\theta^*(r_m;r_p))\right]}. \tag{2}$$

We now introduce three definitions. First, a police officer of race r_p is defined to be *racially prejudiced* if he or she exhibits a preference for searching motorists of one race. Following KPT, we model this preference in the cost of searching motorists.

Definition 1 *A police officer of race r_p is racially prejudiced, or has a taste for discrimination, if $t(M; r_p) \neq t(W; r_p)$.*

Next, we say that police do not exhibit *monolithic behavior* if officers of different races do not use the same search criterion when dealing with motorists of some race.

Definition 2 *The police officers do not exhibit monolithic behavior if $t(r_m; M) \neq t(r_m; W)$ for some $r_m \in \{M, W\}$.*

Note that a monolithic police force does not mean that they are not racially prejudiced: it could be that police officers of both races are equally prejudiced against some race of motorists. Likewise, a non-monolithic police force does not necessarily imply that some racial group of troopers are racially prejudiced: it could be that each group of troopers has the same search cost against all groups of motorists, but that search costs depend on the race of the trooper.

Finally, we say that race-r_p police officers exhibit *statistical discrimination* if they have no taste for discrimination and yet they use different search criterion against motorists with different races.

Definition 3 *Assume $t(M; r_p) = t(W; r_p)$. Then race-r_p police officers exhibit statistical discrimination if $\theta^*(M; r_p) \neq \theta^*(W; r_p)$.*

Officers will choose to use statistical discrimination if the distribution of the signal θ among white and minority motorists is different. When these distributions differ and $t(M; r_p) = t(W; r_p)$ (as assumed), Proposition 1 implies that the race-r_p police will choose search criteria $\theta^*(M; r_p)$ and $\theta^*(W; r_p)$ so that the marginal search success rates against white and minority motorists are both equal to the search cost. This typically implies that $\theta^*(M; r_p) \neq \theta^*(W; r_p)$. One reason why the distribution of the signal θ might be different across motorists of different races is that one group might be more likely to carry contraband. For example, if minority drivers are more likely to carry contraband ($\pi^W < \pi^M$), then it will be optimal for a non-prejudiced officer to search relatively more minority drivers (assume everything else is the same for white and minority drivers), and thus they will set $\theta^*(M; r_p) < \theta^*(W; r_p)$. Another reason why the distribution of θ might be different for whites and minorities is that $f_g^{r_m}(\theta)$ and $f_n^{r_m}(\theta)$ can differ between motorist races. For example, minority drivers not carrying contraband might tend to be more nervous during a stop than whites.[15]

Now we derive some simple implications of the model that will serve as the basis of our empirical test. First, note that if police officers are monolithic, then the cost of searching any given race of motorists is the same, regardless of the race of the officer. That is, $t(W;W) = t(W;M)$ and $t(M;W) = t(M;M)$. If we assume that white and minority troopers face the same population of white motorists and the same population of minority motorists, then Proposition 1 implies that both races of officers will use the same search criterion against a given race of motorists, so that $\theta^*(W;W) = \theta^*(W;M)$ and $\theta^*(M;W) = \theta^*(M;M)$.[16] Thus following from the formula for the search rate (1) and average search success rate (2), we have:

Proposition 2 *If the police officers exhibit monolithic behavior, then* $\gamma(r_m;M) = \gamma(r_m;W)$ *and* $S(r_m;M) = S(r_m;W)$ *for all* $r_m \in \{M, W\}$.

Next, if *none* of the police officers are racially prejudiced, then it immediately follows from Definition 1 that the ranking of $t(r_m;M)$ and $t(r_m;W)$ does not depend on the motorist's race r_m, regardless of whether or not troopers are monolithic.[17] We can illustrate the implication of this using an example where white troopers find searching both minority and white motorists more costly than minority troopers do. More formally this can be written as $t(M;M) = t(W;M) < t(M;W) = t(W;W)$.[18] Because the search threshold given in Proposition 1 is monotonically increasing in $t(r_m;r_p)$ and both white and minority troopers face the same population of white

[15]This scenario is actually quite plausible. Because of the many documented bad past experiences minorities have faced with officers, they might have a stigma that all police officers are out to get them and thus might be very nervous during a stop even if they have not done anything wrong.

[16]We will discuss the validity of this assumption in Section 2.2.

[17]Consider, for illustrative purposes, the case that $t(W;M) < t(W;W)$. Since race-M officers are assumed not to be racially prejudiced, we have $t(W;M) = t(M;M)$. Similarly since race-W officers are not racially prejudiced, we have $t(W;W) = t(M;W)$. Thus it must be the case $t(M;M) < t(M;W)$. Thus $t(r_m;M) < t(r_m;W)$ for all r_m. Similar arguments show that if $t(W;M) > t(W;W)$, then we must have $t(M;M) > t(M;W)$; and if $t(W;M) = t(W;W)$ then we must have $t(M;M) = t(M;W)$. Thus the ranking of $t(r_m;M)$ and $t(r_m;W)$ does not depend on the motorist's race r_m.

[18]Note that the relationship $t(M;r_p) = t(W;r_p)$ does not imply that $\theta^*(M;r_p) = \theta^*(W;r_p)$, because troopers can be engaged in statistical discrimination.

and minority motorists, this implies that $\theta^*(M;M) < \theta^*(M;W)$ and $\theta^*(W;M) < \theta^*(W;W)$. Because the equilibrium search rate given in formula (1) is monotonically decreasing in $\theta^*(r_m;r_p)$, we immediately have that $\gamma(M;M) > \gamma(M;W)$ and $\gamma(W;M) > \gamma(W;W)$, so that race-$M$ officers' search rates will be higher for both races of motorists. Similarly, if $t(M;M) = t(W;M) > t(M;W) = t(W;W)$, then race-$M$ officers' search rates will be lower for both rates of motorists than race-W officers. Finally, if $t(M;M) = t(W;M) = t(M;W) = t(W;W)$, then race-$M$ officers' search rates will be equal to those of race-W officers for both races of motorists.

We can also show that if none of the police officers are racially prejudiced, then the rank order of average search success rates between white and minority troopers for any race of motorists should also be independent of the motorists' race. Recall the previous example where white troopers had a higher overall search cost than minority troopers. We showed this would imply that $\theta^*(M;M) < \theta^*(M;W)$ and $\theta^*(W;M) < \theta^*(W;W)$. The average search success rate with a search criterion θ^* against race-r_m motorist is simply

$$\frac{\pi^{r_m}\left[1 - F_g^{r_m}(\theta^*)\right]}{\pi^{r_m}\left[1 - F_g^{r_m}(\theta^*)\right] + (1 - \pi^{r_m})\left[1 - F_n^{r_m}(\theta^*)\right]},$$

and one can show that it is strictly increasing in θ^*.[19] Thus we have $S(W;M) < S(W;W)$ and $S(M;M) < S(M;W)$. That is, the ranking of $S(r_m;M)$ and $S(r_m;W)$ does not depend on r_m.

The above discussion is summarized in the following proposition:

Proposition 3 *If neither race-M nor race-W of police officers exhibit racial prejudice, then neither the ranking of $\gamma(r_m;M)$ and $\gamma(r_m;W)$ nor the ranking of average search success rates $S(r_m;M)$*

[19] To see this, note that it will be strictly increasing in θ^* if and only if

$$H(\theta^*) = \frac{1 - F_g^{r_m}(\theta^*)}{1 - F_n^{r_m}(\theta^*)}$$

is strictly increasing in θ^*. Note that

$$\begin{aligned} H'(\theta^*) &= \frac{-f_g^{r_m}(\theta^*)\left[1 - F_n^{r_m}(\theta^*)\right] + \left[1 - F_g^{r_m}(\theta^*)\right] f_n^{r_m}(\theta^*)}{\left[1 - F_n^{r_m}(\theta^*)\right]^2} \\ &= \frac{-f_g^{r_m}(\theta^*)\int_{\theta^*}^1 f_n^{r_m}(\theta)\,d\theta + f_n^{r_m}(\theta^*)\int_{\theta^*}^1 f_g^{r_m}(\theta)\,d\theta}{\left[1 - F_n^{r_m}(\theta^*)\right]^2} \\ &= \frac{\int_{\theta^*}^1 \left[f_n^{r_m}(\theta^*) f_g^{r_m}(\theta) - f_g^{r_m}(\theta^*) f_n^{r_m}(\theta)\right] d\theta}{\left[1 - F_n^{r_m}(\theta^*)\right]^2}. \end{aligned}$$

From MLRP, we know that, for all $\theta > \theta^*$,

$$\frac{f_g^{r_m}(\theta)}{f_n^{r_m}(\theta)} > \frac{f_g^{r_m}(\theta^*)}{f_n^{r_m}(\theta^*)},$$

thus the integrand in the numerator is always positive. Thus $H'(\theta^*) > 0$.

and $S(r_m; W)$ *depends on* $r_m \in \{M, W\}$. *Moreover, for any* r_m*, the ranking of* $\gamma(r_m; M)$ *and* $\gamma(r_m; W)$ *should be the exact opposite of the ranking of* $S(r_m; M)$ *and* $S(r_m; W)$.

In our model if race-r_p troopers are not racially prejudiced, we know that race-r_p troopers' *marginal* search success rate against white motorists will be equal to that against minority motorists. But because in our model the marginal motorist's guilty probability is smaller than that of the infra-marginal motorists, we can not conclude that race-r_p troopers' *average* search success rate against white motorists will be equal to that against minority motorists. This is in stark contrast to KPT's model where there is no distinction between marginal and average motorists. Nonetheless, Proposition 3 provides testable implications of our model based on rank orders of observable statistics – the search rates and the average search success rates.

The contrapositive of Proposition 3 is simply that, if the ranking of $\gamma(r_m; M)$ and $\gamma(r_m; W)$, or the ranking of $S(r_m; M)$ and $S(r_m; W)$, depend on r_m, then *at least one racial group* of the troopers exhibit racial prejudice. Without further assumptions, it is not possible to determine which group of troopers are racially prejudiced.

2.1 Empirical Tests

2.1.1 Test for Monolithic Trooper Behavior

Proposition 2 suggests a test for whether troopers of different races exhibit monolithic search behavior that is implementable even when researchers have no access to the signals θ observed by troopers in making their search decisions. Under the null hypothesis that police officers exhibit monolithic behavior, then, for any race of drivers, the search rates and average search success rates against drivers of that race should be independent of the race of the troopers that conduct the searches. That is, under the null hypothesis of monolithic trooper behavior, we must have, for all $r_m \in \{M, W\}$,

$$\gamma(r_m; M) = \gamma(r_m; W), \tag{3}$$

$$S(r_m; M) = S(r_m; W). \tag{4}$$

Any evidence in violation of any of these equalities would reject the null hypothesis.

It is worth pointing out that both equalities (3) and (4) hold if and only if the null hypothesis is true. To illustrate why this is true we need to show that when the null hypothesis is not true we will never satisfy equality (3) and (4). Without loss of generality, suppose that troopers are not monolithic in their search behavior against white motorists ($r_m = W$). That is, $t(W; W) \neq t(W; M)$. If $t(W; W) > t(W; M)$, then, because both white and minority troopers face the same population of white motorists, we know from Proposition 1 that $\theta^* (W; W) > \theta^* (W; M)$, i.e. white troopers will use a more strict search criterion than minority troopers when searching white motorists. This then simultaneously implies that $\gamma(W; W) < \gamma(W; M)$ and that $S(W; W) > S(W; M)$, following from the proof in footnote 19. Thus the test using either (3) and (4) has an asymptotic power of one.

Moreover, the relationship between search rates and average search success rates suggests that, in principle, our model can be refuted. According to our model, whenever $\gamma(W; W) < \gamma(W; M)$, this must be because $\theta^* (W; W) > \theta^* (W; M)$ which directly implies that $S(W; W) > S(W; M)$. Thus if the rank order between the search rates between racial groups of troopers for a given race of motorists is not exactly the opposite of the rank order between the average search success rates, then we know that at least some of the conditions of our model are not satisfied.[20]

2.1.2 Test for Racial Prejudice

Proposition 3 suggests a test for whether some racial groups of troopers exhibit racial prejudice in their search behavior. Under the null hypothesis that none of the racial groups of troopers have racial prejudice, it must be true that both the ranking of search rates for a given race of motorists r_m across the races of troopers $\gamma(r_m; M)$ and $\gamma(r_m; W)$, and the ranking of average search success rates $S(r_m; M)$ and $S(r_m; W)$, do not depend on $r_m \in \{M, W\}$. The null hypothesis will be rejected if the ranking of $\gamma(r_m; M)$ and $\gamma(r_m; W)$, or the ranking of $S(r_m; M)$ and $S(r_m; W)$, depends on the race of the motorists r_m.

This test, however, has an asymptotic power less than one. That is, one may fail to reject the null hypothesis even when it is false. To see this, suppose that the truth is $t(M; M) = t(W; M) < t(M; W) < t(W; W)$. That is, race-$M$ officers are not racially prejudiced, but race-W officers are prejudiced against minorities (race-W officers' cost of searching minority motorists are smaller). In this case, race-W officers will apply higher search criteria toward both races of motorists, and thus the race-W officers' search rates will be lower regardless of the race of the motorists. Thus the null

[20]Of course, if the search rates between racial groups of troopers for a given race of motorists are equal, then the average search success rates between racial groups of troopers for a given race of motorists must also be equal.

would not be rejected even it is false and we commit a type-II error. This is a clear weakness of this test. On the other hand, if we do find evidence against the null hypothesis, we are confident that at least one racial group of troopers is racially prejudiced.[21]

Now we relate our test of racial prejudice to the test proposed in Antonovics and Knight (2004). As we described in the introduction, they use evidence that police officers are more likely to conduct a search if the race of the officer differs from the race of the driver as evidence of racial prejudice. First, it is useful to point out that their test is different from our rank order test proposed above. Consider the following simple example. Suppose that $r_m, r_p \in \{W, M\}$ and let the search rates be as follows: $\gamma(M; M) = .05, \gamma(W; M) = .10, \gamma(M; W) = .20$ and $\gamma(W; W) = .15$. That is, minority officers are more likely to search white motorists than minority motorists, and white officers are more likely to search minority motorists than white motorists. Thus officers in this example are more likely to conduct a search if the race of the motorist is different from their own, causing Antonovics and Knight's test to conclude that racial prejudice is occurring. However, such patterns of search rates satisfy our rank independence condition, that is, $\gamma(r_m; W) > \gamma(r_m; M)$ for $r_m \in \{W, M\}$, and thus our test would not consider this as evidence of racial prejudice. Antonovics and Knight's inference is not justified in our theoretical model without making further assumptions on the signal distributions $f_g^{r_m}$ and $f_n^{r_m}$.

2.2 Discussion of Two Key Assumptions

We made two key assumptions in the description of the model that play important roles in our empirical methodology.

Assumption on the Pool of Motorists Faced by Troopers of Different Races. In the model, we assume that the fraction of race-r_m motorists carrying contraband $\pi^{r_m} \in (0, 1)$ does not depend on the race of the troopers searching them. That is, we assumed that the pools of motorists faced by troopers of different races are the same. This assumption may not be empirically valid if white and minority troopers are systematically assigned to patrol in different locations or time of day (indeed, our raw data indicated that this is the case, see Tables 3 and 4).

[21] If we were to willing to assume that the signal distributions $f_g^{r_m}$ and $f_n^{r_m}$ do not depend on r_m, then one can derive more powerful tests for racial prejudice. But we think such restrictions are too strong to be realistic in empirical applications.

We now propose an empirical method that can resolve this problem even when the raw data does not satisfy this condition. For illustration purposes, suppose that there are two troop stations 1 and 2, each with 100 officers. Suppose that in troop station 1, 80 officers are white and 20 are minorities; in station 2, 60 officers are white and 40 are minorities. Thus, on average 70 percent of the troopers are white and 30 percent are minorities. If the motorists that drive through the patrol areas of stations 1 and 2 differ in their characteristics, then the assumption that on average white and minority troopers face the same pool of motorists may be invalid. To deal with this issue we create artificial samples in the following way. We keep all the minority officers (20 of them) in station 1, but randomly select 47 out of the 80 white officers. Similarly, we keep all the white officers (60 of them) in station 2, but randomly select 26 out of the 40 minority officers. Thus we create an artificial sample of 107 white officers and 46 minority officers. Among the 153 officers in the artificial sample, (roughly) 70 percent of them are whites and 30 percent are minorities, and they are equally likely to be assigned to stations 1 and 2. We can calculate the various search rates and average search success rates in this artificial sample. To alleviate the sampling error, we use independent resampling to create a list of such artificial data sets.

This resampling method can effectively ensure that, when we calculate the search rates and average search success rates, the white and minority officers in the sample are assigned to different trooper stations with equal probability. Thus on average, white and minority officers are facing the same pool of motorists.

Assumption on the Signal Distributions. In the model we allow the signal distributions $f_g^{r_m}$ and $f_n^{r_m}$ to be specific to the racial group of the drivers. This flexibility is important if we intend to use our model as a basis for empirical test. As explained earlier, black and white drivers may exhibit different characteristics in their encounters with highway troopers, and thus imposing f_g^M and f_n^M to be equal to f_g^W and f_n^W, respectively would be very strong and may be empirically implausible. Also note that, since θ is most likely not observable by researchers, we do not want to impose parametric distributional assumptions.

Despite this flexibility, our formulation does assume that the signals of race r_m motorists are drawn from the same distributions independent of police officers' race. For example, we do not allow for the possibility that minority drivers will present a signal that is drawn from one distribution when they are stopped by a minority trooper and another signal that is drawn from a different distribution when they are stopped by a white trooper. This would be a suspicious assumption, for example, if the stops and searches occur in *local* streets. As argued in Donohue and Levitt (2001), a black community may be more willing to cooperate with a black officer, and thus black officers may obtain more information about a black motorist on the streets. However, we maintain that this is a realistic assumption in highway searches. When stopping a black driver on highways, a trooper typically does not have any other citizens to rely on for additional information. Thus any informational advantage that black troopers have about black motorists may not be applicable on the highways. Thus as long as white and black troopers observe the same list of characteristics and summarize them in the same way, this is a valid assumption.

One may also argue that minority drivers might be more nervous with white officers than they are with minority officers, regardless of whether or not they are carrying contraband. But as long as white officers properly take this fact into account, they should put a lower weight on the observed nervousness from a black motorist when they formulate the signal index θ. Thus this argument does not necessarily invalidate our assumption that $f_g^{r_m}$ and $f_n^{r_m}$ do not depend on the race of the police officers r_p.

3 Empirical Results

3.1 Data Description

We now apply the tests described above to data from the Florida State Highway Patrol. The Florida data is composed of two parts. The first is the *traffic* data set that consists of all the stops and searches conducted on all Florida highways from January 2000 to November 2001. For each of the stops in the data set, it includes (among other things) the date, exact time, county, driver's race, gender, ethnicity, age, reason for stop, whether a search was conducted, rationale for search, type of contraband seized, and the ID number of the trooper who conducted the stop and/or search. This part of the data is similar to those used in earlier studies of racial profiling (e.g. KPT 2001 and Gross and Barnes 2002).[22] The unique feature of our data set is the second part, which is the *personnel* data that contains information on each of the troopers that conducted the stops and searches in the traffic data set, including their ID number, date of birth, date of

[22]Even though KPT have data on the stops, they did not use them in their analysis. Gross and Barnes (2002) provided some basic statistics about the stop data.

hiring, race, gender, rank, and base troop station. We merge the traffic data and the personnel data by the unique trooper ID number that appears in both data sets. The merged data set thus provides information about the demographics of the trooper that made each stop and search. After eliminating cases in which there was missing information on the demographics of the trooper that conducted the stop, we end up with 906,339 stops and 8,976 searches conducted by a total of 1,469 troopers.[23] Florida State Highway Patrol troopers are assigned to one of ten trooper stations. Except for trooper station K, which is in charge of the Florida Turnpike, all other stations cover fixed counties. Figure 1 shows the coverage area of different troop stations.

3.2 Empirical Findings

3.2.1 Descriptive Statistics

Table 1 summarizes the means of the variables related to the motorists in our sample. Of the $906,339$ stops we observe, 66.5 percent were carried out against white motorists, 17.3 percent against Hispanic motorists, and 16.2 percent against blacks. In all race categories of the motorists, male motorists account for at least 67 percent of the stopped motorists for all race categories. Among all the motorists that were stopped, 48 percent were in the 16-30 age group, 33.6 percent were in the 31-45 age group and 18.3 percent were 46 and older. Close to 90 percent of stopped motorists have in-state license plates, and close to 70 percent of the stops were conducted in the day time (defined to be between 6am and 6pm).

Of the 8,976 searches we observe, 54.6 percent were performed on white motorists, 23.4 percent on Hispanic motorists, and 22.1 percent against blacks. In all race categories, more than 80 percent of searches were performed on male motorists, and overall, 84.8 percent of searches were against male drivers. Among the motorists that were searched, 58.4 percent were in the 16-30 age group, 31.7 percent were in the 31-45 age group and only 9.9 percent were in the 46 and older age group. Vehicles with in-state plates account for 85.7 percent of the searches, and 52.5 percent of the searches were conducted at night (recall 30.3 percent of the stops were at night). 79.2 percent of searches were not successful (they yielded nothing). Drugs were the most common contraband seized in successful searches (15.1 percent of total searches), followed by alcohol/tobacco (2.1 percent) and

[23] We also eliminated cases where the race of the motorist and trooper was not either white, black, or Hispanic, since there are not enough observations of the other racial groups to consider them.

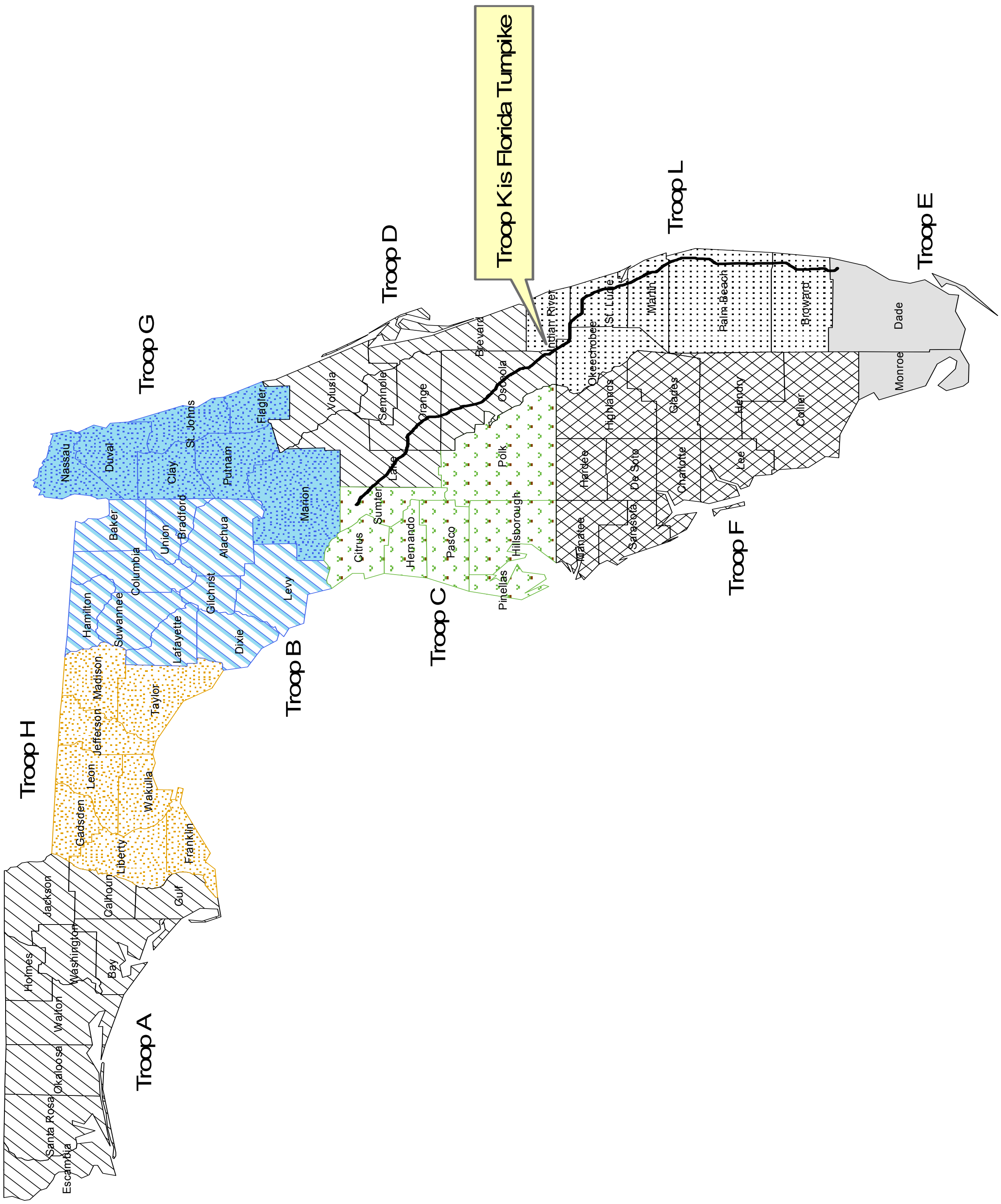

Figure 1: Troop Station Coverage Map

Motoristsʼ Characteristics	Stops: All Stops	Stops: By Motorist Sex: Female	Stops: By Motorist Sex: Male	Searches: All Searches	Searches: By Motorist Sex: Female	Searches: By Motorist Sex: Male
Black	.162 (.368)	.327 (.470)	.673 (.470)	.221 (.415)	.146 (.354)	.851 (.354)
Hispanic	.173 (.378)	.225 (.417)	.775 (.471)	.234 (.423)	.098 (.296)	.902 (.296)
White	.665 (.472)	.319 (.466)	.681 (.466)	.546 (.498)	.178 (.382)	.822 (.382)
Female	.304 (.460)	1.00 (.00)	0.00 (.00)	.152 (.359)	1.00 (.00)	0.00 (.00)
Male	.696 (.460)	0.00 (.00)	1.00 (.00)	.848 (.359)	0.00 (.00)	1.00 (.00)
Age:						
16-30	.481 (.500)	.325 (.468)	.675 (.468)	.584 (.493)	.149 (.356)	.851 (.356)
31-45	.336 (.472)	.295 (.456)	.705 (.456)	.317 (.465)	.162 (.368)	.838 (.368)
46+	.183 (.386)	.269 (.444)	.731 (.444)	.099 (.299)	.136 (.343)	.864 (.343)
License Plate:						
In-state	.899 (.302)	.310 (.462)	.690 (.462)	.857 (.350)	.155 (.362)	.845 (.362)
Out-of-state	.101 (.302)	.252 (.434)	.748 (.434)	.143 (.350)	.132 (.338)	.868 (.338)
Time:						
Day (6am-6pm)	.697 (.459)	.316 (.465)	.684 (.465)	.475 (.499)	.161 (.367)	.839 (.367)
Night	.303 (.459)	.275 (.447)	.725 (.447)	.525 (.499)	.144 (.351)	.856 (.351)
Contraband Seized:						
None				.792 (.406)	.155 (.362)	.845 (.362)
Drugs				.151 (.358)	.137 (.344)	.863 (.344)
Paraphernalia				.015 (.122)	.156 (.364)	.844 (.364)
Currency				.003 (.051)	.174 (.388)	.826 (.388)
Vehicles				.010 (.100)	.154 (.363)	.846 (.363)
Alcohol/Tobacco				.021 (.142)	.151 (.359)	.849 (.359)
Weapons				.006 (.078)	.055 (.229)	.945 (.229)
Other				.003 (.049)	.318 (.477)	.682 (.477)
Number of Observations:	906,339	275,527	630,812	8,976	1,364	7,612

Table 1: Means of Variables Related to Motorists.

Note: Standard errors of the means are shown in parentheses.

drug paraphernalia (1.5 percent).

Table 2 summarizes the means of variables related to the troopers in our sample. The first column shows that in our data, Blacks, Hispanic and whites account for 13.7, 10, and 76.3 percent of the troopers respectively. 89 percent of the troopers are male. The second and the third columns show that white troopers conducted 73 percent of all stops and 86 percent of all searches. The corresponding numbers for black troopers are 16 and 4.6 percent; for Hispanic troopers they are 11.4 and 9.5 percent. Female troopers conducted 9.3 percent of all stops and 6.9 percent of all searches.

3.2.2 Examining the Assumption that Troopers Face the Same Population of Motorists

Before we conduct our tests of monolithic behavior and racial prejudice we first examine whether a crucial assumption of our test, that all troopers face the same population of motorists, are satisfied in the raw data (before resampling). This assumption, of course, is not directly testable, because $\pi^{r_m}, f_g^{r_m}(\theta), f_n^{r_m}(\theta)$ and θ are all unobservable. The best we can do is to examine the distribution of observable motorist characteristics faced by troopers of different races. Table 3 shows the proportions of stopped motorists with given characteristics faced by troopers of different races. The characteristics of motorists reported in the table include race, gender, age, and time of the stops. For each row, we also report in the last column the p-values for Pearson χ^2 tests of the null hypothesis that the proportions of stopped motorists with the characteristics specific to that row are the same for all three race groups of the troopers. As one can see, the hypothesis that troopers of different races face the same population of motorists can be statistically rejected in the raw data, even though the differences are numerically quite small. One may suspect that the reason that troopers of different races are stopping motorists with different characteristics is that Black, Hispanic and White troopers are assigned to different troops. For example, Hispanic troopers are likely to have an over-representation in Troop E (covering Miami in Dade County) relative to Troop A and H (covering counties in the Florida Panhandle). Indeed, Table 4 shows that the allocations of troopers of different races to different troops, and time of the assignment, do not seem random in the raw data. For this reason, we think it is important to conduct the

	Troopers	Stops			Searches		
Troopers'	All	All	By Trooper Sex		All	By Trooper Sex	
Characteristics	Troopers	Stops	Female	Male	Searches	Female	Male
Black	.137	.160	.115	.885	.046	.044	.956
	(.344)	(.366)	(.319)	(.319)	(.208)	(.206)	(.206)
Hispanic	.100	.114	.070	.930	.095	.025	.975
	(.300)	(.318)	(.256)	(.256)	(.293)	(.155)	(.155)
White	.763	.726	.092	.908	.859	.076	.924
	(.425)	(.446)	(.289)	(.289)	(.348)	(.265)	(.265)
Female	.106	.093	1.00	0.00	.069	1.00	0.00
	(.307)	(.291)	(.00)	(.00)	(.254)	(.00)	(.00)
Male	.894	.907	0.00	1.00	.931	0.00	1.00
	(.307)	(.291)	(.00)	(.00)	(.254)	(.00)	(.00)
Ranks:							
Captain	.022	.002	.239	.761	.002	.474	.526
	(.148)	(.041)	(.426)	(.426)	(.046)	(.513)	(.513)
Lieutenant	.070	.013	.023	.977	.007	.000	1.000
	(.255)	(.112)	(.151)	(.151)	(.081)	(.000)	(.000)
Sergeant	.145	.062	.054	.946	.053	.052	.948
	(.352)	(.241)	(.226)	(.26)	(.224)	(.223)	(.223)
Corporal	.147	.112	.068	.932	.071	.030	.970
	(.354)	(.316)	(.252)	(.252)	(.257)	(.170)	(.170)
LEO	.602	.810	.101	.899	.866	.073	.927
	(.490)	(.392)	(.301)	(.301)	(.341)	(.261)	(.261)

Table 2: Means of Variables Related to Troopers.

Note: Standard errors of the means are shown in parentheses.

Motorist's Race	Motorist's Characteristics	White Troopers	Black Troopers	Hispanic Troopers	p-value
White	Male	.679	.684	.701	<.001
	Night stops	.288	.272	.318	<.001
	Age: 16-30	.471	.460	.445	<.001
	Age: 31-45	.325	.341	.349	0.02
Black	Male	.671	.667	.686	<.001
	Night stops	.332	.308	.354	<.001
	Age: 16-30	.514	.514	.507	.001
	Age: 31-45	.340	.344	.356	0.03
Hispanic	Male	.783	.774	.761	<.001
	Night stops	.322	.288	.393	<.001
	Age: 16-30	.516	.497	.494	<.001
	Age: 31-45	.350	.363	.355	0.01

Table 3: Distribution of Characteristics of Stopped Motorists, by Trooper Race in the Raw Data.

resampling methods we described in Subsection 2.2.[24] By construction, in the artificial data we created with the resampling method, troopers of a given race are assigned to different troops with the same probabilities. The Pearson's χ^2 test also reveal that in the artificial sample troopers of different races are assigned to night shifts with the same probability. Thus we can maintain out hypothesis that the distribution of the observable characteristics of the stopped motorists faced by troopers are the same in the artificial sample. We report our test results below using data from the artificial samples.

[24]One may argue that all of the stops occurred on Florida highways, and the drug flow in Florida tends to go from Miami (a city in the southern tip of Florida) to cities in the northeastern United States; that is, drug couriers are moving throughout Florida (except for possibly the panhandle). Thus troopers stationed in different areas are likely to face similar population of drivers, and the differences in the stopped motorists' characteristics reflect the differences in stop behavior of the troopers of different races, rather than the differences in the driver population. It is plausible, but in this paper we take the stopped motorists population as given.

	Troopers' Race		
	White	Black	Hispanic
Troop			
A	.930 (.256)	.054 (.227)	.016 (.124)
B	.889 (.316)	.081 (.274)	.030 (.172)
C	.816 (.389)	.116 (.321)	.068 (.253)
D	.793 (.406)	.117 (.322)	.090 (.287)
E	.412 (.494)	.236 (.426)	.352 (.479)
F	.880 (.326)	.056 (.231)	.063 (.245)
G	.833 (.374)	.135 (.343)	.032 (.176)
H	.886 (.320)	.114 (.320)	0.00 (.00)
K	.698 (.461)	.147 (.355)	.155 (.364)
L	.603 (.491)	.298 (.459)	.099 (.300)
% Night Stops	.283 (.172)	.284 (.192)	.349 (.179)

Table 4: Proportion of Troopers with Different Races by Troop and Time Assignment in the Raw Data.

NOTE: Standard errors of the means are shown in parentheses

3.2.3 Test of Monolithic Behavior

We now implement our test for the hypothesis that troopers of different races exhibit monolithic behavior. Table 5 is the main table. In Panel A, we show the search rate given stop for motorist/trooper race pairs. For example, the first row shows that, of the white motorists stopped by white, black and Hispanic troopers, respectively 0.96, 0.27 and 0.76 percent of them were searched. The last column shows the p-value from the Pearson's χ^2 test under the null hypothesis that troopers of all races search white motorists with equal probability. Specifically, the Pearson's χ^2 test statistic under the null hypothesis all troopers with race in $\mathcal{R}$ search race-r_m motorists with equal probability is given by

$$\sum_{r_p \in \mathcal{R}} \frac{\left(\widehat{\gamma(r_m; r_p)} - \widehat{\gamma(r_m)}\right)^2}{\widehat{\gamma(r_m; r_p)}} \sim \chi^2 (R - 1),$$

where $\widehat{\gamma(r_m; r_p)}$ is the estimated search probability of race-r_p officers against race-r_m motorists, $\widehat{\gamma(r_m)}$ is the estimated search probability against race-r_m motorists unconditional on the race of the officer, and R is the cardinality of the set of troopers' race categories, $\mathcal{R}$.

Panel B presents the average search success rate for given motorist/trooper race pairs. The last column in each row shows the p-value from the Pearson's χ^2 test under the null hypothesis that troopers of all races have the same average search success rate against motorists of race in that specific row. Again the Pearson's χ^2 test statistics under the null hypothesis that all troopers with race in $\mathcal{R}$ have the same average search success rate against race-r_m motorists is given by

$$\sum_{r_p \in \mathcal{R}} \frac{\left(\widehat{S(r_m; r_p)} - \widehat{S(r_m)}\right)^2}{\widehat{S(r_m; r_p)}} \sim \chi^2 (R - 1),$$

where $\widehat{S(r_m; r_p)}$ is the estimated average search success rate of race-r_p officers against race-r_m motorists, and $\widehat{S(r_m)}$ is the estimated average search success rate against race-r_m motorists unconditional on the race of the officers.

As we argued in subsection 2.1.1, under the null hypothesis that troopers exhibit monolithic behavior, $\gamma(r_m; r_p) = \gamma(r_m)$ and $S(r_m; r_p) = S(r_m)$ for all r_p, and thus the Pearson's χ^2 test statistic should be small under the null. The p-values in Table 5 show that we can soundly reject the null hypothesis of monolithic trooper behavior.

Motorist's Race	Trooper Race: White	Trooper Race: Black	Trooper Race: Hispanic	p-value
	Panel A: Search Rate Given Stop (%)			
White	0.96 (6.68E-4)	0.27 (7.73E-4)	0.76 (9.26E-4)	< 0.001
Black	1.74 (1.30E-3)	0.35 (1.42E-3)	1.21 (2.28E-3)	< 0.001
Hispanic	1.61 (1.46E-3)	0.28 (0.76E-3)	0.99 (3.03E-3)	< 0.001
	Panel B: Average Search Success Rate (%)			
White	24.3 (9.43E-3)	39.4 (5.57E-2)	26.0 (2.28E-2)	< 0.001
Black	19.9 (1.26E-2)	26.0 (5.32E-2)	20.8 (2.67E-2)	< 0.001
Hispanic	8.5 (9.78E-3)	21.0 (4.55E-2)	14.3 (6.63E-2)	< 0.001

Table 5: Search Rates and Average Search Success Rates by Races of Motorists and Troopers in the Artificial Data Sets.

NOTE: Standard errors of the means are shown in parentheses.

3.2.4 Test for Racial Prejudice

We have so far provided strong evidence that troopers do not exhibit monolithic search criteria when deciding whether to search motorists of a given race. Now we describe the results from our test for racial prejudice as described in subsection 2.1.2. Under the null hypothesis that none of the racial groups of troopers are racially prejudiced, we argued that the rank order over the search rates $\gamma(r_m; W), \gamma(r_m; B)$ and $\gamma(r_m; H)$, and the rank order over the average search success rates $S(r_m; W), S(r_m; B)$ and $S(r_m; H)$, should both be independent of r_m. From the estimated mean search rates and average search success rates in Table 5, we have for all $r_m \in \{W, B, H\}$,

$$\begin{aligned} \gamma\widehat{(r_m; W)} &> \gamma\widehat{(r_m; H)} > \gamma\widehat{(r_m; B)}, \\ S\widehat{(r_m; W)} &< S\widehat{(r_m; H)} < S\widehat{(r_m; B)}. \end{aligned}$$

We can use simple Z-statistics to formally test that

$$\gamma(r_m; W) > \gamma(r_m; H) > \gamma(r_m; B), \tag{5}$$

$$S(r_m; W) < S(r_m; H) < S(r_m; B). \tag{6}$$

For example, let the null hypothesis be $\gamma(r_m; W) = \gamma(r_m; H)$. We can test it against the one-sided alternative hypothesis $\gamma(r_m; W) > \gamma(r_m; H)$ by using

$$Z = \frac{\gamma\widehat{(r_m; W)} - \gamma\widehat{(r_m; H)}}{\sqrt{\frac{\mathrm{SVar}_W}{n_W} + \frac{\mathrm{SVar}_H}{n_H}}}$$

where n_W and n_H are the number of stops conducted by white and Hispanic officers respectively against race-r_m motorists, and SVar_W and SVar_H are respectively the sample variances of the search dummy variables in the samples of stops against race-r_m motorists conducted by white and Hispanic officers. By the Central Limit Theorem (due to our large sample size), Z has a standard normal distribution under the null hypothesis. The null will be rejected in favor of the alternative at significance level α if $Z \geq z_\alpha$ where $\Phi(z_\alpha) = 1 - \alpha$. When $r_m = W$, the value of the Z-statistic is 27.4 under the null, thus we can reject it in favor of the alternative $\gamma(W; W) > \gamma(W; H)$ at significance level close to 0. Similarly, for the test of the null hypothesis $\gamma(W; H) = \gamma(W; B)$ against $\gamma(W; H) > \gamma(W; B)$, we obtain a Z-statistic of 65, thus again rejecting the null in favor of the alternative. Implementing this test to other races of motorists, we find that the evidence supports inequality (5).

We can use an analogous Z-test to formally test inequality (6) by using

$$Z' = \frac{\widehat{S(r_m;W)} - \widehat{S(r_m;H)}}{\sqrt{\frac{\mathrm{SVar}'_W}{n'_W} + \frac{\mathrm{SVar}'_H}{n'_H}}} \sim N(0,1), \tag{7}$$

where n'_W and n'_H are the number of searches against race-r_m motorists conducted by white and Hispanic officers respectively, and SVar'_W and SVar'_H are respectively the sample variances of the search success dummy variables in the sample of searches against race-r_m motorists conducted by white and Hispanic officers. The null will be rejected in favor of the alternative at significance level α if $Z' \leq -z_\alpha$ where $\Phi(z_\alpha) = 1 - \alpha$. For example when we consider white motorists, we obtain a Z-statistic of -324.1 for white and Hispanic officers, thus we are able to reject the null in favor of the alternative $S(W;W) < S(W;H)$ at a significance level essentially equal to 0. Likewise, we can reject the null $S(W;H) = S(W;B)$ in favor of the alternative $S(W;H) < S(W;B)$ at significance level close to 0 (with a Z-statistic of -254). Implementing this test to other races of motorists, we find that the evidence supports inequality (6).

To summarize, we cannot reject the null hypothesis that troopers are not racially prejudiced. Of course, we would like to emphasize caution in interpreting our finding: while we do not find definitive evidence of racial prejudice, it is still possible that some or all groups of troopers are racially prejudiced. If the latter is true, then we have committed a type-II error as a result of the weak test.

3.2.5 Other Implications from the Tests

It is interesting to note some additional implications from the tests we conducted above. First of all, inequality (5) implies that the search criterion used by troopers against race-r_m motorists have the ranking

$$\theta^*(r_m;W) < \theta^*(r_m;H) < \theta^*(r_m;B).$$

In light of Proposition 1, this implies a ranking over the search costs: for any r_m,

$$t(r_m;W) < t(r_m;H) < t(r_m;B).$$

That is, white troopers seem to have smaller costs of searching motorists of any race, followed by Hispanic troopers. Black troopers have the highest search costs.

Motorist's Race	Search Rate (%)	Average Search Success Rate (%)
White	0.81	25.1
	(.090)	(.434)
Black	1.35	20.9
	(.115)	(.407)
Hispanic	1.34	11.5
	(.115)	(.319)

Table 6: Average Search Success Rates by Race of Motorists in the Raw Data.

Note: Standard errors of the means are shown in parentheses.

Second, as we mentioned at the end of subsection 2.1.1, our model is refuted if, for each r_m, the rank order of the search rates against race-r_m motorists $\gamma(r_m;W), \gamma(r_m;B)$ and $\gamma(r_m;H)$ is not exactly the opposite of the rank order of the corresponding average search success rates $S(r_m;W), S(r_m;B)$ and $S(r_m;H)$. As we showed above, the statistical evidence in our data does not refute our model.

3.2.6 Replicating KPT's Test

Finally, we would like to contrast our findings with those from KPT's test. Recall that KPT's test relies on the prediction from their model that, under the null hypothesis of no racial prejudice, the average search success rates should be independent of the motorists' race. Table 6 shows the search rate and average search success rate for different races of the motorists in the *raw* data,[25] and Table 7 shows the p-values from Pearson's χ^2 test on the hypothesis that the search rates and average search success rates are equal across various race groupings. Their test immediately implies that the troopers show racial prejudice against black and Hispanic motorists, especially the Hispanics. However, as we argued, this conclusion is only valid if their model of motorist and trooper behavior is true.

[25]While KPT's model does make predictions of the search rate, their test does not utilize such information. In fact, they do not have the search rate information in their application to the Maryland data since their data consist of searches only. We include the search rate information in the tables for informative purposes only.

Groupings	Search Rate	Average Search Success Rate
White, Black, Hispanic	< 0.001	< 0.001
White, Black	< 0.001	< 0.001
White, Hispanic	< 0.001	< 0.001
Black, Hispanic	0.798	< 0.001

Table 7: p-Values from Pearson's χ^2 Tests on the Hypothesis that Search Rate and Average Search Success Rate are Equal Across Various Groupings.

4 Conclusion

Black and Hispanic motorists in the United States are much more likely than white motorists to be searched by highway troopers. Is this apparent racial disparity driven by racist preferences by the troopers, or by motives of effectiveness in interdicting drugs? Our paper presents a simple but plausible model of police search behavior, and we define racial prejudice, statistical discrimination and monolithic trooper behavior within the confines of our model. We then exploit the theoretical predictions from this model to design empirical tests that address the following two questions. Are police monolithic in their search behavior? Is racial profiling in motor vehicle searches motivated by troopers' desire for effective policing (statistical discrimination) or by their racial prejudice (racism)? Relative to the seminal research in Knowles, Persico and Todd (2001), our model allows troopers of different races to behave differently, thus allowing us to examine non-monolithic trooper behavior; moreover, our model does not yield, and the subsequent empirical test does not rely on, the convenient, but in our view unrealistic, implication that all drivers of the same race carry contraband with the same probability regardless of characteristics other than race, which is the vital prediction underlying their tests. We also propose a resampling method to deal with raw data sets where one of the major assumptions underlying our model and empirical tests is violated. Our tests require data sets with race information about both the motorists and troopers. When applied to vehicle stop and search data from Florida, our tests can soundly reject the null hypothesis that troopers of different races are monolithic in their search behavior, but fail to reject the null hypothesis that none of the racial groups of troopers are racially prejudiced. Finally we would like to emphasize that our test for racial prejudice is relatively conservative in that we may not always conclude there is racial prejudice when it is actually present. Although our test is a low-power one, which implies a high probability of type-II error will occur, the positive side of this is that when we do find evidence of racial prejudice it is rather conclusive.

A Appendix: A Model with Endogenous Drug Carrying Decisions.

In Section 2 we assumed that the proportion of motorists in race group r_m is exogenously given as $\pi^{r_m} \in (0,1)$. For the purpose of testing for monolithic behavior and racial prejudice, this partial equilibrium approach suffices. However, for other purposes such as public policy considerations like reducing crimes and the "war on drugs," one may want to know how any changes in trooper behavior may affect the motorists' drug carrying decisions.[26] One needs an equilibrium model to address such questions. In this appendix, we propose a simple model. We show that closing our partial equilibrium model in Section 2 is easy; moreover, such an equilibrium model has nice equilibrium uniqueness properties under reasonable conditions. This is in contrast to the labor market statistical discrimination models where multiple equilibria naturally arise and are the driving force for statistical discrimination (see, among others, Coate and Loury 1993).

Consider a single motorist race group r_m, and two trooper racial groups, r_p and r_p'.[27] Suppose that in the trooper population a fraction α is of race r_p and the remainder fraction $1-\alpha$ is of race r_p'. Suppose that Nature draws for each driver a utility cost of carrying contraband $v \in \mathbb{R}_+$ from CDF G with a continuous density. The utility cost v represents feelings of fear experienced by a driver from the act of carrying contraband. If a driver carries contraband and is not caught, he/she derives a benefit of $b > 0$. If a guilty driver is searched and thus arrested, he/she experiences an additional cost (over and above v) of c_g. If a driver does not carry contraband, he/she does not incur the utility cost of v. But the inconvenience experienced by an innocent driver when he/she is searched is denoted by c_n. Naturally we assume that $c_g > c_n$. We assume that a driver's realization of v is his or her private information; b, c_g and c_n are constants known to all drivers and police

[26] See Persico (2002) for an analysis on how racially blind search policies may affect the total crimes committed by motorists.

[27] Because we are only considering one race group of motorists, we will omit r_m from the subsequent notation. Having more than one racical groups of motorists will not change any of the results below.

officers. Each driver decides whether to carry contraband.

As before, we normalize the benefit of each arrest to the police officer to be one, and for notational simplicity, the cost of search for a race-r_p trooper is written as $t_p \in (0,1)$ and that for a race-r_p' trooper is $t_p' \in (0,1)$. As in Section 2, troopers observe noisy but informative signals regarding whether or not a driver is carrying contraband: if a driver is guilty, the signal $\theta \in [0,1]$ is drawn from PDF $f_g(\cdot)$; if the driver is not guilty, then θ is drawn from PDF $f_n(\cdot)$. As before f_g/f_n is strictly increasing in θ. Let F_g and F_n denote the corresponding CDFs of f_g and f_n. We assume that a trooper wants to maximize the total number of convictions minus the cost of searching cars.

We first suppose that a proportion π of drivers choose to carry contraband and analyze the optimal search behavior of the troopers. Let $\Pr(G|\theta)$ denote the posterior probability that a driver with signal θ is guilty of carrying illicit drugs, which is given by

$$\Pr(G|\theta,\pi) = \frac{\pi f_g(\theta)}{\pi f_g(\theta) + (1-\pi) f_n(\theta)}.$$

A race-r_p trooper will decide to search a driver with signal θ if and only if

$$\Pr(G|\theta,\pi) - t_p \geq 0;$$

which, from the MLRP, is equivalent to $\theta \geq \theta_p^*(\pi)$ where $\theta_p^*(\pi) \in [0,1]$ is the unique solution to

$$\Pr(G|\theta,\pi) = t_p.$$

Obviously $\theta_p^*(\pi)$ is strictly decreasing in π. Similarly, race-r_p' troopers will search a motorist if and only if the motorist's signal θ exceeds $\theta_{p'}^*(\pi)$ where $\theta_{p'}^*(\pi)$ solves

$$\Pr(G|\theta) = t_p'.$$

Now suppose that race-r_p and race-r_p' troopers use search criteria of θ_p^* and $\theta_{p'}^*$ respectively. The expected payoff of a driver with utility cost v from carrying contraband is given by

$$\overbrace{\left[\alpha F_g\left(\theta_p^*\right) + (1-\alpha) F_g\left(\theta_{p'}^*\right)\right] b}^{\text{Term 1}} - \overbrace{\left\{\alpha\left[1 - F_g\left(\theta_p^*\right)\right] + (1-\alpha)\left[1 - F_g\left(\theta_{p'}^*\right)\right]\right\} c_g}^{\text{Term 2}} - v$$

where Term 1 is the probability of not being caught multiplied by the benefit from drugs if the motorist is not caught. Note that a fraction α of the troopers are of race-r_p and use a search criterion of θ_p^*, and $1-\alpha$ of the troopers use $\theta_{p'}^*$. Thus the expected probability of not being caught is $\alpha F_g\left(\theta_p^*\right)+(1-\alpha)F_g\left(\theta_{p'}^*\right)$. Term 2 is the expected probability of being caught multiplied by the cost of being caught with illicit drugs. Of course, the driver suffers a disutility v whenever he or she carries drugs.

The expected payoff of a driver, whose utility cost is v, from not carrying contraband is simply the inconvenience cost of being searched by mistaken troopers:

$$-\left\{\alpha\left[1-F_n\left(\theta_p^*\right)\right]+(1-\alpha)\left[1-F_n\left(\theta_{p'}^*\right)\right]\right\}c_n.$$

Thus a driver with utility cost realization v will decide to carry illicit drugs if and only if $v\leq v^*\left(\theta_p^*,\theta_{p'}^*\right)$ where

$$\begin{aligned} v^*\left(\theta_p^*,\theta_{p'}^*\right) &= \left[\alpha F_g\left(\theta_p^*\right)+(1-\alpha)F_g\left(\theta_{p'}^*\right)\right]b-\left\{\alpha\left[1-F_g\left(\theta_p^*\right)\right]+(1-\alpha)\left[1-F_g\left(\theta_{p'}^*\right)\right]\right\}c_g \\ &\quad+\left\{\alpha\left[1-F_n\left(\theta_p^*\right)\right]+(1-\alpha)\left[1-F_n\left(\theta_{p'}^*\right)\right]\right\}c_n. \end{aligned} \tag{A1}$$

Thus if the troopers follow search criteria θ_p^* and $\theta_{p'}^*$ respectively, the proportion of drivers who will choose to carry contraband is given by $G\left(v^*\left(\theta_p^*,\theta_{p'}^*\right)\right)$.

An equilibrium of the model is a triple $\left(\pi,\theta_p^*,\theta_{p'}^*\right)$ such that:

$$\Pr\left(G|\theta_p^*,\pi\right) = t_p \tag{A2}$$

$$\Pr\left(G|\theta_{p'}^*,\pi\right) = t_{p'} \tag{A3}$$

$$G\left(v^*\left(\theta_p^*,\theta_{p'}^*\right)\right) = \pi \tag{A4}$$

The existence of equilibrium follows directly from Brouwer's Fixed Point Theorem. Now we show that in fact for any CDF G with non-negative support (i.e., $v\in\mathbb{R}_+$), the equilibrium is *unique*. Suppose that there are two equilibria in which the proportion of guilty motorists are π and $\tilde{\pi}$ with $\pi>\tilde{\pi}$. Observe from (A1) that $v^*(0,0)=c_n-c_g<0$ and

$$\begin{aligned} \frac{\partial v^*\left(\theta_p^*,\theta_{p'}^*\right)}{\partial\theta_p^*} &= \alpha c_n f_n\left(\theta_p^*\right)\left[\frac{f_g\left(\theta_p^*\right)}{f_n\left(\theta_p^*\right)}\frac{b+c_g}{c_n}-1\right], \\ \frac{\partial v^*\left(\theta_p^*,\theta_{p'}^*\right)}{\partial\theta_{p'}^*} &= \alpha c_n f_n\left(\theta_{p'}^*\right)\left[\frac{f_g\left(\theta_{p'}^*\right)}{f_n\left(\theta_{p'}^*\right)}\frac{b+c_g}{c_n}-1\right]. \end{aligned}$$

By the MLRP, we know that there exists $\left(\widehat{\theta_p^*},\widehat{\theta_{p'}^*}\right)\in[0,1)^2$ such that $v^*\left(\theta_p^*,\theta_{p'}^*\right)$ is strictly increasing in both θ_p^* and $\theta_{p'}^*$ when $\left(\theta_p^*,\theta_{p'}^*\right)>\left(\widehat{\theta_p^*},\widehat{\theta_{p'}^*}\right)$. Since $v^*(0,0)<0$ and the support of G is non-negative, we have $G\left(v^*(0,0)\right)=0$. Moreover, $G\left(v^*\left(\theta_p^*,\theta_{p'}^*\right)\right)$ will be zero for all

$\left(\theta_p^*, \theta_{p'}^*\right) \leq \left(\widehat{\theta_p^*}, \widehat{\theta_{p'}^*}\right)$. Thus any $\left(\theta_p^*, \theta_{p'}^*\right) \leq \left(\widehat{\theta_p^*}, \widehat{\theta_{p'}^*}\right)$ cannot be part of the equilibrium (because if $\pi = 0$, the optimal thresholds should be 1 from the troopers' best response). Thus in both equilibria of the model, we must have $\left(\theta_p^*, \theta_{p'}^*\right) > \left(\widehat{\theta_p^*}, \widehat{\theta_{p'}^*}\right)$ and $\left(\tilde{\theta}_p^*, \tilde{\theta}_{p'}^*\right) > \left(\widehat{\theta_p^*}, \widehat{\theta_{p'}^*}\right)$. That is, both equilibria lie in the region where $v^*(\cdot, \cdot)$ is strictly increasing in both arguments. If $\pi > \tilde{\pi}$, equilibrium conditions (A2) and (A3) imply that $\theta_p^* < \tilde{\theta}_p^*$ and $\theta_{p'}^* < \tilde{\theta}_{p'}^*$, therefore $0 < v^*\left(\theta_p^*, \theta_{p'}^*\right) < v^*\left(\tilde{\theta}_p^*, \tilde{\theta}_{p'}^*\right)$. But then it implies that $\tilde{\pi} > \pi$, a contradiction.

References

[1] Antonovics, Kate L. and Brian G. Knight (2004). "A New Look at Racial Profiling: Evidence from the Boston Police Department." Unpublished Manuscript, University of California, San Diego, June 15, 2004.

[2] Ayres, Ian (2001). *Pervasive Prejudice? Unconventional Evidence of Race and Gender Discrimination.* Chicago: University of Chicago Press.

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

[4] Becker, Gary S. (1957). *The Economics of Discrimination.* Chicago: University of Chicago Press.

[5] Coate, Stephen and Glenn C. Loury (1993). "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review,* **Vol.** 83, 1220-1240.

[6] Borooah, Vani K (2001). "Racial Bias in Police Stops and Searches: An Economic Analysis." *European Journal of Political Economy*, **Vol.** 17, 17-37.

[7] Dharmapala, Dhammika and Stephen L. Ross (2003). "Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence." Working Paper No. 2003-12, University of Connecticut.

[8] Donohue, John J. III and Steven D. Levitt (2001). "The Impact of Race on Police and Arrests." *Journal of Law and Economics*, **Vol.** XLIV, 367-394.

[9] Engel, Robin Shepard, Jennifer M. Calnon and Thomas J. Bernard (2002). "Theory and Racial Profiling: Shortcomings and Future Directions in Research." *Justice Quarterly*, **Vol.** 19, 249-273.

[10] Farrell, Amy, Jack McDevitt, and Deborah Ramirez (2000). "A Resource Guide on Racial Profiling Data Collection Systems: Promising Practicies and Lessons Learned." U.S. Department of Justice.

[11] Hernandez-Murillo, Ruben and John Knowles (2004). "Racial Profiling or Racist Policing?: Distinguishing between Competing Explanations for Discrimination." Forthcoming, *International Economic Review*.

[12] Knowles, John, Nicola Persico and Petra Todd (2001). "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy,* **Vol.** 109, 203-228.

[13] National Research Council (2004). *Measuring Racial Discrimination*. Panel on Methods for Assessing Discrimination, Rebecca M. Blank, Marilyn Dabady and Constance F. Citro, Editors.

[14] Persico, Nicola (2002). "Racial Profiling, Fairness, and Effectiveness of Policing." *American Economic Review*, **Vol.** 92, 1472-97.

[15] Riksheim, E.C. and S.M. Chermak (1993). "Causes of Police Behavior Revisited." *Journal of Criminal Justice*, **Vol.** 21, 353-382.

[16] Sherman L.W. (1980). "Causes of Police Behavior: The Current State of Quantitative Research." *Journal of Research in Crime and Delinquency*, **Vol.** 17, 69-100.