NBER WORKING PAPER SERIES

RACIAL BIAS IN MOTOR VEHICLE SEARCHES:
THEORY AND EVIDENCE

John Knowles
Nicola Persico
Petra Todd

Working Paper 7449
http://www.nber.org/papers/w7449

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1999

An earlier version of this paper was circulated as CARESS Working Paper #99-06. We thank Hanming Fang, John Ham, George Mailath, Andrew Postlewaite, Raphael Rob, Ken Wolpin and the participants to the University of Pennsylvania Empirical/Applied Theory seminars for helpful comments. We thank the Maryland ACLU for providing us with data and information. Persico and Todd are grateful to the NSF for support under #SBR-9905564 and #SBR-9730688, respectively. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Racial Bias in Motor Vehicle Searches: Theory and Evidence
John Knowles, Nicola Persico, and Petra Todd
NBER Working Paper No. 7449
December 1999
JEL No. J70, K42

ABSTRACT

African-American motorists in the United States are much more likely than white motorists to have their cars searched by police checking for illegal drugs and other contraband. The courts are faced with the task of deciding on the basis of traffic-search data whether police behavior reflects a racial bias. We discuss why a simple test for racial bias commonly applied by the courts is inadequate and develop a model of law enforcement that suggests an alternative test.

The model assumes a population with two racial types who also differ along other dimensions relevant to criminal behavior. Using the model, we construct a test for whether racial disparities in motor vehicle searches reflect racial prejudice, or instead are consistent with the behavior of non-prejudiced police maximizing drug interdiction. The test is valid even when the set of characteristics observed by the police is only partially observable by the econometrician. We apply the test to traffic-search data from Maryland and find the observed black-white disparities in search rates to be consistent with the hypothesis of no racial prejudice.

Finally, we present a simple analysis of the tradeoff between efficiency of drug interdiction and racial fairness in policing. We show that in some circumstances there is no trade-off; constraining the police to be color-blind may achieve greater efficiency in drug interdiction.

John Knowles
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
jknowles@econ.sas.upenn.edu

Nicola Persico
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
persico@econ.sas.upenn.edu

Petra Todd
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
petra@athena.sas.upenn.edu

1 Introduction

African-American motorists in the United States are much more likely than white motorists to have their cars searched by police checking for illegal drugs and other contraband. In the state of Maryland, for example, African Americans represented 63% of motorists searched by state police on the I-95 highway between January 1995 and January, 1999, but only represented 18% of motorists on the road.[1] While it is conceivable that African-American motorists are more likely to commit the types of traffic offenses that police use as pretexts for vehicle checks, traffic studies and police testimony suggest that blacks and whites are not distinguishable by their driving habits. An alternative explanation for the racial disparity in traffic searches is that race is one of the criteria police officers use in deciding whether to search cars. This explanation, known as "racial profiling," is the basis of several recent lawsuits against state governments. The issue has also attracted attention in political spheres, forcing the resignation of the New Jersey chief of police and provoking the U.S. President to describe racial profiling as a "morally indefensible, deeply corrosive practice."[2]

Evidence of racial profiling is often interpreted as an indication of racist preferences on the part of the police. The task of deciding whether racism is a factor in police traffic searches falls on the courts, which consider a variety of statistical evidence. The case for discrimination rests largely on the observation that the proportion of African-Americans among the drivers searched by police far exceeds the proportion in the general population of drivers. This simple comparison is the basis of expert witness testimonies in several legal cases.[3] A refined version of the test estimates the probability

[1] Based on our own tabulations and on Lamberth (1996), a study commissioned by the Maryland ACLU.

[2] "Clinton Order Targets Racial Profiling," *Associated Press*, June 9, 1999.

[3] In the 1993 case of *Maryland v. Wilkins*, for example, a statistician testified that

of being searched as a function of race and other observable characteristics thought to be related to criminal propensity. If race has no explanatory power in the regression, this is taken as evidence of no discrimination.[4]

The drawback of this type of test for discrimination is that it requires data on the full set of characteristics that a police officer uses in deciding whether to search a motorist.[5] If some characteristics are missing from the data, then race may have explanatory power due to omitted-variable bias. Even if race is found to be insignificant, there is still the possibility that police target individuals with certain characteristics because those characteristics are correlated with race and not because they are good predictors of criminality. Conditioning on those characteristics would naturally lead researchers to conclude in favor of the hypothesis of no racism. Thus, the validity of this procedure as a test for discrimination hinges crucially on judgments about what constitutes a set of admissible conditioning variables and on whether the analyst has access to the full set of variables.

Even if these types of tests find evidence of discrimination, they are not informative about the motivation for discrimination. Police may use race as a criterion in traffic stops because they are trying to maximize arrests and race helps predict criminality, or because they prefer stopping one racial group over another. We call the first type of situation *statistical discrimination*, using the terminology of Arrow (1973). An equilibrium exhibits statistical discrimination if police are not racially prejudiced and yet one

"The disparities are sufficiently great that, taken as a whole, they are consistent with and strongly support the assertion that the Maryland State Police are targeting the community of black motorists for stop, detention and investigation...".

[4]See Donohue (1999).

[5]A training manual issued by the Illinois State Police highlights some indicators of criminal activity. These include tinted windows, religious paraphernalia used to divert suspicion, and attorney's business cards. We report the complete list in Appendix A.

race is searched more often than another.[6] In contrast, we say that an officer is *racially prejudiced* if, cæteris paribus, he has a preference for searching motorists of a particular race. Thus, we model prejudice as a taste for discrimination, following Becker (1957). Prejudice is a property of the officer's utility function, while statistical discrimination is a property of equilibrium.

In this paper, we propose a test for distinguishing between statistical discrimination and racist preferences. This test looks at the success rate of searches across races, and is derived from a simple model of law enforcement via police searches. A key advantage of the test is that it is feasible even when the data comprise only a subset of the variables used by the police in deciding whether to search a motorist. In fact, while more variables allow for a more powerful test, the test that we propose can be carried out when race is the only characteristic observed.

Our model assumes that the police maximize the number of arrests, net of the cost of searching motorists.[7] An arrest is made when a motorist is searched and is found carrying contraband in his/her car. We assume that motorists take into account the probability of being searched in deciding whether to carry contraband. The police make their search decision on the basis of observable characteristics of the motorist, including race. Some of these characteristics may be informative about a driver's propensity to carry contraband. Prejudice is introduced into the model as a difference in the cost to the police of searching drivers of different groups. A key implication of the model is that if a police officer has the same cost of searching two subgroups of the population, and if these two subgroups are searched at equilibrium, then the returns from searching must be equal across these subgroups. For

[6]We present a formal definition of statistical discrimination in Section 2.

[7]An alternative model is that police maximize the number of arrests given a certain amount of resources (number of police officers, say). In terms of the implications tested in this paper, the two models are equivalent.

example, suppose searching one subgroup of motorists yielded a higher return. Then police would always search these motorists, and they would in turn react by carrying contraband less often, until the returns to searching are equalized across groups.

Notice that, if the returns to searching are equal across *all* subgroups distinguishable by police, they must also be equal across *aggregations* of these subgroups, which is what we can distinguish in the data. Thus, equality of the returns to searching can be tested without knowing all the characteristics observed by the police.

We test the implication that returns to searches are equal across groups using data on drug-related highway searches gathered by the Maryland State Police. The Maryland data attracted national attention when the ACLU asked that the police be found in contempt of court, using these data as evidence of racial discrimination in violation of the terms of the settlement of a 1993 class-action suit.[8] In our data, cars of African-American motorists are searched much more frequently than those of white motorists. However, the probability that a searched driver is found carrying contraband is very similar across races. Thus, we cannot reject the hypothesis that the disparity in the probability of being searched is due purely to statistical discrimination. In other words, it is possible to explain the black-white disparities observed in the Maryland motorist search data without recourse to racial prejudice.

Of course, effectiveness of drug interdiction is not the only relevant criterion in evaluating outcomes. An innocent driver may be stopped more often only because he belongs to a particular racial group. In the last section of the paper, we consider the costs that statistical discrimination imposes on inno-

[8]The class-action suit was filed by the ACLU on behalf of Robert L. Wilkins, an African-American attorney who was stopped and searched by the Maryland State Police. The ACLU contends that the data show a "continuing pattern of race discrimination in drug interdiction activities carried out along the I-95 corridor."

cent drivers and discuss the trade-offs between racial fairness in policing and efficiency.[9] We demonstrate that, in some circumstances, requiring police to disregard race can actually lead to an increase in efficiency of drug interdiction. This is because a race-blind search policy effectively pools in the same group motorists with different criminal propensities, subjecting them to the same probability of being searched. At equilibrium, this leads motorists with a high propensity to carry drugs to increase their drug-carrying activity, and low-propensity motorists to decrease it, so that the average level of drug-carrying in the population does not change relative to the equilibrium with no constraints on police behavior. We show that due to this effect, imposing a race-blind search policy may result in an outcome with the same level of drug-carrying and a lower cost of interdiction. This is interesting because it suggests that forcing police to adopt racially fair behavior may be desirable purely from the point of view of efficiency in fighting crime.

1.1 Legal Background

The judicial standpoint on racial profiling is not clear-cut. The courts, when confronted with racial profiling cases, have often concluded that race or ethnicity can be used as a factor in determining the likelihood that a person is engaging in or has committed a crime, as long as the use of race is reasonably related to law enforcement and is not a pretext for racial harassment (Kennedy, 1997). For example, in *United States v Weaver*, the U.S. Court of Appeals upheld the validity of airport searches of young black males prompted by suspicions that members of a Los Angeles black gang were bringing in cocaine.[10] In *United States v Martinez-Fuerte*, the U.S. Supreme

[9] Christopher Darden, the African-American prosecutor in the O.J. Simpson case, says that to survive traffic stops he "learned the rules of the game years before...Don't move. Don't turn around. Don't give some rookie an excuse to shoot you."

[10] *United States v Weaver*, 966 F.2d at 394, n.2.

Court upheld the legality of stopping cars at border checkpoints with occupants of Mexican ancestry on the grounds that these cars were more likely to be transporting illegal aliens.[11] Another case where the justification for the use of race is perhaps more questionable is *State v. Dean*, where the Arizona Supreme Court ruled that it was permissible to use race as a contributing factor in stopping and questioning a Mexican-American driver who seemed out of place in a predominantly white neighborhood.[12] In *United States v Nicholas*, however, the courts ruled it impermissible to stop a black driver with out-of-state plates, maintaining that this did not constitute sufficient grounds for inferring criminality.[13] In *United States v Laymon*, a judge suppressed incriminating evidence found in a vehicle, because he believed the police officer did not have sufficient justification for searching the car and had used race as a factor in the decision to search.[14]

Whether discrimination is deemed reasonable or not by the courts depends on assessments about the degree to which discrimination assists in apprehending criminals, the benefits of apprehending criminals and the costs imposed on people erroneously searched or detained.[15] In evaluating the legality of racial disparities in law enforcement, the courts have clearly sought

[11] *United States v Martinez-Fuerte,* 428 U.S. 543 (1976).

[12] Here the court found it important that race was only one of several factors taken into account. *State v. Dean*, 543 P.2d 425, 427 (Arizona, 1975).

[13] *United States v Nicholas,* 448, F.2d 622 (CA 8 1971).

[14] *United States v Laymon*, 730 F.Supp. 332 (D. Colo. 1990)

[15] Also, there is some debate in the courts over whether 'reasonableness' is an appropriate criterion to use in racial profiling cases. Pursuant to the Equal Protection Clause of the 14th Amendment, reasonableness is considered insufficient justification for government officials to discriminate on racial grounds. Instead, racially discriminatory government actions are to be subjected to a higher level of scrutiny than reasonableness (called *strict scrutiny*) and are to be upheld only if there is a compelling justification for the racial distinction that "is narrowly tailored to advance the project at hand." (See Kennedy, 1997)

to determine the motivation for discriminating. Sometimes discrimination motivated by efficiency reasons is considered permissible, whereas discrimination motivated by racial prejudice is never permissible. The standard regression-based test for discrimination described earlier is only informative about whether a racial disparity in car searches exists and is silent on the question of motivation. In contrast, the test we propose distinguishes between different motivations for racial disparities.

1.2 Related Literature

The theoretical model we develop belongs to the literature on optimal auditing. Early auditing models, such as Becker (1968) and Stigler (1970), examined citizens' incentives to misbehave under an exogenous probability of being audited. The more recent literature on optimal auditing, mainly dealing with income reporting and tax evasion, assumes that both parties, the auditor and the auditee, behave strategically (see Reinganum and Wilde (1986), Border and Sobel (1987), and Scotchmer (1987)). To the best of our knowledge, ours is the first paper that attempts an empirical test of an optimal auditing model.

In our model, if police are prejudiced, then at equilibrium the returns to searching members of the discriminated-against group are below average. The idea that a taste for racial discrimination leads to lower profits for the discriminators originated with Becker (1957) and is also discussed in Epstein (1992). The empirical link between profitability and discrimination has been investigated previously in the context of mortgage lending. Berkovec *et al.* (1998) find that loans granted to minorities perform worse than loans granted to nonminorities.[16]

[16] This evidence would be consistent with reverse discrimination. However, the degree to which this happens, i.e. how much worse minority loans perform, is found to be independent of the degree of competition of the loan markets. Since differences in loan

2 The Model

This section describes the model of police and motorist behavior that underlies our empirical work. We assume a continuum of police officers and motorists. Let $r \in \{A, W\}$ denote the race of the motorist, which is assumed to be observable by the police officer. Let c denote all characteristics other than race that are potentially used by the officer in the decision to search cars.[17] For expositional ease, we treat c as a one-dimensional variable. All the results in this section extend straightforwardly to the case where c is multidimensional. Let $F(c|W)$ and $F(c|A)$ denote the distribution of c in the white and African-American populations, respectively.

Police officers search motorists. We assume that each officer can choose from any type (c, r). Police maximize the total number of convictions minus a cost of searching cars. The marginal cost of searching a motorist of race r is denoted by t_r. Our model assumes that the benefit of each arrest is 1 so that the cost is scaled as a fraction of the benefit. A low t can be interpreted as society placing a high value on drug interdiction. To avoid trivial cases, assume $t_W, t_A \in (0, 1)$. Let G denote the event that the motorist searched is found guilty. (In our data, G corresponds to being found with drugs in the car).

We assume that motorists consider the probability of being searched in deciding whether to carry contraband. If they do not carry contraband, their payoff is zero whether or not the car is searched. If they carry contraband, their payoff is $-j(c, r)$ if they are searched and $v(c, r)$ if not searched. We can interpret $v(c, r)$ as the expected value of carrying drugs, and $j(c, r)$ as the expected cost of being convicted.[18] We assume that both $j(c, r)$ and

performance are not reduced by the degree of competition, the authors conclude that these differences do not reflect a taste for discrimination.

[17] The variable c may be unobserved or only partially observed by the econometrician.

[18] If there were discrimination in the court system leading to higher penalities for minor-

$v(c,r)$ are positive.

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

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

The motorist chooses to carry contraband if this expression is greater than zero. When the expression is zero, motorists are willing to randomize between carrying and not carrying. We denote the probability that a motorist of type c,r carries contraband by $P(G|c,r)$.[19]

The officer chooses the probability $\gamma(c,r)$ of searching each motorist of type c,r. The officer solves

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

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

Next, we introduce two definitions. First, a police officer is defined to be racially prejudiced if his/her utility function exhibits a preference for searching motorists of one race. We model this as a difference in the cost of searching motorists.

Definition 1 *We say that the police officer is racially prejudiced, or has a taste for discrimination, if* $t_A\neq t_W$.

ity drivers found with contraband, this could be thought of as operating through $j(c,r)$.

[19]We do not allow for the possibility of false accusation by police or planting of evidence, as do Donohue and Levitt (1998) in a different context from ours.

Next, we say that an equilibrium exhibits statistical discrimination if police officers have no taste for discrimination and yet the probability of being searched differs by race. Define the probability that a motorist of race r is searched as $\gamma(r) = \int \gamma(c, r)\, dF(c|r)$.

Definition 2 *Assume $t_A = t_W$. Then we say that an outcome exhibits statistical discrimination if $\gamma(W) \neq \gamma(A)$.*

An alternative definition of statistical discrimination would require that $\gamma(c, W) \neq \gamma(c, A)$ for some c, i.e. blacks are searched at different rates than whites *with the same observable characteristics* c. This definition is more stringent than Definition 2, in the sense that if $\gamma(c, W) \neq \gamma(c, A)$, then one expects that $\gamma(W) \neq \gamma(A)$. For our purposes, it is more convenient to use Definition 2.

2.1 Equilibrium

We construct an equilibrium where motorists randomize over whether to carry contraband and police officers randomize over whether to search them.[20] For a motorist of type c, r to be willing to randomize, expression (1) must equal zero. Solving for γ yields $\gamma^*(c, r) = v(c, r) / [v(c, r) + j(c, r)]$. This determines the police officer's searching intensity. The number $\gamma^*(c, r)$ is between zero and one, so at equilibrium the officer randomizes over whether to search each type c, r.

For a police officer to be willing to randomize, it must be that $P^*(G|c, r) = t_r$ for all c, r. At equilibrium, for all c

$$P^*(G|c, A) \quad = \quad t_A$$

[20]The implication that policemen and motorists randomize may look unrealistic, since it requires agents to be indifferent across actions. However, there is a simple interpretation of these mixed strategies that does not require agents to actually flip coins. This is discussed in Section 2.3.

$$
\begin{aligned}
P^*(G|c,W) &= t_W \\
\gamma^*(c,A) &= \frac{v(c,A)}{[v(c,A)+j(c,A)]} && (2) \\
\gamma^*(c,W) &= \frac{v(c,W)}{[v(c,W)+j(c,W)]} && (3)
\end{aligned}
$$

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

$$P^*(G|c,A) = t = P^*(G|c,W). \tag{4}$$

Notice that this does not imply $\gamma^*(c,W) = \gamma^*(c,A)$. In other words, given c, the equilibrium search intensity may be higher for African Americans even in the absence of prejudice. This happens if $\frac{v(c,W)}{[v(c,W)+j(c,W)]} < \frac{v(c,A)}{[v(c,A)+j(c,A)]}$, i.e. if the expected value of carrying drugs is higher or the cost of being convicted lower for black motorists, after conditioning on observables c. $\gamma^*(c,W) > \gamma^*(c,A)$ may indicate that race proxies for some variable that is unobservable by the police officer and is correlated with both race and crime. Possible examples of such an unobservables are educational attainment or the earnings potential of the motorist.

2.2 Testing for Prejudice

Equation (4) provides a test for prejudice ($t_W \neq t_A$) that is implementable even in the absence of data on c and on γ^*. It suffices to have data on the frequency of guilt by race, conditional on being searched,

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

Using (4) to substitute for $P(G|c,r)$ we get

$$D(W) = t = D(A), \tag{5}$$

which is the implication that we test in the data below.

In the model, there is nothing special about the characteristic "race." Thus, the analog of (5) should hold for any other observed characteristics. Equality of posterior frequencies of guilt should hold true across any characteristic on which the police officer conditions his searching decision. In section 3, we test condition (5) using race and gender variables.

The empirical evidence is consistent with equation (5), which we interpret as indicating absence of prejudice against African Americans on the part of the police. At the same time, our data indicate that African Americans are searched proportionately more often than whites, i.e. $\gamma(A) > \gamma(W)$. This disparity may be due to two reasons. It is possible that at equilibrium $\gamma(c, A) \neq \gamma(c, W)$, i.e. blacks are more likely to be searched than whites *with the same observable characteristics* c. If true, this may indicate that race proxies for some variable that is unobservable by the police officer and is correlated with both race and crime. The observed disparity may also arise from differences in the distribution of characteristics c among races (i.e. $F(c|W) \neq F(c|A)$. Distinguishing between these two alternatives would require comprehensive data on c.

2.3 Discussion of the Model

In our model, motorists respond to the probability of being searched. This assumption is key to obtaining a test for racial prejudice that can be applied without data on all the characteristics police use in the search decision. If the probability of being guilty was exogenous, i.e. independent of the probability of being searched, testing for prejudice would require data on the full set of characteristics c. To see this, denote by $P(G|c, r)$, the (now, exogenously given) probability that a motorist with characteristic c and race r is found guilty if searched. Supposing that $P(G|c, r)$ is increasing in c, police officers choose two cutoffs k_W and k_A, and they search any motorist of race r with a

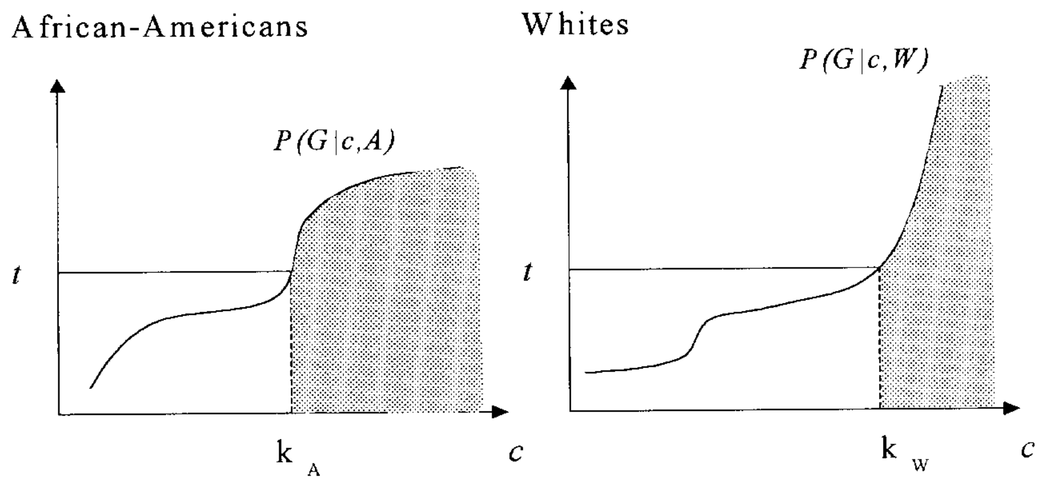


Figure 1: Model with exogenous guilt probabilities.

c greater than k_r. In the absence of prejudice, maximizing behavior on the part of police implies that the probability that types k_W and k_A are guilty should equal the marginal cost of searching motorists. See Figure 1.

Without data on c, we cannot determine who are the marginal motorists in the data. Furthermore, this model has no predictions on the average guilt probability of supramarginal types (the shaded areas in Figure 1), which is what we observe in the data. Thus, we cannot estimate the equilibrium implication of a model where motorists do not react to the probability of being searched.

Now, let us return to the model where motorists react to the probability of being searched. A number of stylized features of the model can be made more realistic without jeopardizing the test for racial prejudice. First, it may seem counterintuitive that at equilibrium motorists randomize over whether to carry drugs. Second, the model assumes that the characteristics c in the individuals' utility function are the same as those observed by the police officer (i.e. *no private information*). We can obtain a version of our model where motorists retain private information that is not observable by the police officer, and motorists never randomize, by "purifying" the mixed strategies. Purification is performed by adding a random variable X to the utility function of each motorist. The realization of X represents an individual motorist's idiosyncratic component of his propensity to carry drugs, which is information. Given a certain probability of being stopped, those motorists with a high realized value of X strictly prefer to carry drugs, and those with small values of X strictly prefer not to carry drugs. Thus, at equilibrium no motorist randomizes. Notice that, at equilibrium, if police are not racist, the returns to searching motorists of any type c, r are the same. If this were not the case, police would search for sure groups with the highest probability of being found with drugs, which cannot happen at equilibrium. This extension is outlined in Appendix B.

Another stylized feature of our model is that characteristics c and r are given, so that individuals do not choose their characteristics. In reality, some characteristics — such as tinted windows — can plausibly be viewed as endogenous. When characteristics can be purchased, some types with characteristics closely linked to criminal behavior will choose to purchase more innocuous characteristics. At equilibrium, the police will take this into account when computing the probability of being guilty of motorists in a certain group. Appendix B.2 extends the model to incorporate endogenous characteristics and shows that the same test still applies.

3 Empirical Results

3.1 Data Description

The data we analyze were collected as part of the settlement of a lawsuit filed in February 1993 by the ACLU, challenging as unconstitutional the Maryland State Police's alleged use of a "racial profile" as a basis for stopping, detaining and searching motorists. In the settlement, the state agreed to maintain detailed records on motorist searches and to file quarterly reports with the Court and the ACLU. In November, 1996 the Maryland ACLU asked a federal court in Baltimore to hold the Maryland State Police agency in contempt of court and to impose a $250,000 penalty, based upon evidence showing that state police violated a 1995 court decree by continuing a pattern of race discrimination in drug interdiction activities carried out along the Interstate 95 corridor. Because the data are only available for the time period after the initiation of the first ACLU lawsuit, the estimates reported below cannot be construed as describing police behavior prior to the legal interventions.

Our data consists of 1590 observations on all car searches on a stretch of I-95 in Maryland from January, 1995 through January, 1999.[21] The data provide information on the race and sex of the motorist as well as on the year, make and model of the vehicle, and the date, time and location of the search. We also know whether the police officer requested consent to search the car and, if consent to search was not requested, the probable cause that the police officer invoked to search the car. The probable-cause information provides insight into what types of characteristics are considered grounds for initiating searches.[22] We also know whether dogs were used in the search and whether contraband (typically illegal drugs) was found. If any was recovered, we know what type and how much (in grams). Finally, the data includes the name of the police officer performing the search.

It is important to note that our data pertains only to motorists who were both stopped and searched; it does not include motorists who were

[21]The searches were conducted in Baltimore, Cecil, Harford, Howard, and Prince George's Counties.

[22]For example, 'third-party-vehicle' (a vehicle not owned by the driver) is often listed as a grounds for requesting consent to search. Appendix A lists other observables that police sometimes use.

stopped but not searched. While data on stops might reflect differences in driving habits between different race and sex groups, search data only includes motorists who police officers suspect to be carrying contraband. The total population searched is also the relevant one for the denominator of the success rate $P(G|c,r)$.

Another important observation is that although the decision to search a car is usually contingent on the driver's having committed some traffic violation, in practice this does not pose a constraint for police who seem to be able to stop and search, if they want to, almost any motorist on the highway. This is because the vast majority of drivers along I-95 commit traffic violations (mostly speeding).[23] In fact, part of the racial profiling controversy is that state troopers often use minor traffic violations, such as exceeding the speed limit by 5 miles, as pretexts to stop and search the motorist, other occupants and the vehicle.[24]

Table 1 summarizes the means and variances of the subset of variables used in our empirical tests. Of the 1590 total searches, 1007 or 63.4% were performed on African-Americans, 466 or 29.3% on whites, 97 or 6.1% on Hispanics, and the remaining (1.3%) on other race/ethnic groups. Female motorists were rarely searched: a total of 117 female motorists appear in the data, compared with 1473 men. Marijuana is the drug most commonly found (23% of the times), and it is not uncommon for drivers to be carrying up to three different types of drugs, as well as drug paraphernalia. About a third

[23]According to a study designed by John Lamberth, 98.1% of all cars on a stretch of the New Jersey Turnpike were clearly exceeding speed limits. See *New Jersey v. Soto*, 1996.

[24]In 1986 for instance, the Drug Enforcement Agency trained 27,000 police officers in 48 states in the use of pretext stops to find drugs in vehicles. According to the ACLU, the training materials in these and similar programs "implicitly" encourage the targeting of minority motorists. The practice of using discretionary stops as pretexts in this way was supported by the Supreme Court in *Whren and Brown vs. US* in 1996, which held that any traffic violation was a legal basis for stopping a motorist.

of the searches occur during the hours of midnight-6am.

Figure 2 plots the proportion of drivers searched who were African-American against time. The circle size is proportional to the inverse of the standard deviation of the estimates, with the larger circles having larger sample sizes. The figure reveals a downward trend over time in the proportion of African-American drivers searched and an upward trend in the proportion of white drivers searched.[25] There is no clear trend for female motorists. If police practices changed over time, the issue arises as to whether the model's predictions hold in subperiods of the data. This is discussed below.

To implement our test for detecting tastes for discrimination, we need to define what it means to be guilty. We classify as guilty anyone found carrying any amount of drugs of the following types: marijuana, heroine, cocaine, crack, PCP, LSD, and methadone.[26] Thus, guilty corresponds to the police having reported finding drugs and not to any court decision. The proportions of African-American and white drivers found to be carrying drugs are plotted in Figure 3.

3.2 Test Results

Our test for prejudice compares the probability of being found guilty conditional on a subset of observed characteristics. The model has a strong implication; namely, that no matter what the set of characteristics, the probability of being guilty should be the same across the groups. That is, the test compares

$$\Pr(G = 1|r, c) = \Pr(G = 1) \text{ for all } r, c$$

[25]These trends are statistically significant in regressions of the proportions on a linear time trend. These results are reported in Appendix B.

[26]A small number of individuals were found in possession of barbituates (such as valium); these we did not classify as guilty.

where c is a set of characteristics and r is the race indicator variable. The probability can be estimated using a parametric procedure such as probit or logistic regression.

If the regressors are discrete, then our test for equality of guilt probabilities across groups corresponds to a test of whether the coefficients associated with all the conditioning variables except the intercept are jointly equal to zero. This is a more stringent requirement than the conventional test of whether the coefficient on race is statistically different from zero.(which is usually applied to the probability of being searched)

A drawback of using a parametric approach to testing is that the test is generally only valid if the systematic component and distribution of the error component of the model are correctly specified. An alternative, simpler approach to testing equality of guilt probabilities across groups is to condition on characteristics nonparametrically through cell means, which is not subject to the criticism that the probability model could be misspecified.[27] Here we report results from Pearson chi-squared tests of association within classification tables. In the appendix, we compare the nonparametric results to those from a probit model, which leads to identical conclusions.

Tables 2a-2c compare the probability of being found guilty of carrying drugs, by race and sex. Although African-American motorists are more likely to be searched by police, the proportion found guilty among whites and African-Americans is nearly identical (0.32 vs. 0.34). This finding is consistent with the hypothesis of no racial prejudice. For Hispanics, however, the proportion guilty is 0.11, which is significantly lower than for African Americans or whites. According to our model, this finding implies racial

[27] A problem that is often encountered using nonparametric estimators is that the cell sizes can become small as the number of conditioning variables increases. Fortunately, our test for racial prejudice requires only that a subset of the covariates be used, making a nonparametric test feasible.

prejudice against Hispanics.[28]

A comparison of the guilty proportions by sex shows that they are also similar (0.32 for men and 0.36 for women). A breakdown by both race and sex shows similar patterns, although guilty proportions are somewhat lower for white women than for African-American women.(0.22 vs. 0.44)

Table 3 reports p-values from Pearson chi-squared tests of no association between guilt and race. The test rejects the null at conventional significance levels when all race groups are used in the test, but does not reject the null when the sample is restricted to African-Americans and whites. The test also does not reject the null hypothesis of no association between guilt and the sex of the motorist (p-value=0.32). When we condition on both race and sex, the joint test rejects the null hypothesis when Hispanic males are included, but not when they are excluded.[29]

Our data were collected during three distinct time periods, the first when the police were being audited as part of a law suit, the second when the audit period was over but the police were still gathering data as part of the settlement agreement, and the third after the filing of a second lawsuit alleging that the police were still discriminating. These distinct periods are apparent in Figure 4, which shows how many cars were searched in different time periods. During the audit period and after the filing of the second lawsuit, more searches were conducted.

Tests combining all time periods might mask variation over time in police search behavior. To address this concern, we perform identical tests after disaggregating the data by the three time periods: prior to May, 1996, in between June, 1996 and December, 1998, and after December 1998. The

[28]However, a caveat is in order: when Hispanics are found with drugs, they appear to be found with very large quantities relative to other races. Unfortunately, because we have so few guilty Hispanics in our data, this feature is not statistically significant.

[29]There are no Hispanic women in our data, so the rejection is due to Hispanic males.

tests performed on disaggregated data generally yield the same conclusions as those performed on the full sample, not rejecting equality of the probabilities within each of time periods when Hispanic males are excluded from the sample. These results are reported in Appendix C.

4 Efficiency vs. Fairness

In the context of our model, it might be thought that there is a tradeoff between efficiency and racial fairness; if police are forced to stop both races at the same rate, efficiency would seem to be compromised. However, we now demonstrate that this is not necessarily the case. Constraining the police to use a color-blind search strategy may result in the same level of deterrence at a lower cost. This is somewhat surprising because one might think that police, by trading off the number of arrests with the cost of searching, are effectively maximizing efficiency. In the following, we show that because this is not the case, there might be benefits from forcing the police to forego racial profiling.

First, we need to define what we mean by efficiency in drug interdiction. This definition is not meant to capture Pareto-efficiency, as we do not take into account the utility of motorists.

Definition 3 *Given two outcomes z and z', we say that z' is more efficient than z if both the number of motorists carrying drugs and the total cost of searching are weakly lower in z' than in z, and at least one of the two is strictly lower in z'.*

Efficient outcomes may not be fair. Assume that a person who does not carry contraband suffers a cost when searched. If efficiency requires that persons from one race be searched more frequently, they will suffer a higher expected cost of being searched. We say that an outcome is racially fair when individuals of different races with the same observables c, experience the same expected cost of being searched.

Definition 4 *An outcome is racially fair if for all c we have $\gamma(c,W) = \gamma(c,A)$.*

One way of implementing the fair outcome is to constrain the police to search both races at the same rate, conditional on c.

To focus on the trade-off between fairness and efficiency, it is useful to consider a more restricted version of our model. Assume that $t_A = t_W = t$ and c is absent, so the only observable on which the police can condition is race. Assume further that the value of carrying drugs is higher for African-Americans than for whites, i.e. $\frac{v(W)}{[v(W)+j(W)]} < \frac{v(A)}{[v(A)+j(A)]}$. We refer to these two groups as A and W.

Consider now the behavior of police under the "fairness" constraint that $\gamma(A) = \gamma(W) = \gamma$. We show that the tradeoff between fairness and efficiency depends crucially on the fraction of A's in the population relative to the cost of search. Let $\varphi(r)$ represents the fraction of people with race r in the entire population.

Proposition 1 *If $t < \varphi(A)$, imposing fairness generates a less efficient outcome. If $t > \varphi(A)$, imposing fairness generates a more efficient outcome.*

Proof. When police are unconstrained, persons in groups A and W will be searched at rates $\gamma^*(A)$ and $\gamma^*(W)$ respectively, with $\gamma^*(A) > \gamma^*(W)$. At equilibrium, the benefit of law enforcement is t, the probability that any motorist carries contraband. The cost is

$$t\left[\gamma^*(A)\varphi(A) + \gamma^*(W)\varphi(W)\right].$$

Denote with the superscript $**$ the equilibrium of the constrained game where police are forced to adopt $\gamma(A) = \gamma(W) = \gamma$.

Case 1: $t < \varphi(A)$

We now construct an equilibrium where members of group W do not carry drugs, while members of group A randomize.

For A's to be willing to randomize it must be that $\gamma^{**} = \gamma^*(A)$. At this level of γ, W's strictly prefer not to carry drugs, consistent with the proposed equilibrium.

For the police to be willing to randomize it must be that $t = P^{**}(G|A)\varphi(A)$, which implicitly defines $P^{**}(G|A)$. Notice that $P^{**}(G|A)$ is well-defined because by assumption $\varphi(A) > t$.

The benefit of interdiction at the constrained equilibrium is equal to the average probability that a motorist carries contraband at equilibrium, $P^{**}(G|A)\varphi(A) = t$. The cost at the constrained equilibrium is, after substituting $\gamma^*(A)$ for γ^{**},

$$t\left[\gamma^*(A)\varphi(A) + \gamma^*(A)\varphi(W)\right].$$

Comparing the constrained equilibrium with the unconstrained one, we notice that the benefit is the same, but the number of searches necessary to achieve that benefit differs. Indeed, the search cost in the unconstrained equilibrium is $t\left[\gamma^*(A)\varphi(A) + \gamma^*(W)\varphi(W)\right]$, so the efficiency cost of fairness (the difference in search costs) is

$$t\varphi(W)\left[\gamma^*(A) - \gamma^*(W)\right]. \tag{6}$$

This quantity is positive because, by assumption, $\gamma^*(A) > \gamma^*(W)$.

Case 2: $t > \varphi(A)$

In this case, at the constrained equilibrium all A's carry drugs with probability one and W's carry drugs with probability $P^{**}(G|W)$. This probability solves

$$P^{**}(G|W)\varphi(W) + \varphi(A) = t.$$

The left-hand side describes the equilibrium probability that a random driver is guilty. The equality guarantees that a police officer is indifferent between searching and not. $P^{**}(G|W)$ is well-defined since $t > \varphi(A)$.

Police search all motorists at rate $\gamma^{**} = \gamma^{*}(W)$, which makes whites indifferent between carrying drugs or not, and A's strictly prefer to carry drugs.

The equilibrium probability that a random individual in the population carries drugs is t, so the benefits of interdiction are the same as in the unconstrained equilibrium. However, in the constrained equilibrium all motorists are searched at the same rate $\gamma^{*}(W)$, while in the unconstrained equilibrium, some motorists (those in the A group) are stopped at rate $\gamma^{*}(A) > \gamma^{*}(W)$. Thus, the cost of interdiction is higher at the unconstrained equilibrium. The efficiency gain from imposing fairness is

$$t\varphi(A)\left[\gamma^{*}(A) - \gamma^{*}(W)\right] \tag{7}$$

∎

Let us discuss the equilibrium in the case $t > \varphi(A)$. Imposing the race-blind constraint effectively pools together the two races, so that motorists of both races experience the same probability of being searched. At equilibrium, this causes blacks to increase their probability of carrying drugs, and whites to decrease their amount of drug-carrying so as to exactly offset the increased likelihood that the pooled group carries drugs. Thus, although blacks carry drugs with probability one, the average amount of drug-carrying in the population is unchanged from the unconstrained equilibrium. At the constrained equilibrium, whites are indifferent between carrying drugs or not, and the level of interdiction required to make whites indifferent is lower than the level of interdiction in the unconstrained equilibrium. Imposing the race-blind constraint achieves the same level of drug-carrying in the population, but the cost of interdiction decreases relative to the unconstrained equilibrium.

We have shown that imposing a constraint on police behavior may achieve a more efficient outcome. This results from two features of the model. First, the fact that police target the number of arrests. In a world with more than one type of motorist, this is not equivalent to the social goal of targeting the aggregate rate of drug carrying in the population. Second, the fact that the drug-carrying decision is endogenous; pooling together two types of motorists may, at equilibrium, cause some types to decrease their probability of carrying drugs, to accommodate the behavior of other types of motorists who have a higher propensity of carrying drugs.

One must keep in mind that the above efficiency analysis equates the benefits of drug interdiction to the reduction of contraband on the road. In practice, the benefits and costs of drug interdiction are more complex. Subject to this caveat, however, one could use expressions (6) and (7) to measure the costs (or benefits) of imposing fairness. One needs the probabilities that police search motorists of different races at the unconstrained optimum, the racial proportions of the population, and the costs of search t.

5 Summary and Conclusions

Given the key role of statistical testing in detecting discrimination, it is important to know what assumptions on the behavior of motorists and troopers are needed to justify different types of tests. In this paper, we developed a simple equilibrium model of law enforcement via traffic searches and considered its implications for testing for racial prejudice in policing. Existing tests for discrimination typically regress an indicator for whether a motorist is searched on a number of characteristics and check whether race has any additional explanatory power. We discussed two disadvantages of these sorts of tests. First, their validity relies crucially on which set of variables are considered admissible nondiscriminatory variables that police can use in searching cars and on whether those variables are available in the data. Second, they are only informative on whether a disparity by race exists, and not about the motivation for the disparity. The question of motivation plays a prominent role in racial profiling court cases.

Our equilibrium model of police and motorist behavior provides a test for whether racial disparities in motor vehicle searches reflect prejudice, or instead are consistent with a non-prejudiced police maximizing drug interdiction. The test is based on the success rates of police searches, and compares the probabilities that various subgroups of the population are found guilty of carrying contraband when searched. Our test is valid even when the set of characteristics observed by the police is only partially observable by the econometrician, and so it is less demanding than the standard test in terms of data requirements. This important feature results from explicitly modeling the reaction of motorists to the probability of being searched.

In our model, at equilibrium both races should have the same probability of carrying drugs, yet one race may be searched more often than another. Actually, the fact that some subgroups are searched more often than another may be *necessary* to sustain equality in the proportions guilty across subgroups. Thus, it is ironic that differences in search intensities by race, coupled with equality in the proportions guilty by race, has been used in court to argue that police are racist in searching a greater fraction of cars of African American motorists.[30]

Our empirical results for the Maryland data showed that the probabilities of being found with drugs are equal across African-Americans and whites,

[30]¿From a memorandum prepared by ACLU lawyers: "MSP's own data demonstrates that this racial distortion is unnecessary to successful drug interdiction. ...[Indeed,] MSP data shows that statewide, police find contraband on black and white motorists at equal rates." Mertens and Jeon (1996).

which is consistent with maximizing behavior by troopers who are not racially prejudiced against African-Americans. However, our findings for Hispanic males are consistent with a taste for discrimination against this group.

Statistical discrimination, even if not due to racial prejudice, may be considered unfair, because innocent drivers experience different probabilities of being searched depending on their race. In the context of our model, it might be thought that there is a tradeoff between efficiency of drug interdiction and racial fairness. However, our investigation of the tradeoff showed that achieving the fair outcome need not entail a cost in terms of efficiency. In some cases, requiring that police implement a color-blind search strategy can increase efficiency of drug interdiction. This observation suggests that there may exist an efficiency rationale for imposing racially fair behavior on the police.

To the best of our knowledge, this paper presents the first empirical test of an optimal auditing model. Although this paper focuses on traffic searches, our analysis extends straightforwardly to some other similar settings; for example, our test could be applied to analyze the behavior of security and customs agents in airports, where it is alleged that minorities and foreigners are unfairly targeted in baggage and passenger searches.[31]

[31]See for instance *Anderson v. Cornejo*, 1999 (No. 97 C 7556).

References

[1] Arrow, K. (1973) "The Theory of Discimination." In *Discrimination in Labor Markets*, O. Ashenfelter and A. Rees, eds., Princeton University Press, Princeton NJ.

[2] Becker, G. (1957) *The Economics of Discrimination*, Chicago University Press.

[3] Becker, G. (1968) "Crime and Punishment: An Economic Approach" *Journal of Political Economy*, 76, p. 169-217.

[4] Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel and Timothy H Hannan (1998) "Discrimination, Competition, and Loan Performance in FHA Mortgage Lending, " *The Review of Economics and Statistics*, p. 241-250.

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

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

[7] Donohue, John and Levitt, Steven (1998): "The Impact of Race on Policing, Arrest Patterns, and Crime" NBER Working Paper # W6784.

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

[9] Kennedy, Randall (1997): *Race, Crime and the Law*, New York: Pantheon Books, p. 136-167.

[10] Lamberth, John (1996) Report in connection with the case *Wilkins v. Maryland State Police*, Civil Action No. CCB-93-483.

[11] Mertens William J. and Deborah A. Jeon (1996) "Memorandum in support of plaintiff's motion for enforcement of settlement agreement and for further relief in the United States District Court for the District of Maryland," *Wilkins at al. v. Maryland State Police.* Civil Action No. CCB-93-468

[12] Reinganum, J. and L. Wilde (1986) "Equilibrium Verification and Reporting Policies in a Model of Tax Compliance" *International Economic Review*, 27, p. 739-760.

[13] Scotchmer, Suzanne (1987) "Audit Classes and Tax Enforcement Policy" *American Economic Review.* Vol. 77 (2). p 229-33.

[14] Stigler, G (1970) "The optimum enforcement of laws" *Journal of Political Economy*, 78, p. 526-536.

A Observable Indicators of Criminal Activity

The following indicators of criminal activity are listed in the police officer's manual for the Valkyrie police (p. 13-15):

Citizen's band radios, Cellular telephones, Pre-paid phone cards, Tinted Windows, Radar Detectors, Perfumes, Duct tape, Pagers, Screws, Handles and knobs, Inability to completely roll down windows (which may hide concealment of drugs in the doors), Religious paraphernalia used to divert suspicion, Police materials used in attempt to show support for law enforcement, High odometer mileage, particularly on late model vehicles, Switches and buttons, which may activate electronic compartment doors, Large amounts of cash, Attorney's business cards, Too little or too much luggage for stated length of trip, Signs of recent drug use, Hiding places, Weapons, Only one key in ignition or no trunk key, Maps from source cities or states, map turned to locations other than those mentioned by occupants, Leased vehicles: 'Leased vehicles are used frequently by drug traffickers. Many times these vehicles are rented from airports. The person(s) authorized to drive the vehicle should be noted on the lease agreement'; Third-party vehicles: 'Question when the owner of the vehicle is not at the scene of the stop or if the occupants cannot tell you the owner's/lessee's name without looking at the registration or lease agreement themselves; Cashier's checks, Bondo (a compound applied to the car's exterior)

Source: Expert witness testimony by Professor John Donohue in *ACLU v. State of Illinois.*

B Extensions

B.1 Mixed Strategies and Private Information

In the model, motorists have no private information about any feature of their own utility function (of course, they do have private information on whether they are carrying drugs). We now show how the model can be modified to allow for private information, yielding an equilibrium where motorists do not randomize over the decision to carry drugs. The procedure, called "purification" of mixed strategies, is standard. The modified model preserves the main testable implication that at equilibrium the guilt probability is constant across groups. A simple example suffices to illustrate the idea.

Consider an agent who randomizes between actions a and b with $\Pr(a) = p$. Then utility maximization implies that the agent is indifferent between the two actions, $v(a) = v(b)$. To purify this mixed strategy, we can imagine that the utility from action a is really $v(a) + X$, where X is a random variable with the property that $\Pr(X > 0) = p$.[32] If this is the case, the agent chooses a with probability p, and is never indifferent between actions a and b. This perturbed model is equivalent in terms of outcomes to the mixed strategy, but the agent is never indifferent between the two actions and, as a consequence, never randomizes. The random variable X can be thought of as the agent's private information on his own propensity to take action a.

In the context of our model, we can imagine a class of individuals, all with characteristics c, r. Let i denote a generic individual in that class, and let i's propensity to carry contraband be $v(c,r) + X^i_{c,r}$ where $X^i_{c,r}$ is a random variable independent across all motorists in that class. We can imagine that motorist i in class c, r knows his realized value $x^i_{c,r}$. Then, given a cer-

[32] X may have small support, so that the two models may be quite close in terms of primitives.

tain $\gamma(c,r)$, motorist i carries contraband if and only if $\gamma(c,r)[-j(c,r)] + [1-\gamma(c,r)]\left[v(c,r)+x^i_{c,r}\right] > 0$. Equating the left hand side of this inequality to zero, we find a threshold value $\overline{x}_{c,r}$ such that a motorist carries contraband if and only if $x_{c,r} > \overline{x}_{c,r}$. This threshold determines the fraction of individuals in class c, r who carry contraband. Suppose that this fraction is higher in class c, r than in class c', r'. Then police officers will never search an individual in class c', r', which implies that $\gamma(c', r') = 0$. But this decreases the threshold in class c', r' until the equilibrating process equalizes the probabilities of carrying contraband across classes.

B.2 Endogenous Characteristics

Our model assumed that observables c and r are inherent characteristics of the motorist. We now relax this assumption and show that the model can be extended to the case where motorists can incur a small cost to change their characteristics. We show that, if police are non racist, motorists of both races have a probability t of being found with drugs, regardless of their observable characteristics.

To demonstrate this, it suffices to show that, if there is only one race, all motorists who are stopped are found guilty with probability t. This is because the maximization problem of police can be rewritten as two separate maximizations, one for African-Americans and one for Whites.

So, consider a world where there is only one race, and where there in only one observable characteristics that can take on two values: c and c' (e.g. license plates for two different states).[33] Assume that $v(c') > v(c)$ and $j(c') = j(c) = 1$. Thus, motorists with plate c' have a higher criminal

[33]Police often target motorists with license plates from certain states considered to be drug source states (e.g. Florida, Texas) or drug destination states (e.g. New York, Illinois, Pennsylvania).

propensity than motorists with license plate c. If, as in the above model, license plates are a fixed characteristic of each individual, then at equilibrium $\gamma^*(c') > \gamma^*(c)$. The utility of type c' is

$$v(c')(1-\gamma^*(c')) - \gamma^*(c') = 0.$$

Suppose now that it is possible for a type c' to purchase a license plate c at a cost of e. If e is sufficiently small, $\gamma^*(c')$ and $\gamma^*(c)$ cannot be part of an equilibrium, since type c' would prefer to carry drugs and incur the cost e to mimic type c, thus lowering his probability of being searched. Denote by $\varphi(c)$ and $\varphi(c')$ the fraction of motorists in the total population with characteristic c and c'. Below we describe the equilibrium of the model where characteristics can be purchased for small e. The equilibrium variables are denoted by superscript P. It is important to distinguish between the *type*, by which we mean the original characteristic of a motorist, and the *observable characteristic*, which refers to what the police officer observes after types c' have decided how to represent themselves. Recall that t is the cost of searching a car.

Case A: $\varphi(c') > t$.

There is an equilibrium where:

- Types c do not carry drugs, $P^P(G|c) = 0$
- Types c' randomize with probability τ over whether to incur the cost e. Those who incur e then carry drugs with probability one, while those who do not incur e carry drugs with probability $P^P(G|c', \text{Not } e) = t$. This makes police officers willing to randomize over searching motorists with observable characteristic c'. Types c' choose probability τ so that

  $$\frac{\tau\varphi(c')}{\tau\varphi(c') + \varphi(c)} = t.$$

  This expression on the left is the fraction of motorists with observable characteristic c who are guilty. Equating this to t makes police officers willing to randomize between searching motorist with characteristic c or not. Because $\varphi(c') > t$ by assumption, τ is well-defined.

- Motorists with observable characteristic c' are searched at rate $\gamma^P(c') = \gamma^*(c')$. This makes those motorists willing to randomize over whether to carry drugs.

- Motorists with observable characteristic c are searched at a rate $\gamma^P(c)$ defined by

$$v(c')\left(1-\gamma^P(c)\right) - \gamma^P(c) - e = 0.$$

  This choice of $\gamma^P(c)$ makes types c' indifferent between incurring cost e or not (after type c' incurs e he strictly prefers to carry drugs, since e is sunk). Also, since $v(c) < v(c')$, this choice of $\gamma^P(c)$ makes types c strictly prefer to not carry drugs for e sufficiently small.

Case B: $\varphi(c') < t$.

In this case there is an equilibrium where:

- Types c' incur the cost e and carry drugs for sure.

- Types c randomize between carrying drugs or not with probability $P^P(G|c)$ solving

$$\varphi(c') + P^P(G|c)\,\varphi(c) = t.$$

  This makes police officers willing to randomize between searching motorist with observable characteristic c or not. Because $\varphi(c') < t$ by assumption, $P^P(G|c)$ is well-defined.

- Motorists with observable characteristic c are searched at a rate $\gamma^P(c)$ defined by

$$v(c)\left(1-\gamma^P(c)\right) - \gamma^P(c) = 0.$$

This choice of $\gamma^P(c)$ makes types c indifferent between carrying drugs or not. Also, since $v(c) < v(c')$, this choice of $\gamma^P(c)$ makes types c' strictly prefer to carry drugs for e sufficiently small.

- Motorists with observable characteristic c' (of which there are none at equilibrium) are searched at rate $\gamma^P(c') \geq \gamma^*(c')$. This makes motorists of type c' strictly prefer to incur the cost e, for e sufficiently small.

In both cases, if a motorist is stopped, the probability that he is carrying drugs is t, regardless of his observable characteristic. That is what we wanted to show.

C Supplementary Test Results

In this appendix, we report test results based on data that is disaggregated into three time periods, roughly corresponding to the three different periods of data collection as described in the text. We also compare nonparametric test results with those from a parametric model for the probability of guilt.

Table B.1a compares the probability of being found with drugs by race for the three different time periods. Table B.1b and B.1c present analogous results when the breakdown is by sex and by sex and race. In each time period, the probabilities are close for whites and African Americans and for men and women. The probabilities are consistently lower for the Hispanic group. Table B.2 reports p-values from Pearson chi-squared tests performed within each of the time periods. As with the combined sample, we always reject the null that probabilities are equal across race groups when Hispanics are included in the test but do not reject when the test is performed only on whites and African Americans. When we break the groups down by race and sex across different time periods, we marginally reject the null for the third time period (p-value of 0.04). For this period, the probability of being guilty is lower for white females and for African American males, which would be consistent in our model with prejudice operating against white females and African American males.

Table B.3 reports test results analogous to those already discussed, except where the conditional probabilities of guilt are estimated by a logit model.[34] The table footnotes report p-values from tests for equality of coefficients under different specifications. In the first column, the coefficient on race is constrained to be the same across time. The inference is the same as for Table 3 in the text. In the second specification, we allow for a race-specific time

[34]Because the race category "other" has small sample sizes, we exclude people in this category in estimation.

trend in the probability guilty. The coefficient on time is never significant for any of the race groups. In the last column, we allow the effect of race to be different for the three time periods. Again, the inference is the same as for Table B.3.

Table 1
Means and Standard Deviations of Variables used in Analysis
(standard deviations in parentheses)

	All Obser-vations	**By Race**				**By Sex**	
		Black	**Hisp.**	**White**	**Other**	**Female**	**Male**
Black	0.63	1.00	0.00	0.00	0.00	0.64	0.63
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)	(0.01)
White	0.29	0.00	0.00	1.00	0.00	0.35	0.29
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)	(0.02)
Hispanic	0.06	0.00	1.00	0.00	0.00	0.00	0.07
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Female	0.07	0.07	0.00	0.09	0.22	1.00	0.00
	(0.01)	(0.008)	(0.00)	(0.01)	(0.09)	(0.00)	(0.00)
Guilty	0.33	0.35	0.12	0.32	0.47	0.38	0.32
	(0.01)	(0.02)	(0.03)	(0.02)	(0.11)	(0.05)	(0.01)
Cocaine	0.08	0.10	0.03	0.03	0.37	0.09	0.08
	(0.01)	(0.01)	(0.02)	(0.01)	(0.11)	(0.03)	(0.007)
Marijuana	0.23	0.23	0.23	0.26	0.41	0.21	0.23
	(0.01)	(0.01)	(0.01)	(0.02)	(0.11)	(0.04)	(0.01)
Crack Cocaine	0.04	0.05	0.01	0.01	0.22	0.06	0.04
	(0.005)	(0.01)	(0.01)	(0.004)	(0.09)	(0.02)	(0.005)
Heroine	0.02	0.02	0.03	0.03	0.22	0.06	0.02
	(0.003)	(0.004)	(0.02)	(0.01)	(0.09)	(0.02)	(0.004)
Morphine	0.001	0.00	0.00	0.002	0.00	0.00	0.001
	(.001)	(0.00)	(0.00)	(0.002)	(0.00)	(0.00)	(0.001)
Other Drugs	0.01	0.00	0.00	0.01	0.00	0.01	0.02
	(0.002)	(0.00)	(0.00)	(0.005)	(0.00)	(0.01)	(0.003)
Paraphernalia	0.01	0.003	0.010	0.02	0.00	0.01	0.01
	(0.002)	(0.002)	(0.010)	(0.006)	(0.00)	(0.01)	(0.002)
Night	0.43	0.46	0.44	0.35	0.51	0.47	0.43
(12am-6am)	(0.01)	(0.02)	(0.05)	(0.02)	(0.11)	(0.05)	(0.01)
Number of Observations	1582	1002	97	463	20	117	1465

Table 2a
Proportion of Vehicles Searched Found with Drugs
by Race/Ethnicity

	Not Guilty	Guilty
African American	0.66	0.34
White	0.68	0.32
Hispanic	0.87	0.11

Table 2b
Proportion of Vehicles Searched Found with Drugs
by Sex

	Not Guilty	Guilty
male	0.68	0.32
female	0.64	0.36

Table 2c
Proportion of Vehicles Searched Found with Drugs
by Race/Ethnicity and Sex

		Not Guilty	Guilty
male	**African American**	0.66	0.34
	White	0.67	0.33
	Hispanic	0.89	0.11
	Other	0.68	0.32
female	**African American**	0.56	0.44
	White	0.78	0.22
	Hispanic	*	*
	Other	100.00	*

Table 3
P-values on Pearson Chi-Squared Tests on
Hypothesis that Proportion Guilty is Equal Across Various Groups

Groups	χ^2	p-value
race (African American, Hispanic and white)	21.59	<0.001
race (African American, White)	0.97	0.33
sex (male, female)	0.82	0.37
sex and race (African American, Hispanic, white and male, female)	26.97	<0.001
sex and race (African American, white and male or female)	6.29	0.10

Table B.1a
Proportion of Vehicles Searched Found with Drugs
by Race/Ethnicity

	Period 1		Period 2		Period 3	
	Not Guilty	Guilty	Not Guilty	Guilty	Not Guilty	Guilty
African American	0.66	0.34	0.62	0.38	0.69	0.31
White	0.72	0.28	0.71	0.29	0.64	0.36
Hispanic	0.88	0.12	0.87	0.13	0.91	9.00

Table B.1b
Proportion of Vehicles Searched Found with Drugs
by Sex

	Period 1		Period 2		Period 3	
	Not Guilty	Guilty	Not Guilty	Guilty	Not Guilty	Guilty
male	0.68	0.32	0.67	0.33	0.69	0.31
female	0.63	0.37	0.64	0.36	0.65	0.35

Table B.1c
Proportion of Vehicles Searched Found with Drugs
by Race/Ethnicity and Sex

		Period 1		Period 2		Period 3	
		Not Guilty	Guilty	Not Guilty	Guilty	Not Guilty	Guilty
male	**African American**	0.66	0.34	0.63	0.37	0.71	0.29
	White	0.73	0.27	0.69	0.31	0.62	0.38
	Hispanic	0.88	0.12	0.87	0.13	0.91	0.09
	Other	0.70	0.30	0.50	0.50	0.80	0.20
female	**African American**	0.64	0.36	0.52	0.47	0.48	0.52
	White	0.56	0.44	0.00	100.00	0.79	0.21
	Hispanic	n/a	n/a	n/a	n/a	n/a	n/a
	Other	100.00	0.00	n/a	n/a	n/a	n/a

Table B.2
P-values on Pearson Chi-Squared Tests on
Hypothesis that Proportion Guilty is Equal Across Various Groups

Groups	Period 1 χ^2	Period 1 p-value	Period 2 χ^2	Period 2 p-value	Period 3 χ^2	Period 3 p-value
race (African American, Hispanic and white)	8.68	0.01	8.90	0.01	9.72	0.01
race (African American, White)	2.41	0.12	2.54	0.11	1.18	0.28
sex (male, female)	0.7225	0.40	0.11	0.74	0.20	0.66
sex and race (African American, Hispanic, white and male, female)	9.9481	0.04	13.99	0.007	17.00	0.002
sex and race (African American, white and male or female)	3.66	0.30	7.47	0.06	8.27	0.04

Table B3
Parameter Estimates for Probit Model of Conditional Probability of being 'Guilty'
Models without Covariates
(p-values from Hypothesis Tests shown in footnote)

Variable	Model (1)[(a)]	Model (2) [(b)]	Model (3) [(c)]
Indicator for white	-0.46	-0.66	…
	(0.06)	(0.13)	
Indicator for black	(-0.38	-0.32	…
	(0.04)	(0.07)	
Indicator for Hispanic	-1.16	-1.20	…
	(0.16)	(0.32)	
Indicator for white * time	…	0.007	…
		(0.004)	
Indicator for black * time	…	-0.003	…
		(0.004)	
Indicator for Hispanic * time	…	0.002	…
		(0.011)	
indicator for white * period 1	…	…	-0.58
			(0.11)
indicator for white * period 2	…	…	-0.53
			(0.13)
indicator for white * period 3	…	…	-0.34
			(0.09)
indicator for black * period 1	…	…	-0.39
			(0.05)
indicator for black * period 2	…	…	-0.27
			(0.09)
indicator for black * period 3	…	…	-0.45
			(0.09)
indicator for Hispanic * period 1	…	…	-1.17
			(0.28)
indicator for Hispanic * period 2	…	…	-1.13
			(0.29)
indicator for Hispanic * period 3	…	…	-1.17
			(0.28)

(a) P-value from test of hypothesis white = black = Hispanic is 0.0001. P-value from test that white = black is 0.2523.
(b) P-value from test of hypothesis black = white = Hispanic for both intercept and time trend is 0.0001. P-value from test that black = white for both intercept and time trend is 0.0530.
(c) P-value from test of hypothesis that black = white = Hispanic within all time periods is 0.0007. P-value from test that black = white for all time periods is 0.2266.

Figure 2

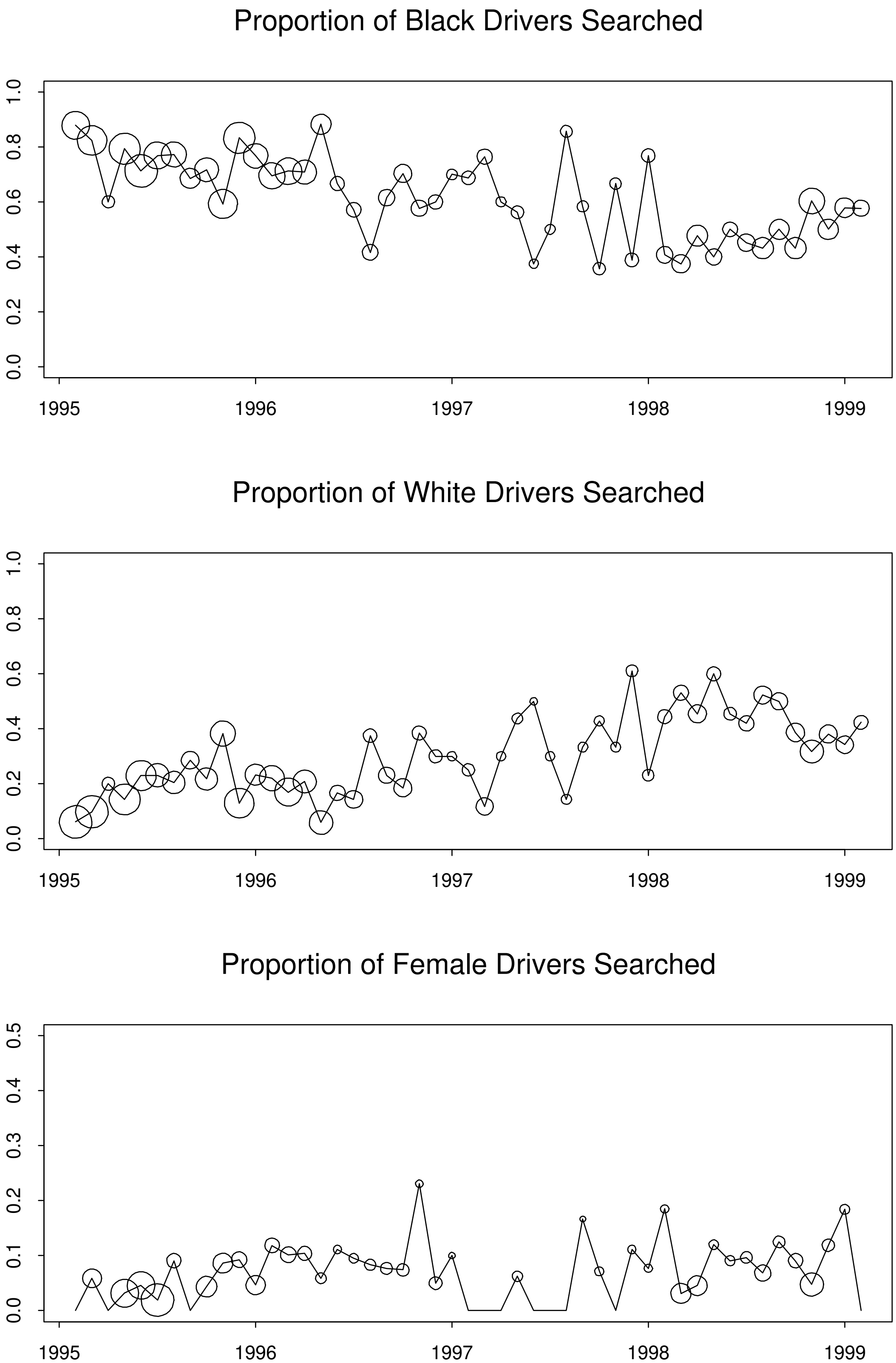
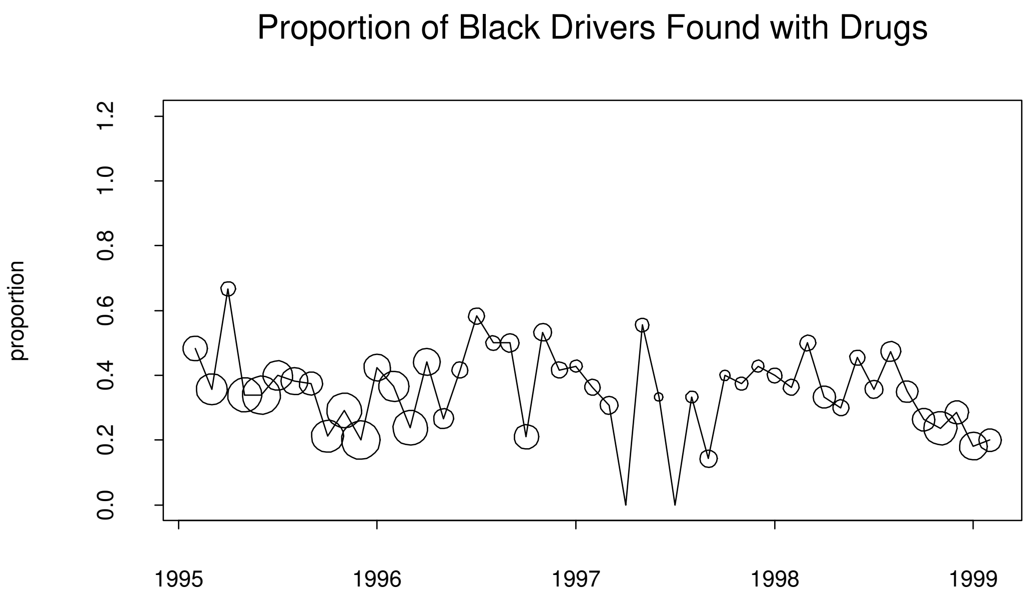

Proportion of Black Drivers Searched
0.0
0.2
0.4
0.6
0.8
1.0
1995
1996
1997
1998
1999
Proportion of White Drivers Searched
0.0
0.2
0.4
0.6
0.8
1.0
1995
1996
1997
1998
1999
Proportion of Female Drivers Searched
0.0
0.1
0.2
0.3
0.4
0.5
1995
1996
1997
1998
1999

Figure 3


Proportion of Black Drivers Found with Drugs
proportion
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1995
1996
1997
1998
1999

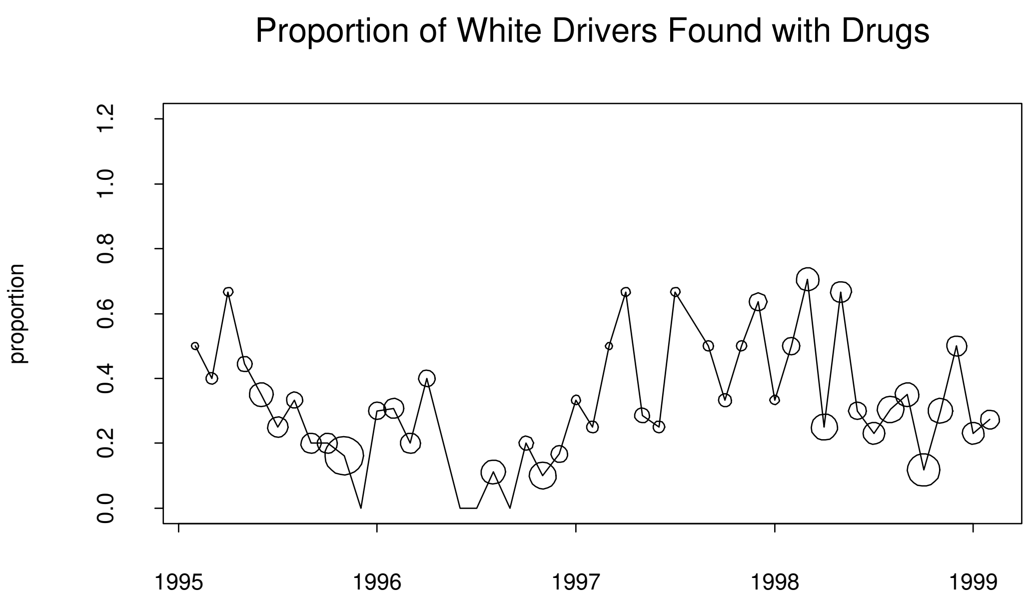

Proportion of White Drivers Found with Drugs
proportion
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1995
1996
1997
1998
1999

Figure 4

Number of Cars Searched by Maryland Troopers on I-95