A New Look at Racial Profiling: Evidence from the Boston Police Department

Kate L. Antonovics and Brian G. Knight[*†]

January 17, 2007

Abstract

This paper provides new evidence on the role of preference-based versus statistical discrimination in racial profiling using a unique data set that includes the race of both the motorist and the officer. We build upon the model presented in Knowles, Persico and Todd (2001) and develop a new test for distinguishing between preference-based and statistical discrimination. In particular, we show that if statistical discrimination alone explains differences in the rate at which the vehicles of drivers of different races are searched, then, all else equal, search decisions should be independent of officer race. We then test this prediction using data from the Boston Police Department. Consistent with preference-based discrimination, our baseline results demonstrate that officers are more likely to conduct a search if the race of the officer differs from the race of the driver. We then investigate and rule out two alternative explanations for our findings: officers are better at searching members of their own racial group and the non-random assignment of officers to neighborhoods.

[*] Respectively, Department of Economics, University of California, San Diego, email: kantonov@ucsd.edu, and Department of Economics, Brown University, email: Brian_Knight@brown.edu.

[†] We thank Peter Arcidiacono, Eli Berman, Richard Carson, Kim Sau Chung, Hanming Fang, Arthur Goldberger, Roger Gordon, Nora Gordon, Winfried Koeniger, Finis Welch and an anonymous referee for their comments. We have also benefitted from discussions with Amy Farrall at Northeastern University and Carl Walter at the Boston Police Department. Finally, we are indebted to Bill Dedman at *The Boston Globe* for providing us with our data.

A New Look at Racial Profiling: Evidence from the Boston Police Department

To date, there have been over 200 court cases involving allegations of racial and ethnic profiling against law enforcement agencies in the United States. Typically, the focus in these cases has been on uncovering why law enforcement officials treat individuals from different racial groups differently. The courts have tended to uphold racially biased policing patterns when they can be reasonably justified by racial differences in crime rates, but have consistently ruled against what appear to be purely racist policing practices. The problem, of course, is that it is not easy to empirically distinguish between these two possibilities.

Economists have now joined the debate over racial profiling, and a number of recent papers have attempted to determine whether the observed racial disparities in policing patterns are best explained by models of statistical discrimination or by models of preference-based discrimination (see, for example, Knowles, Persico and Todd (2001), Hernández-Murillo and Knowles (2003), Anwar and Fang (2006) and Dharmapala and Ross (2004)).

Statistical discrimination arises because law enforcement officials are uncertain about whether a suspect has committed a particular crime. If there are racial differences in the propensity to commit that crime, then the police may rationally treat individuals from different racial groups differently. On the other hand, preference-based discrimination arises because the police have discriminatory preferences against members of a particular group and act as if there is some non-monetary benefit associated with arresting or detaining members of that group. Thus, preference-based discrimination raises the benefit (or, equivalently, lowers the cost) of searching motorists from one group relative to those from some other group.[1]

This debate among economists over the sources of racial disparities in policing patterns roughly parallels the debate over racial profiling within the court system. That is, statistical discrimination approximately corresponds to the type of behavior that the courts have tended to uphold, while preference-based discrimination approximately corresponds to the type of behavior that the courts have tended to condemn.

In this paper, we attempt to understand the reasons for observed racial differences in the rate at which the vehicles of African-American, Hispanic and white motorists are searched during traffic stops. We build upon the model of police search developed in Knowles, Persico and Todd (2001) (hereafter, often, KPT) and develop an alternative mechanism for distin-

[1] For an extended discussion of models of statistical discrimination and of preference-based discrimination see Arrow (1973) and Becker (1954), respectively.

guishing between these two forms of discrimination that does not rely upon the probability of guilt conditional on search. In particular, we show that if statistical discrimination alone explains differences in the rate at which African-American and white drivers are searched, then, all else equal, search decisions should be independent of the race of the police officer. Thus, we argue that if searches are more likely to occur when the race of the officer differs from the race of the driver, then this provides evidence of preference-based discrimination.

We then apply our test to a unique data set in which we are able to match the race of the officer to the race of the driver for every traffic stop made by officers in the Boston Police Department for the two-year period starting in April 2001.[2] Thus, in addition to being able to discern differences in the likelihood that motorists from different racial groups are subject to search, we are also able to determine whether these patterns differ depending on the race of the officer. We find that if the race of the officer differs from the race of the motorist, then the officer is more likely to conduct a search than otherwise. We argue that our results cannot be explained by standard models of statistical discrimination and, instead, are consistent with preference-based discrimination. In addition, we rule out the possibility that our findings are driven by officers being better able to search members of their own racial group and by the way in which officers are assigned to neighborhoods.

Some Initial Trends in the Data

In order to motivate our model and the analysis that follows, it is worthwhile to first highlight a few patterns in our data. For now, these patterns are merely meant to be suggestive, and we will discuss the data in greater detail below.

Table 1 presents, by officer race and motorist race, the probability that a motorist's car is searched during a traffic stop. Looking at the last column, we see that both Hispanics and blacks are almost twice as likely as are whites to have their cars searched. This differential search pattern could be the result of preference-based discrimination. However, it is also consistent with statistical discrimination. That is, if blacks and Hispanics are more likely to carry drugs or other contraband than are whites, then it is also possible that they are also more likely than whites to raise the suspicion of the police. Thus, the last column of Table 1 simply reiterates the well-known fact that racial disparities in search rates exist, but does not offer any insight into why those disparities might arise.

Columns 2-4, however, are more revealing and provide some evidence that motorists are

[2] For an alternative discussion of these data, see the series of articles by Bill Dedman and Francie Latour (2003).

more likely to be searched if the officer making the stop is from a different racial group than the motorist. For example, the probability that a white motorist is searched is 0.40 percent if the officer is white and 0.62 percent if the officer is black. Similarly, the probability that a black motorist is searched is 0.82 percent if the officer is black but 0.97 percent if the officer is white. Interestingly, the table also reveals that black motorists and white motorists are *more* likely to be searched by an officer from their own racial group than by a Hispanic officer. This pattern is hard to interpret however, since it may partially reflect the fact that Hispanic officers are less likely to search motorists from any racial group than are African-American and white officers.

In order to insure that the patterns in Table 1 are not driven by a small number of officers who issue an unusually large number of tickets, Table 2 weights each citation by the inverse of the number of citations given by the officer issuing the citation. Thus, Table 2 effectively gives the mean search rate across officers giving equal weight to each officer who made at least one stop. Since officers who issue a large number of tickets are less likely to conduct searches than officers who issue a small number of tickets, the search probabilities are generally larger in Table 2 than in Table 1. However, as in Table 1, we see that motorists tend to be searched at lower rates when the officer making the stop is a member of the motorist's own racial group than when there is a mismatch between the race of the officer and the race of the motorist.

Abstracting at this stage from issues of statistical significance and other possible concerns, the patterns in Tables 1 and 2 appear to be inconsistent with standard models of statistical discrimination in which racial differences in the rate at which motorists are searched arise because the police believe that motorists from some racial groups are more likely to have contraband than are motorists from other groups. Assuming that these beliefs must be correct in equilibrium, there should be no difference in the rate at which officers from different racial groups search the vehicles of motorists from a particular racial group. On the other hand, preference-based discrimination could explain these patterns. In particular, if officers favor members of their own racial group, then we would expect search rates to be lower when there is a match between the race of the officer and the race of the motorist.

However, two alternative explanations for the patterns in Tables 1 and 2 also come to mind. First, officers may be better able to search motorists who are members of their own racial group. Second, officers may not be randomly assigned to neighborhoods. For example, if white officers are assigned to neighborhoods in which crimes are more likely to be committed by blacks than whites, and if black officers are assigned to neighborhoods in which crimes are more likely to be committed by whites than blacks, then we might expect that, for the city as a whole, white officers would be more likely than black officers to search the cars of black motorists. We address both of these alternative explanations in the final sections of the paper.

The Model

In this section, we develop a simple model of police search behavior and propose a test to distinguish between preference-based discrimination and statistical discrimination.[3] We then relate our test to the existing literature.

In the model, individuals are either African-American or white, denoted by a and w, respectively. In addition, motorists are distinguished by some characteristic, c, that is potentially useful to the police in determining whether or not to search a motorist's car. For now, we assume that both the police and the econometrician observe these driver characteristics; the case in which the econometrician does not observe driver characteristics is investigated in the next section.

In deciding whether or not to carry contraband, motorists weigh the benefit of carrying contraband against the penalty of being caught. If a motorist does not carry contraband, then his payoff is zero regardless of whether or not his car is searched. If a motorist of type (c,r) does carry contraband, then he incurs an idiosyncratic carrying cost Z, which is distributed in the population according to the function $G(\cdot)$. In addition, he faces cost $j(c,r)$ if his car is searched and benefit $\nu(c,r)$ if his car is not searched. Letting $\gamma^j(c,r)$ denote the probability that officers from group j search motorists from group r and letting ρ denote the proportion of officers who are African-American, the expected payoff to carrying contraband for a motorist of type (c,r) with idiosyncratic cost Z is given by

$$-\gamma(c,r)j(c,r) + [1-\gamma(c,r)]\nu(c,r) - Z,$$

where $\gamma(c,r) = \rho\gamma^a(c,r) + (1-\rho)\gamma^w(c,r)$.

The police cannot perfectly observe whether a motorist of type (c,r) is guilty of carrying contraband. Instead, it is assumed that police maximize the expected payoff from making an arrest, which is normalized to one, minus the cost of search, which is assumed to depend on the match between the race of the officer and the race of the motorist. Let $t_r^j \in (0,1)$ denote the cost to officers from group j of searching motorists from group r. Finally, let U denote a mean-zero idiosyncratic search cost, and let $H(\cdot)$ denote the distribution of such

[3] We follow the notation used in Knowles, Persico and Todd (2001).

costs across officers.[4,5]. Denoting $\pi(c,r)$ as the probability that a motorist of type (c,r) is guilty of carrying contraband, the payoff to officers from group j of searching motorists of type (c,r) can be written as

$$\pi(c,r) - t_r^j - U.$$

An equilibrium for motorists of type (c,r) occurs whenever police officers are playing a best response to motorists and whenever motorists are playing a best response to the average behavior of police. That is, an equilibrium for motorists of type (c,r) occurs at any $\pi^*(c,r)$, $\gamma^{a*}(c,r)$, and $\gamma^{w*}(c,r)$ such that

$$\gamma^{a*}(c,r) = H(\pi^*(c,r) - t_r^a)$$

$$\gamma^{w*}(c,r) = H(\pi^*(c,r) - t_r^w)$$

and

$$\pi^*(c,r) = G(-\gamma^*(c,r)j(c,r) + [1-\gamma^*(c,r)]\nu(c,r))$$

where $\gamma^*(c,r) = \rho\gamma^{a*}(c,r) + (1-\rho)\gamma^{w*}(c,r)$.

We now examine how preference-based discrimination and statistical discrimination influence the probability that officers search motorists and the probability that motorists carry contraband. Police officers in this model are defined to have racially discriminatory preferences if the cost of search depends on the race of the motorist, so that $t_a^j \neq t_w^j$ for $j = a, w$. To see how such preference-based discrimination affects the equilibrium outcome in this model, Figure 1 displays the equilibrium outcome for African-American motorists with characteristic c under the assumption that average search costs are the same for officers from different racial groups (so that $t_w^w + t_a^w = t_a^a + t_w^a$) and officers from at least one group of officers have discriminatory preferences against drivers of the other race (so that $t_a^a < t_w^a$ or $t_w^w < t_a^w$). It is easy to verify that these assumptions imply $t_a^w < t_a^a$, so that the cost of searching African-American motorists is lower for white officers than it is for African-American officers.[6] As Figure 1 reveals, given these assumptions, white officers are more likely than African-American officers to search African-American motorists in equilibrium.

On the other hand, if average search costs are the same across different racial groups, then in the absence of racially discriminatory preferences ($t_a^a = t_w^a$ and $t_w^w = t_a^w$), it is clear that

[4] Interestingly, our data suggest that officers do vary in their preferences for search. We observe substantial variation in the likelihood that officers search motorists whom they have pulled over.

[5] Note that we do not incorporate a resource constraint on the total time that police spend searching; that is, officers can search all drivers if they so choose. The implications of this assumption, which is also employed in the baseline model of Knowles, Persico and Todd (2001), is discussed below.

[6] These assumptions also imply that $t_w^a < t_w^w$, so that the cost of searching white motorists is lower for African-American officers than it is for white officers.

$t_a^a = t_a^w$, so that, in equilibrium, white and African-American officers will be equally likely to search African-American motorists. Thus, one implication of our model is that, controlling for average differences in the cost of search between African-American and white officers, there should be no difference in the rate at which officers from different racial groups search drivers of any given race in the absence of preference-based discrimination. This insight forms the basis of the empirical strategy that we employ.

In contrast to preference-based discrimination, statistical discrimination arises whenever $-j(c, r)$ and $\nu(c, r)$ vary by r, so that there exist racial differences in the net benefit of carrying contraband. To see how statistical discrimination affects the equilibrium in this model, Figure 2 shows the case in which, even among motorists with the same observable characteristic, c, the net benefit of carrying contraband is higher for African Americans than it is for whites, but the average cost of search is assumed to be constant, so that $t_r^j = t$, $\forall j, r$. In this case, for any given search probability, γ, African Americans are more likely than whites to carry contraband, and, as the figure reveals, in equilibrium, the police are more likely to search African Americans than whites ($\gamma^*(c, a) > \gamma^*(c, w)$). Note, however, that as long as there are no racial differences in the average cost of search ($t_w^w + t_a^w = t_a^a + t_w^a$) and as long as officers do not have discriminatory preferences ($t_a^j = t_w^j$, $\forall j$), the probability of search will not depend upon the interaction between the race of the driver and the race of the motorist.

Thus, in this paper, we test for the presence of preference-based discrimination by examining whether, controlling for average differences in search costs across officers, the likelihood that officers search drivers from a given racial group depends upon the race of the officer. If so, this suggests that differences in search rates arise because of preference-based rather than statistical discrimination. We discuss the details of our empirical methods in the next section.

In independent work, Anwar and Fang (2006), whose paper we became aware of after developing the first draft of our paper, employ a similar test to distinguish between statistical discrimination and preference-based discrimination. Their test is like ours in that it requires information on both the race of the police officer and the race of the motorist. In contrast to our model, however, Anwar and Fang assume that c is not perfectly known to the motorist at the time the motorist decides whether to carry contraband. Rather, c is a random variable whose distribution depends upon the motorist's decision to carry contraband, and, thus, c serves as a noisy signal to officers of likelihood that the motorist is guilty. From this behavioral model, Anwar and Fang develop a test that employs information on both the probability that motorists are searched and the probability that motorists are found to be carrying contraband conditional on being searched.[7] It is not obvious *a priori* that either our test or the Anwar and Fang test dominates the other. Rather, since each test is derived from a specific behavioral model, the appropriateness of each test depends upon the appropriateness of the underlying behavioral assumptions. Interestingly, when applied to our data, both tests provide support for the existence of discriminatory preferences among police.[8]

Figure 2 also reveals that the probability that motorists from different racial groups carry contraband differs even in the absence of preference-based discrimination. This contrasts with the prediction of Knowles, Persico and Todd (2001) in which the probability of guilt conditional on search will be the same for all motorists in the absence of preference-based discrimination. Based on this prediction, KPT propose to test for the presence of preference-based discrimination by examining whether the probability of guilt conditional on search differs across racial groups. As it turns out, the model in KPT is a special case of the model presented above in which 1) there is no heterogeneity in officer search costs ($U = 0$ for every officer) and 2) the cost of searching motorists from group r is the same for all officers so that $t_r^a = t_r^w = t_r$ for $r = a, w$. Without heterogeneity in officer search costs, the officer best response function in Figure 2 is step-shaped, and their prediction thus holds. As Figure 2 suggests, however, any alteration that smooths out the officer best response function may invalidate their test.[9]

Interestingly, Persico and Todd (2004) show that if officers face binding resource constraints, then even in the presence of heterogeneity in officer search costs, the probability of guilt conditional on search still will be the same for all motorists in the absence of preference-based discrimination. The intuition is that when officers are constrained, officers will focus their search activities on the group for which the probability of guilt conditional on search is the highest and this, in turn, will lower the net benefit to carrying drugs and the likelihood that motorists do so. As a result, the probability of guilt conditional on search will be the same for all motorists.

Whether or not officers face binding resource constraints in practice is an open question and may vary from application to application. Like us, Anwar and Fang (2006) and Bjerk

[7]Due to the limited number of searches in our dataset, we are only able to make limited use of information on conditional guilt probabilities.

[8]In addition to these differences in methodology, there are important differences in the nature of the data used in the two papers. Our dataset comes from the city of the Boston, and many of the traffic stops occurred on neighborhood streets. Traffic stops analyzed in Anwar and Fang, by contrast, come exclusively from highways.

[9]Other alterations to the model presented in KPT can also lead to differences in the probability of guilt conditional on search in the absence of preference-based discrimination. See, for example, Anwar and Fang (2006), Bjerk (2005) and Dharmapala and Ross (2004).

(2005) assume that officers do not face binding capacity constraints, and we feel that this assumption is appropriate for our analysis of traffic stops by the Boston Police Department. In our data, searches are rare: as shown in Table 1, under 1,000 searches occurred in Boston over the period April 1, 2001-January 31, 2003. Moreover, search rates vary significantly across officers, suggesting that resource constraints are unlikely to be binding for the many officers who search at low intensity in our data.

Empirical Strategy

In this section, we discuss how we test our model's prediction. Recall that our model predicts that, controlling for average racial differences in officer search costs, there should be no difference in the rate at which officers from different racial groups search motorists of type (c, r) in the absence of preference-based discrimination. Thus, assuming that the officer's race, the motorist's race, and c are known, this implication can be tested. Below we discuss what happens if c is unobserved. However, in order to establish the link between our model and our empirical strategy, it is useful to start with the case in which c is observed. Also, while our empirical analysis includes Hispanics, we focus on the simpler two-groups case of whites and blacks in this section. It should be clear, however, that all of the results here are similar in a framework with three groups.

In order to motivate the probit model that we employ, note that officers of race j search drivers of race r and characteristics c with the following probability:

$$\Pr(search|j, c, r) = H\left[\pi(c, r) - t_r^j\right].$$

Note that equilibrium guilty probabilities $[\pi(c, r)]$ are independent of officer race, which is revealed to drivers in the model only after the decision over whether or not to carry contraband has been made; this independence is key to our identification strategy.[10]

In order to estimate preference-based discrimination separately by officer race, we would ideally estimate a Probit model with a full set of interactions between officer race and driver race. Unfortunately, this fully specified model is perfectly collinear and thus cannot be estimated.[11] We can, however, feasibly estimate the following restricted Probit model:

[10] Drivers do, however, know the distribution of officer race, which is assumed to be the same for all motorists.

[11] To see this, consider the following fully-specified probit model:

$$\Pr(search|j, c, r) = H\left(\beta_0 + \beta_1 c + \beta_2 1[j = a] + \beta_3 1[r = a] + \beta_4 1[j = a, r = w] + \beta_5 1[j = w, r = a]\right),$$

Then, take the difference between the final two regressors:

$$\Delta = 1[j = a] \times 1[r = w] - 1[j = w] \times 1[r = a] = 1[j = a] - 1[r = a]$$

$$\Pr(search|j,c,r) = H\left(\beta_0 + \beta_1 c + \beta_2 1[j=a] + \beta_3 1[r=a] + \beta_4 mismatch\right),$$

where mismatch $= 1[j=a, r=w] + 1[j=w, r=a]$ indicates a traffic stop in which the race of the officer differs from the race of the driver. Given that we cannot identify racial prejudice separately for African-American and white officers, we assume that they are equally prejudiced ($t_a^a - t_w^a = t_w^w - t_a^w$).[12] Under this assumption, we can write the following relationships between the theoretical and empirical specifications for this model.[13]

Relationship	Interpretation
$\beta_2 = t_w^w - t_a^a$	Cost differences by officer race
$\beta_3 = \pi(c,a) - \pi(c,w)$	Statistical discrimination
$\beta_4 = t_a^a - t_w^a = t_w^w - t_a^w$	Racial prejudice

With data on the race of both the driver and the officer, we can thus distinguish between racial profiling based upon statistical discrimination, which is captured by the coefficient on driver race (β_3), and racial profiling based upon prejudice, which is captured by the coefficient on mismatch (β_4).

An implicit assumption underlying this Probit formulation is that search rates are separable in driver characteristics (c) and driver race ($1[r=a]$). That is, officers do not condition on driver characteristics in a manner that differs between black and white drivers. While this formulation appears to be restrictive, it is straightforward to incorporate an interaction between driver characteristics and driver race ($c \times 1[r=a]$) into the econometric specification. Let β_5 be the coefficient on this interaction term. Then, the above relationships are identical except for the expression for statistical discrimination $[\pi(c,a) - \pi(c,w)]$, which was previously equal to β_3 and now equals $\beta_3 + \beta_5 c$. Intuitively, any conditioning on driver characteristics that differs between African and white drivers should be considered statistical discrimination. Importantly, however, the interpretation of the coefficient on mismatch is unchanged $[\beta_4 = t_a^a - t_w^a = t_w^w - t_a^w]$ and still captures racial prejudice. As noted below,

Thus, this difference (Δ) equals a linear combination of the first two regressors. Our inability to estimate this fully specified model is not surprising since, even if c is a constant, there are only four possible cases of driver / officer interactions but five parameters.

[12] If blacks and whites are not equally prejudiced, then our estimates will uncover the average level of prejudice across black and white officers.

[13] In order to derive these relationships, consider the following four possible cases of driver/officer interactions for both the theoretical and empirical models: 1) $j=w, r=w$, 2) $j=a, r=a$,3) $j=w, r=a$,4) $j=a, r=w$. One can then show that $\beta_0 + \beta_1 c = \pi(c,w) + t_w^w$. Using this relationship, the three key parameters can then be solved for.

the coefficient on mismatch is positive in our empirical application even after including these interactions between driver race and driver characteristics.

Consider next the case in which driver characteristics (c) are unobserved to the econometrician. We show below that, under assumptions of normality and random matching of officers and drivers, our approach retains the ability to distinguish between racial prejudice and statistical discrimination *even if unobserved driver characteristics are correlated with driver race*. Intuitively, the coefficient on driver race absorbs any unobserved differences between black and white drivers, and the coefficient on mismatch is thus not contaminated by the presence of these unobserved characteristics.

Recall that, according to the probit model, officers search if the following expression holds:

$$\beta_0 + \beta_1 c + \beta_2 1[j = a] + \beta_3 1[r = a] + \beta_4 mismatch - U > 0$$

where $U \sim N(0, 1)$. Assume next that unobserved driver characteristics are normally distributed with a mean that varies by race:

$$c = c_r - \sigma\varepsilon, \; r = a, w$$

where $\varepsilon \sim N(0, 1)$ and is assumed to be independent of both driver race (r) and officer characteristics (U, j).[14,15] We refer to the assumption of independence between unobserved driver characteristics and mismatch as random matching. This random matching assumption will be satisfied if ε is independent of mismatch; we will describe below what is identified under a special case in which this assumption is violated.

Substituting in the above expression for unobserved driver characteristics, officers of race j search drivers of race r if:

$$\beta_0 + \beta_1 c_w + \beta_2 1[j = a] + [\beta_3 + \beta_1(c_a - c_w)]1[r = a] + \beta_4 mismatch - U - \sigma\beta_1\varepsilon > 0$$

[14] The assumption that c is a scalar is not crucial and can be generalized. In particular, allow an $N \times 1$ vector of unobserved driver characteristics (C) to vary according to driver race and a random vector: $C = C_r - E$, where C_r and E are both $N \times 1$ vectors, and the components of E are assumed to be distributed jointly normal with covariance matrix Σ. In this case, the unconditional probit can be written as follows:

$$\Pr(search|j, r) = H\left[\frac{\beta_0 + \beta_1 C_w + \beta_2 1[j = a] + [\beta_3 + \beta_1(C_a - C_w)]1[r = a] + \beta_4 mismatch}{\sqrt{1 + \beta_1 \Sigma \beta_1'}}\right]$$

where β_1 is now a $1 \times N$ vector.

[15] As shown in Yatchew and Griliches (1985), without the normality assumption, which is made here for reasons of tractability, the presence of unobserved characteristics leads to complicated asymptomatic bias formulas in probit models. In particular, the asymptotic bias formulas depend on the cumulative distribution function for unobserved characteristics. Applying this lesson to our analysis, if traffic stops in which the race of the driver differs from the race of the officer are also stops in which drivers disproportionately carry contraband, then the coefficient on mismatch could be asympotically biased in either direction.

Under the assumption that U and ε are independently distributed, $U - \sigma\beta_1\varepsilon \sim N(0, 1+\beta_1^2\sigma^2)$ and the probability of search, unconditional on driver characteristics, is given as follows:

$$\Pr(search|j,r) = H\left[\frac{\beta_0 + \beta_1 c_w + \beta_2 1[j=a] + [\beta_3 + \beta_1(c_a - c_w)]1[r=a] + \beta_4 mismatch}{\sqrt{1+\beta_1^2\sigma^2}}\right],$$

We can thus define the unconditional probit parameters $(\gamma_0, \gamma_2, \gamma_3, \gamma_4)$ as follows:

$$\begin{aligned}
\gamma_0 &= \frac{\beta_0 + \beta_1 c_w}{\sqrt{1+\beta_1^2\sigma^2}} \\
\gamma_2 &= \frac{\beta_2}{\sqrt{1+\beta_1^2\sigma^2}} \\
\gamma_3 &= \frac{\beta_3 + \beta_1(c_a - c_w)}{\sqrt{1+\beta_1^2\sigma^2}} \\
\gamma_4 &= \frac{\beta_4}{\sqrt{1+\beta_1^2\sigma^2}}
\end{aligned}$$

Using these definitions and the relationships listed above between the theoretical parameters and the probit parameters conditional on driver characteristics, we can thus relate the probit parameters unconditional on driver characteristics to the theoretical parameters as follows:

Relationship	Interpretation
$\gamma_2 = (t_w^w - t_a^a)/\sqrt{1+\beta_1^2\sigma^2}$	Cost differences by officer race
$\gamma_3 = [\pi(c,a) - \pi(c,w) + \beta_1(c_a - c_w)]/\sqrt{1+\beta_1^2\sigma^2}$	Statistical discrimination
$\gamma_4 = (t_a^a - t_w^a)/\sqrt{1+\beta_1^2\sigma^2} = (t_w^w - t_a^w)/\sqrt{1+\beta_1^2\sigma^2}$	Racial prejudice

These relationships yield several key insights. First, results from the case in which the econometrician observes and does not observe driver characteristics are identical if officers do not rely on driver characteristics in their search decisions ($\beta_1 = 0$). In addition, if there is no heterogeneity other than race in unobserved characteristics ($\sigma = 0$), then the coefficients on officer race and mismatch are unchanged. The coefficient on driver race (γ_3), however, is altered and now captures both statistical discrimination based purely upon race $[\pi(c,a) - \pi(c,w)]$ and statistical discrimination based upon driver characteristics that vary according to race ($\beta_1(c_a - c_w)$); without further information, we cannot distinguish between these two forms of statistical discrimination. However, even if $\beta_1 \neq 0$ and $\sigma \neq 0$, our approach retains the ability to distinguish between statistical discrimination, in whatever form it may take, and racial prejudice ($\gamma_4 = (t_a^a - t_w^a)/\sqrt{1+\beta_1^2\sigma^2}$). In fact, the presence of unobserved driver characteristics only serves to bias our analysis away from measuring racial prejudice due to the scaling factor ($\sqrt{1+\beta_1^2\sigma^2}$), which exceeds one.

Without the assumption of random matching, our empirical strategy may no longer directly measure racial prejudice. While any dependence may invalidate our test, we focus here on differences in the mean of unobserved characteristics as this is the most tractable case.[16] In particular, if $c = c_r + \eta\text{mismatch} - \sigma\varepsilon$, then the probit specification is given as follows:

$$\Pr(search|j,r) = H\left[\frac{\beta_0 + \beta_1 c_w + \beta_2 1[j=a] + [\beta_3 + \beta_1(c_a - c_w)]1[r=a] + (\beta_4 + \beta_1\eta)mismatch}{\sqrt{1+\beta_1^2\sigma^2}}\right]$$

Thus, the coefficient on mismatch will capture both racial prejudice (β_4) and non-random matching ($\beta_1\eta$), and the assumption of random matching is crucial to our identification strategy. This assumption could be violated, for example, if officers are assigned to neighborhoods in which drivers are predominantly of the other race and also have unobserved characteristics that make them more likely to carry drugs. We address this issue in the empirical analysis to follow by studying how police officers are assigned to neighborhoods in Boston.[17]

Data

In July 2000, the Massachusetts legislature passed Chapter 228 of the Acts of 2000, *An Act Providing for the Collection of Data Relative to Traffic Stops*. Among other things, this statute required that, effective April 1, 2001, the Registry of Motor Vehicles collect data on the identifying characteristics of all individuals who receive a citation or who are arrested. The data collected by the State contain a wide variety of information including: the age, race and gender of the driver, the year, make and model of the car, the time, date and location of the stop, the alleged traffic infraction, whether a search was initiated and whether the stop resulted in an arrest.

The statute also required the Registry of Motor Vehicles to collect data on warnings. However, citing budgetary shortfalls, the Registry only compiled data on warnings for two months. Thus, for most of the time period under investigation, we do not observe stops for which an officer merely issued a written or verbal warning. That is, unless an officer issued a citation, the stop does not appear in our data outside of the two-month period. We will address this data limitation in the empirical results section to follow.

We were also able to obtain officer-level data from the Boston Police Department. These

[16] We have also investigated the case in which the variance of unobserved characteristics depends upon mismatch. In this case, $c = c_r - (\sigma + \eta mismatch)\varepsilon$, and mismatch enters the Probit specification in a non-linear manner. Thus, a direct link to the Probit specification that we estimate, in which mismatch enters linearly, is no longer possible.

[17] Anwar and Fang (2006) also note that problems may arise if officers have more information about motorists from their own racial group than about motorists from other racial groups. Thus, mismatch must also be independent of the amount of information available to officers.

data contain, among other things, information on the officer's race, gender, rank and number of years on the force. For the subset of citations issued by officers in the Boston Police Department, we are then able to match the officer-level data to the citation-level data collected by the state. In total, we are able to match officer-level data to over 112,473 citations issued by 1,369 officers, representing just over 80 percent of the citations issued by officers in the Boston Police Department in our data. That is, for approximately 20 percent of the citations issued by an officer in the Boston Police Department in our data, we were unable to identify the officer who issued the citation.

We restrict our sample in a number of ways. First, we delete the 6 citations for which contradictory race information was recorded. In addition, we drop citations issued by Asian officers (23 officers in total), and 7,732 citations issued to Asian, Native American and Middle Eastern motorists. As a result, all of the motorists and officers in our data are either black, white or Hispanic. We also drop the ten citations that were issued to motorists outside the city of Boston. This may have happened, for example, if an officer followed a speeding driver outside of the city limits. Finally, we drop about 4,500 observations with missing information on the race, age and residence of the driver and whether an accident occurred. Once these restrictions have been made we are left with 95,855 citations issued by 1,317 officers.

Of considerable concern is the fact that the search variable is missing for over 18 percent of the citations in our data. When filling out a citation, officers are required to check either "yes" or "no" to indicate whether a search was conducted. If an officer neglected to check either box, then the search variable is missing in our data. We do not know why officers failed to check this box. One possibility is that they were careless. Another is that they did not fully understand how to fill out the citation and generally only checked the "yes" box if they conducted a search but otherwise left the question blank. There is substantial variation across officers in the proportion of citations for which the search variable is left missing; some officers appear to have been better at accurately filling out the citation than others. There are a number of ways of dealing with these missing values. We pick the method that we think is the best and then check to see if our results are robust to alternative procedures. In our baseline specification, if the officer indicated that a search was conducted for all citations in which search was non-missing, then we assume that when the search variable is missing, no search was conducted. Then, we drop all officers for whom search is missing for more than 10 percent of the citations that those officers issue. Doing so eliminates approximately 24 percent of the citations (and 48 percent of the officers) in our data.[18] For the remaining

[18]In calculating the percentage of citations for which search is missing, we do not include citations in which the race of the driver is missing.

685 officers, we drop observations for which search is missing, and are left with a sample comprising 70,652 citations. Tables 1 and 2 are calculated using this sample.

Table 3 presents some basic summary statistics. The first column includes all of the citations for which our baseline search measure is missing, whereas the second column includes all of the citations for which our baseline search measure is available. Thus, comparing these first two columns provides some idea as to whether the citations for which search is missing differ systematically from those where it is not. Among citations for which search is missing, accidents are about twice as likely to have occurred as among citations for which search is not missing. There is also some variation across the first two columns in the percentage of citations that are issued in each neighborhood, reflecting the fact that officers in some districts were less likely to leave the search question blank than were officers in other districts. Otherwise, citations for which the search variable is missing appear to be quite similar to those for which it is not.[19] The last three columns of Table 3 show the average characteristics of the citations in our sample broken down by the race of the officer issuing the citation. We see that drivers are disproportionately issued citations by officers from their own racial group. As we will see below, this may reflect the fact that officers are more likely to issue tickets in districts in which a large portion of the population (and so, presumably, the drivers) are in the same racial group as the officer. Indeed, this is also reflected in the fact that there is variation across the last three columns in the proportion of citations issued in different neighborhoods. Finally, we see that black officers are more likely to issue citations at night and less likely to issue citations at which an accident has occurred than either white or Hispanic officers.

Search Patterns in the Boston Police Department

In this section we test our model's theoretical predictions. For the time being we abstract from the possibility that there exist racial differences in officers' abilities to assess the guilt of motorists from different racial groups and the possibility that officers may be non-randomly matched with motorists from different racial groups.

We start by replicating the results presented in KPT. To do so, we use a probit model to study the probability of search and the probability of guilt conditional on search. In order

[19] We also estimated probit models for whether or not the search variable was missing as a function of officer and driver characteristics. The mismatch coefficient turns out to be negative but statistically insignificant. This insignificance suggests that the omission of missing observations is not driving our results. Even if the coefficient were statistically significant, this result would only serve to bias us against finding preference-based discrimination under the assumption that non-searches were more likely to be coded as missing observations. That is, our data are missing non-searches in which the race of the officer and driver were likely to match.

to determine how the probability of search and the probability of guilt conditional on search differ depending on the race of the driver, we include indicator variables for whether the driver is black or Hispanic (so that white drivers are our omitted category). We also include as controls indicator variables for whether the stop occurred at night (6pm-5am), whether the driver was below the age of 26, whether the driver was male, whether the driver was from in state, whether the driver was from in town, and whether an accident had occurred. In addition, we include indicator variables that control for the district in which the stop occurred. In Table 4 (and in all remaining relevant tables) we report the coefficients of our probit model. Column 1 presents the results from the probit model of the probability of search, and column 2 presents the results for the probability of guilt conditional on search. In these first two columns, each citation receives equal weight. However, concern that these results are driven by a small number of officers who issue an unusually large number of citations prompted us to repeat the analysis in columns 1 and 2, but weight each citation by one over the number of citations given by the officer issuing that citation. The last two columns of Table 4 present the results of these weighted probits.

Our results are sometimes sensitive to whether or not we weight citations in this fashion. In fact, the merits of weighting depend upon the question that you wish to answer. If you are interested in understanding the behavior of the average officer, the weighted probits provide a better description of the data since officers who issue a large number of tickets do not exert a disproportionate impact on the estimates. On the other hand, if you are interested in understanding search outcomes for the average motorist who receives a citation, then the unweighted probits are more appropriate. In this paper, we are interested in understanding the search decisions of officers and, in particular, whether their behavior is consistent with preference-based discrimination. Thus, we believe that the results of the weighted probits are the most appropriate. For completeness, however, we present the results of both the weighted and unweighted probits for several specifications.

As the first column of Table 4 indicates, black drivers are more likely to have their cars searched than are white drivers. This result also holds for the weighted probit in column 3. Like Knowles, Persico and Todd, we find no evidence that the probability of guilt conditional on search differs by the race of the driver. In particular, in both columns 2 and 4, the coefficient of the indicator variable for whether the driver is not statistically different from zero. Given the small number of searches in our dataset, however, the standard errors are large and our test may thus fail to detect statistically significant differences even if they exist in the underlying population.

KPT interpret the finding that the probability of guilt conditional on search is identical across racial groups as evidence against preference-based discrimination. However, as the discussion in the preceding section highlights, once the model of KPT is generalized to allow for heterogeneity in officer search costs, this prediction no longer holds. As an alternative method for distinguishing between preference-based discrimination and statistical discrimination, our model predicts that if statistical discrimination alone explains differences in the rate at which African-American and white drivers are pulled over, then there should be no difference in the rate at which officers from different racial groups search drivers from any given racial group. To examine this, we again use a probit model to analyze the probability of search. Here, in addition to controlling for the race of the driver, we also include indicator variables for the race of the officer as well as an indicator variable that is equal to 1 if the race of the officer differs from the race of the driver (we call this indicator "mismatch"). Table 5 presents our results. In the first three columns, each citation receives equal weight, and each column includes a progressively broader set of controls. In the last three columns each citation is weighted by one over the number of citations given by the officer issuing the citation. In all six columns, the coefficient on our mismatch indicator is positive and statistically different from zero at standard significance levels. Thus, our results indicate that officers are more likely to search motorists who are not members of the officer's racial group. As mentioned before, this finding is inconsistent with standard models of statistical discrimination. Our results also suggest that Hispanic officers are more likely to conduct searches than are white officers, and the second and third columns suggest that officers are more likely to search motorists who are black, young or involved in an accident.[20]

As mentioned previously, our estimates may be biased if the mismatch between the race of the officer and the race of the motorist is correlated with motorist characteristics (c) that are not included in our regressions. Thus, it is comforting that the point estimate on mismatch changes very little as we add more regressors, suggesting that mismatch tends not to be correlated with unobserved motorist characteristics. We investigate this potential bias more fully below by empirically analyzing the assignment of officers to neighborhoods in Boston.[21]

[20] As noted in the empirical strategy section, we have also estimated models in which the effect of driver characteristics is allowed to vary between black and white drivers. These results, not reported here, are similar to those in the baseline analysis. That is, after controlling for the interaction between driver race and driver characteristics, the coefficient on mismatch remains positive and statistically significant.

[21] We have also investigated the possibility of estimating models with officer fixed effects. Given that searches are relatively rare events in our data, however, most of the variation in our data is across, rather than within, officers and these fixed effects models are not well identified. In particular, when we run analogous officer fixed effects logit models, which drop all observations without within-officer variation in the dependent variable, 64 percent of stops and 79 percent of officers are dropped, and the coefficient on mismatch becomes statistically insignificant, likely reflecting the loss of power associated with this significantly reduced sample size.

A positive coefficient on our mismatch variable could be driven either by discrimination on the part of white officers against black drivers or by discrimination on the part of black officers against white drivers. Thus, as noted above, we cannot separately identify differences in racial prejudice by officer race. Thus, for example, our results should not be taken as evidence that black motorists in the Boston area are the subject to discrimination by white officers. Rather, our results simply indicate that the interaction between the race of the motorist and the race of the officer is positively related to the probability that the motorist is searched, a pattern that is consistent with preference-based discrimination.

While the coefficient on mismatch is statistically significant in all of the specifications of Table 5, we should also evaluate the magnitude of these results. Conventional computation of the marginal effect of the variable mismatch is somewhat nonsensical as mismatch is an interaction term and thus changes in mismatch also require changes in either officer or driver race. Thus, in order to gauge the magnitude of these results, we calculate the probability of search for each driver who was stopped by an officer of different race under the counterfactual scenario in which the officer was instead of the same race as the driver. Using the coefficients in column 6 of Table 5, we calculate that searches would have occurred in 1.5 percent of these counterfactual own-race stops; this is substantially lower than the 3.5 percent predicted probability with which searches occurred using the actual officer race. [22]

As mentioned earlier, the search variable is missing for over 18 percent of the citations in our data. To see whether our results are sensitive to the way in which we treat these missing values, we conduct a number of robustness checks, the results of which are presented in Table 6. In the first column, we run the same basic specification as above with our full set of controls, but include in the analysis officers for whom the search variable is missing in more than 10 percent of the citations they issue. In the second column, we repeat the analysis in column 1 but assume that if search was missing, then no search was conducted. The motivation for this assumption is the notion that officers may be more likely to leave the search question blank if no search was conducted. This obviously increases our sample size substantially. Finally, in column 3, we repeat the analysis in column 1 except that if all of an officer's non-missing search citations indicate that a search was conducted, then we assume that no search was conducted for all of the missing observations. As shown, the point estimates drop in size relative to the comparable estimate using our baseline search measure. However, the mismatch coefficient remains statistically different from zero in both column 2

[22]Consistent with the raw averages in Table 1, the effect of changes in officer race associated with the unweighted results, those of column 3 of Table 7, are smaller in magnitude. In particular, searches would have occurred in 0.6 percent of these counterfactual own-race stops, relative to the 0.8 percent predicted probability with which searches occurred using the actual officer race.

and column 3.[23]

In addition, we have estimated a variety of other specifications, and the full results are available in Antonovics and Knight (2004). First, in order to focus on relations between blacks and whites, we deleted observations with either Hispanic drivers or officers. Interestingly, the coefficient on mismatch remains positive and statistically significant. Second, we restricted attention to citations that are issued by Patrol Officers; the remaining officers are some manner of either Deputies, Detectives, Sergeants or Captains. While the coefficients remain positive for this subsample, they are now statistically insignificant. One possible explanation for this result is that patrol officers tend to be less experienced than non-patrol officers and, as we discuss below, our results are weaker for inexperienced than experienced officers. Third, as noted above, our data include written warnings for the two-month period of April-May 2001. If officers tend to issue more warnings to drivers of their own race, then our data after this two-month period may include only the own-race interactions in which drivers had committed the most severe infractions. To address this concern, we restricted our sample to stops that occur within this two-month period. As expected, the coefficients on mismatch are larger than those in the baseline analysis, and the coefficients remain statistically significant at conventional levels. Fourth, as suggested by Grogger and Ridgeway (2004), we examine citations issued at night. The idea is that officers are less likely to know a motorist's race prior to pulling them over when it is dark outside and thus any selection into stops may be less problematic at night. Interestingly, the coefficient on mismatch is larger than in the baseline analysis and remains statistically significant.[24]

Asymmetric Search Ability

One concern is that our results may be driven by the fact that officers may be more successful at finding contraband in cars that are driven by motorists who are in their own racial group. For example, Donohue and Levitt (2001) find evidence that own-race policing may be more efficient than cross-race policing. To see how this would affect our model, let ϕ_r^j denote the probability that an officer from group j is successful of searching a motorist from group r;

[23] Recall that in our baseline search measure we drop officers for whom the search variable is missing for more than 10 percent of the citations issued by that officer. We have also experimented with changing that 10 percent cutoff. Lowering the cutoff (to say 5 percent or 3 percent), tends to strengthen our results, while increasing the cutoff tends to weaken them. This is reflected in column 1 of Tables 8 and 9 where the cutoff is effectively 100 percent (all officers are included).

[24] We also attempted to examine drivers who were pulled over for going more than 15mph and more than 20mph over the posted speed limit, since the police arguably have less discretion in whether or not to pull over these motorists. However, the results were sensitive to how we handled the large number of observations with missing information on mph and so we do not report them here.

the baseline model is one in which this probability equals one for all officers and drivers. In this generalized model, the payoff to an officer from group j to searching a motorist from group r is given by

$$\phi_r^j \pi(c,r) - t_r - U$$

The higher is ϕ_r^j, the *higher* will be the benefit to officers from group j of searching a motorist from group r. Thus, if officers are better at finding contraband when the motorist is a member of the officer's own racial group, then we would expect officers to be *more* likely to search motorists from their own racial group and our estimates will tend to *understate* the extent of preference-based discrimination.[25]

To empirically address the possibility that asymmetric information drives our results, we examine whether our results hold among officers with more than 10 years of experience. The idea is that if officers become better at searching motorists from a particular group as their exposure to that group increases, then, assuming there are decreasing returns to experience, officers with substantial experience should be equally able to search the cars of motorists from different racial groups.

Table 7 presents the results; the first three columns focus on citations issued by officers with less than 10 years of experience while the last three columns focus on citations issued by officers with 10 or more years of experience. We chose 10 years as our cutoff because it is close to the average experience level of officers in our data, approximately 12 years. However, our results are not sensitive to the exact cutoff experience level that we employ.

As shown, the coefficients on the mismatch indicator are small and statistically insignificant for inexperienced officers but large and statistically significant for experienced officers. Thus, these results suggests that our findings of preference-based discrimination are *not* driven by differences in the ability of officers to accurately search the cars of motorists from their own racial group. Rather, this analysis suggests that our results are stronger when we examine only experienced officers, for whom we would expect the likelihood of a successful search to be independent of the match between the officer's race and the driver's race.

[25] Like Anwar and Fang (2006) and Bjerk (2005) we have also considered a model in which officers observe a noisy signal of a motorist's guilt that is unknown to motorists at the time they make their decision about whether to carry contraband, but that is correlated with the likelihood that they carry contraband. One might expect officers to receive more informative signals from motorists who are in the officer's racial group than from those who are not. In an appendix available from the authors upon request, we show that changes in the information content of the signal deliver ambiguous predictions about search behavior.

The Assignment of Officers to Neighborhoods

As discussed in the section on the econometric specification, if there is some relevant characteristic, c, that is not included in our regressions and that is not independent of mismatch, then our estimate of β_4, the coefficient on mismatch, may be biased. One of the most plausible explanations for the source of this bias is that officers may not be randomly assigned to different neighborhoods within the city.

Suppose, for example, that white officers are disproportionately assigned to neighborhoods in which blacks commit a large fraction of the drug trafficking offenses. We would expect white officers to be more likely than black officers to search the cars of black motorists, even in the absence of preference-based discrimination. It seems unlikely that officers would be assigned to neighborhoods in this fashion, but it is worth examining how the Department allocates officers across the city.

Officers in the Boston Police Department are assigned to one of 11 districts. These districts correspond to well-defined geographic areas within the city and are the primary organizational units for the Department. Figure 3 indicates both the name and location of these 11 districts. In addition, the Boston Police Department has a "Same Cop Same Neighborhood" (or "SC/SN") policing policy. Under SC/SN, patrol officers are assigned to a neighborhood beat within each district, and spend no less than 60 percent of their shift in that beat. The intent of SC/SN is to enable officers to become familiar with the local community to which they are assigned and, thus, be more effective at preventing crime. While our data contain information on the district to which an officer was assigned at the time he or she issues a citation, we lack information on the officer's neighborhood beat.

In Table 8, we compare the racial composition of the population aged 18 and over in each district with the racial composition of the officers who are assigned to that district. As the table shows, in districts in which a relatively large percentage of the population is white, a relatively large proportion of the officers assigned to that district are white. Similarly, in districts in which a relatively large proportion of the population is black, a relatively large proportion of the officers assigned to that district is black, and the same pattern holds for Hispanics. For whites, the correlation between the fraction of the population aged 18 and older in each district who is white and the fraction of officers in that district who is white is 0.751. For blacks, Asians and Hispanics the analogous correlations are 0.844, 0.575 and 0.885, respectively. To some extent, these patterns may reflect intentional assignment patterns on the part of officials at the Boston Police Department. However, officers also have some discretion about the district to which they are assigned. In any case, officers appear to patrol areas in which the majority of residents are members of the officer's racial group.

However, even if officers tend to be assigned to districts in which the majority of residents are members of their own racial group, mismatch may not be independent of the likelihood that the motorist is guilty. For example, if African-Americans go to predominantly white neighborhoods to sell drugs and if whites go to predominantly African-American neighborhoods to buy drugs, then officers in predominantly white neighborhoods may deliberately target African-American motorists and officers in black predominantly black neighborhoods may deliberately target white motorists. In order to address this concern, Table 9 provides results from only those traffic stops in which the motorist was a resident of the district in which he or she was pulled over. Presumably, when an officer stops a motorist, the officer can observe (by looking at the motorist's driver's license) whether the motorist is a resident of the district and, thus, whether he or she is a "suspicious outsider". As the Table reveals, the coefficient on mismatch remains positive and is statistically different from zero in all three specifications.[26] Thus, even if we focus on stops that take place in the district in which the motorist resides, we find that officers are more likely to conduct a search if the motorist is not a member of the officer's racial group.

In addition, in Table 10, we break districts down into two categories: those that are racially diverse and those that are racially homogeneous. The idea is that in racially diverse neighborhoods, a driver's race is less likely to signal that the driver is out of place. Using Table 8, we categorize East Boston, Roxbury/Mission Hill, Jamaica Plain and Hyde Park as diverse, while we categorize Boston Central (A-1 and D-4), South Boston, Allston/Brighton and West Roxbury/Roslindale as homogeneous.[27] We drop citations that were issued in Dorchester and North Dorchester because these districts differ substantially in their racial composition and our data do not allow us to distinguish between the two. Interestingly, we find that the results are somewhat stronger in diverse neighborhoods, suggesting that our baseline results are not driven by the possibility that police target drivers whose race differs from that of the local population.[28]

Conclusion

The fact that minority drivers tend to be searched more frequently than whites during traffic stops has led to a significant legal and academic debate over the underlying motivations for these differences. One possible explanation is statistical discrimination. Another is

[26] Our results are qualitatively similar when we only examine blacks and whites.
[27] Charlestown is included in East Boston.
[28] Our results are qualitatively similar when we only examine blacks and whites.

preference-based discrimination. In this paper, we develop a new test for uncovering the motivations of police in searches during traffic stops.

We start by generalizing the model of police search developed in Knowles, Persico and Todd (2001) and develop a new test for distinguishing between preference-based discrimination and statistical discrimination. In particular, our model predicts that if statistical discrimination alone accounts for racial disparities in the rate at which motorists from different racial groups are subject to search, then there should be no difference in the rate at which *officers* from different racial groups search drivers from any given group.

We test this hypothesis using data from the Boston Police Department. Our results strongly suggest that officers are more likely to conduct a search if the race of the motorist differs from the race of the officer. We then test whether this pattern could be explained by differences in the ability of officers to search members of their own racial group. We find no evidence that this sort of search asymmetry explains our results. We also show that the manner by which officers are assigned to neighborhoods within the city does not account for our empirical findings. Rather, our results suggest that preference-based discrimination plays a substantial role in explaining differences in the rate at which motorists from different racial groups are searched during traffic stops.

References

[1] Antonovics, Kate and Knight, Brian "A New Look at Racial Profiling." NBER working paper 10634, 2004.

[2] Anwar, Shamena and Fang, Hanming "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence.", *American Economic Review*, March 2006, *96*(1), 127-151.

[3] Arrow, Kenneth "The Theory of Discrimination," in O. Ashenfelter and A. Rees, eds., *Discrimination in Labor Markets.* Princeton, NJ: Princeton University Press, 1973, 3-33.

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

[5] Bjerk, David. "Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate." *Mimeo, McMaster University,* March 2005.

[6] Dedman, Bill and Latour, Francie. "Traffic Citations Reveal Disparity in Police Searches." *The Boston Globe,* January 6, 2003, pp.A1.

[7] Dedman, Bill and Latour, Francie. "Race, Sex and Age Drive Ticketing." *The Boston Globe,* July 20, 2003.

[8] Dedman, Bill and Latour, Francie. "Punishment Varies by Town and Officer." *The Boston Globe,* July 21, 2003.

[9] Dedman, Bill and Latour, Francie. "Troopers Fair, Tough in Traffic Encounters." *The Boston Globe,* July 22, 2003.

[10] Dharmapala, Dhammika and Ross, Stephen L. "Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence." *Contributions to Economic Analysis and Policy,* 2004, *3*(1), 1-21.

[11] Donohue, John J. and Levitt, Steven D. "The Impact of Race on Policing and Arrests." *Journal of Law and Economics,* October 2001, 44, 367-394.

[12] Grogger, Jeffrey and Ridgeway, Greg. "Testing for Racial Profiling in Traffic Stops from Behind a Veil of Darkness."*Journal of the American Statistical Association,* September 2006, *101*(475), 878-87.

[13] Hernández-Murillo, Ruben and Knowles, John. "Racial Profiling or Racist Policing?: Testing in Aggregated Data." *International Economic Review,* August 2004, *45*(3), 959-89.

[14] Knowles, John; Persico Nicola; and Todd, Petra. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy,* February 2001, *109*(11), 203-229.

[15] Persico, Nicola and Todd, Petra. "Using Hit Rate Tests to Test for Racial Bias in Law Enforcement: Vehicle Searches in Wichita," NBER Working Paper 10947, December 2004.

[16] Yatchew, Adonis and Griliches, Zvi. "Specification Error in Probit Models." *Review of Economic Studies,* February 1985, *67*(1), 134-39.

Figure 1: Equilibrium Outcome with Preference-Based Discrimination

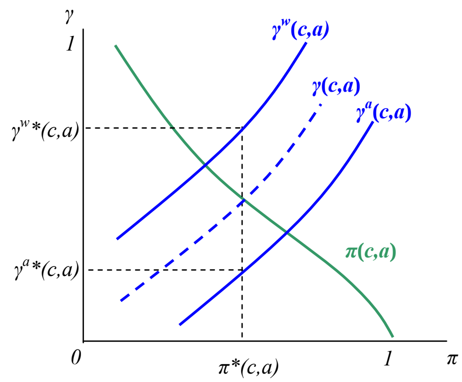


Figure 2: Equilibrium Outcome with Statistical Discrimination

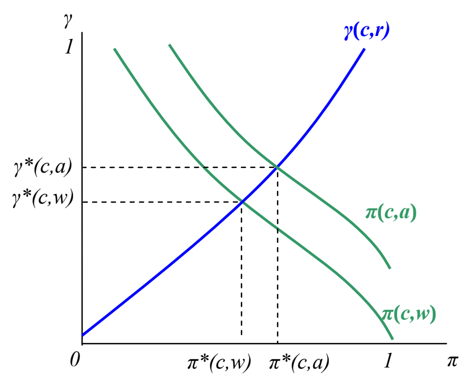
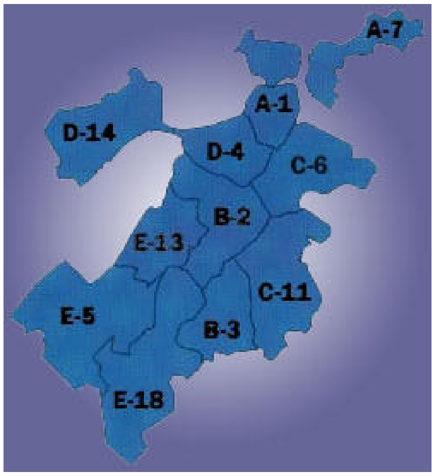


Figure 3: City of Boston, Police Districts



A-1 Downtown/Beacon Hill/Chinatown/Charlestown
A-7 East Boston
B-2 Roxbury/Mission Hill
B-3 Mattapan/North Dorchester
C-6 South Boston
C-11 Dorchester
D-4 Back Bay/Sound End/Fenway
D-14 Allston/Brighton
E-5 West Roxbury/Roslindale
E-13 Jamaica Plain
E-18 Hyde Park

Table 1: Probability of Search by Officer Race and Driver Race

(Standard Deviation of Sample Mean in Parentheses)

	Officer Race			
Driver Race	White	Black	Hispanic	All
White	0.40%	0.62%	0.25%	0.46%
	(0.04%)	(0.07%)	(0.09%)	(0.04%)
	n=22471	n=11132	n=3256	n=36859
Black	0.97%	0.82%	0.49%	0.87%
	(0.09%)	(0.09%)	(0.15%)	(0.06%)
	n=13131	n=9116	n=2258	n=24505
Hispanic	0.97%	0.82%	0.38%	0.85%
	(0.14%)	(0.16%)	(0.19%)	(0.10%)
	n=5058	n=3164	n=1066	n=9288
All	0.65%	0.73%	0.35%	0.65%
	(0.04%)	(0.06%)	(0.07%)	(0.03%)
	n=40660	n=23412	n=6580	n= 70652

Note: Includes only officers for whom the search variable is missing for at most 10% of all citations. Stops involving drivers from other racial groups are not included.

Table 2: Probability of Search by Officer Race and Driver Race Weighted by Inverse of Number of Citations

(Standard Deviation of Sample Mean in Parentheses)

	Officer Race			
Driver Race	White	Black	Hispanic	All
White	1.18%	2.59%	2.31%	1.99%
	(0.47%)	(0.66%)	(1.91%)	(0.39%)
	n=404	n=138	n=46	n=588
Black	5.85%	1.95%	0.49%	4.46%
	(1.18%)	(0.69%)	(0.21%)	(0.82%)
	n=361	n=135	n=42	n=538
Hispanic	4.05%	4.74%	0.30%	3.86%
	(1.41%)	(2.42%)	(0.17%)	(1.11%)
	n=265	n=110	n=37	n=412
All	3.45%	2.64%	1.37%	3.10%
	(0.52%)	(0.53%)	(0.96%)	(0.38%)
	n=470	n=163	n=52	n=685

Note: Includes only officers for whom the search variable is missing for at most 10% of all citations. Stops involving drivers from other racial groups are not included. For each officer, observations weighted by one over the number of citations given by that officer.

Table 3: Summary Statistics

(Standard Deviation in Parentheses)

Variable	Baseline Search Missing	Primary Sample			
	All Officers	All Officers	White Officers	Black Officers	Hispanic Officers
White Driver	57.3%	52.2%	55.3%	47.5%	49.5%
	(49.5%)	(50.0%)	(49.7%)	(49.9%)	(50.0%)
Black Driver	30.8%	34.7%	32.3%	38.9%	34.3%
	(46.2%)	(47.6%)	(46.8%)	(48.8%)	(47.5%)
Hispanic Driver	11.9%	13.1%	12.4%	13.5%	16.2%
	(32.4%)	(33.8%)	(33.0%)	(34.2%)	(36.8%)
Mismatch	49.1%	53.8%	49.7%	61.1%	83.8%
	(50.0%)	(49.9%)	(49.7%)	(48.8%)	(36.8%)
Baseline Search	-	0.7%	0.7%	0.7%	0.3%
		(8.0%)	(8.1%)	(8.5%)	(5.9%)
Stop at Night	30.4%	30.4%	26.6%	36.9%	30.4%
	(46.0%)	(46.0%)	(44.2%)	(48.3%)	(46.0%)
Young Driver (Age<26)	24.7%	24.2%	24.2%	23.8%	26.2%
	(43.1%)	(42.8%)	(42.8%)	(42.6%)	(44.0%)
Male Driver	71.8%	68.1%	69.4%	65.5%	69.8%
	(45.0%)	(46.6%)	(46.1%)	(47.5%)	(45.9%)
In-State Driver	93.8%	93.6%	93.3%	94.1%	93.4%
	(24.1%)	(24.6%)	(25.1%)	(23.6%)	(24.7%)
In-Town Driver	48.7%	51.1%	49.0%	54.3%	52.4%
	(50.0%)	(50.0%)	(50.0%)	(49.8%)	(49.9%)
Accident	2.5%	1.3%	1.5%	0.9%	1.3%
	(15.7%)	(11.3%)	(12.0%)	(9.7%)	(11.5%)
Allston-Brighton	7.5%	6.7%	8.2%	4.6%	4.6%
	(26.4%)	(25.0%)	(27.5%)	(21.0%)	(21.0%)
Boston Central	20.4%	13.1%	12.9%	12.1%	18.0%
	(40.3%)	(33.7%)	(33.5%)	(32.6%)	(38.4%)
Charlestown-East Boston	10.4%	6.3%	8.5%	3.2%	3.2%
	(30.5%)	(24.2%)	(27.9%)	(17.5%)	(17.7%)
Dorchester-Mattapan	21.8%	20.1%	20.6%	20.2%	16.5%
	(41.3%)	(40.0%)	(40.4%)	(40.1%)	(37.1%)
Hyde Park	0.7%	0.9%	0.7%	1.3%	0.4%
	(8.6%)	(9.3%)	(8.3%)	(11.3%)	(6.4%)
Jamaica Plain	2.4%	2.5%	3.2%	0.4%	6.1%
	(15.3%)	(15.7%)	(17.5%)	(6.2%)	(23.9%)
Roslindale	0.5%	1.1%	1.3%	1.0%	0.7%
	(7.4%)	(10.6%)	(11.2%)	(10.1%)	(8.2%)
Roxbury	13.0%	17.6%	18.6%	15.1%	20.0%
	(33.6%)	(38.1%)	(38.9%)	(35.8%)	(40.0%)
South Boston	6.0%	4.0%	4.7%	3.5%	1.8%
	(23.7%)	(19.7%)	(21.1%)	(18.4%)	(13.2%)
Number of Officers	922	685	470	163	52
Number of Citations	25,203	70,652	40,660	23,412	6,580

Table 4: Probability of Search and Guilt Conditional on Search Officer Race Excluded

	Unweighted Probits		Weighted Probits	
	Search	Guilt	Search	Guilt
Black Driver	0.213***	-0.472	0.387***	-0.622
	(0.059)	(0.388)	(0.144)	(0.464)
Hispanic Driver	0.144	-0.228	0.219	0.262
	(0.108)	(0.409)	(0.163)	(0.452)
Stop at Night	0.154	0.012	0.201*	-0.487
	(0.101)	(0.329)	(0.116)	(0.349)
Young Driver (Age<26)	0.087**	-0.314	0.110	-0.413
	(0.038)	(0.236)	(0.129)	(0.367)
Male Driver	0.064	-0.188	0.096	-0.062
	(0.046)	(0.261)	(0.123)	(0.365)
In-State Driver	0.084		0.246	
	(0.092)		(0.194)	
In-Town Driver	0.028	-0.030	0.032	0.045
	(0.036)	(0.335)	(0.105)	(0.402)
Accident	0.854***	-0.138	0.022	0.481
	(0.153)	(0.433)	(0.188)	(0.531)
Neighborhood Controls	YES	YES	YES	YES
Observations	70,652	369	70,652	369

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Probability of Search, Baseline Specification

	Unweighted Probits			**Weighted Probits**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Black Driver	0.207***	0.175***	0.190***	0.167	0.144	0.204
	(0.059)	(0.059)	(0.057)	(0.126)	(0.126)	(0.142)
Hispanic Driver	0.173	0.120	0.083	0.061	0.023	-0.006
	(0.115)	(0.118)	(0.106)	(0.166)	(0.174)	(0.176)
Black Officer	0.027	0.044	0.058	-0.134	-0.115	-0.085
	(0.185)	(0.188)	(0.165)	(0.134)	(0.134)	(0.135)
Hispanic Officer	-0.251	-0.260	-0.225	-0.487*	-0.511*	-0.501**
	(0.182)	(0.180)	(0.176)	(0.279)	(0.269)	(0.249)
Mismatch	0.099*	0.110*	0.124**	0.354***	0.355***	0.345***
	(0.057)	(0.057)	(0.059)	(0.126)	(0.125)	(0.121)
Stop at Night		0.144	0.156		0.207*	0.208*
		(0.104)	(0.107)		(0.123)	(0.117)
Young Driver (Age<26)		0.087**	0.092**		0.101	0.106
		(0.038)	(0.038)		(0.128)	(0.126)
Male Driver		0.077	0.069		0.100	0.088
		(0.054)	(0.044)		(0.128)	(0.122)
In-State Driver		0.116	0.084		0.255	0.254
		(0.096)	(0.093)		(0.182)	(0.185)
In-Town Driver		0.013	0.029		-0.015	0.025
		(0.046)	(0.036)		(0.099)	(0.105)
Accident		0.867***	0.863***		0.036	0.018
		(0.144)	(0.151)		(0.179)	(0.188)
Neighborhood Controls	NO	NO	YES	NO	NO	YES
Observations	70,652	70,652	70,652	70,652	70,652	70,652

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Probability of Search, Robustness Checks

	(1)	(2)	(3)
Black Driver	0.284**	0.192*	0.198*
	(0.111)	(0.106)	(0.104)
Hispanic Driver	0.068	0.035	0.021
	(0.128)	(0.130)	(0.129)
Black Officer	0.022	-0.073	-0.052
	(0.119)	(0.106)	(0.106)
Hispanic Officer	-0.111	-0.289*	-0.128
	(0.161)	(0.157)	(0.156)
Mismatch	0.109	0.259***	0.189**
	(0.106)	(0.091)	(0.091)
Stop at Night	0.239***	0.195**	0.217**
	(0.090)	(0.089)	(0.084)
Young Driver (Age<26)	0.183**	0.162*	0.242***
	(0.092)	(0.091)	(0.088)
Male Driver	0.165	0.128	0.115
	(0.102)	(0.101)	(0.098)
In-State Driver	0.099	0.156	0.075
	(0.172)	(0.140)	(0.168)
In-Town Driver	0.110	0.033	0.118
	(0.093)	(0.084)	(0.085)
Accident	0.079	-0.028	-0.045
	(0.162)	(0.133)	(0.141)
Neighborhood Controls	YES	YES	YES
Observations	79,337	95,855	79,369

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Note: In column (1), we include officers for whom the search variable is missing in more than 10 percent of the citations they issue and drop all citations in which search is missing. In column (2), we assume that if search was missing, then no search was conducted. Column (3) is identical to column (1), except that if all of an officer's non-missing search citations indicate that a search was conducted, then we assume that no search was conducted for all of the missing observations.

Table 7: Probability of Search, Unexperienced vs. Experienced Officers
Weighted Probits

	Inexperienced Officers (<10 years)			Experienced Officers (>10 years)		
	(1)	(2)	(3)	(4)	(5)	(6)
Black Driver	-0.127	-0.087	-0.028	0.416**	0.348*	0.468**
	(0.130)	(0.138)	(0.151)	(0.201)	(0.201)	(0.219)
Hispanic Driver	0.090	0.100	0.106	0.073	-0.057	-0.108
	(0.172)	(0.168)	(0.184)	(0.281)	(0.286)	(0.279)
Black Officer	-0.043	-0.013	-0.062	-0.172	-0.172	-0.085
	(0.174)	(0.170)	(0.157)	(0.190)	(0.189)	(0.196)
Hispanic Officer	-0.755***	-0.780***	-0.758***	-0.109	-0.153	-0.152
	(0.170)	(0.170)	(0.180)	(0.430)	(0.425)	(0.392)
Mismatch	0.165	0.210	0.200	0.494**	0.486**	0.474**
	(0.134)	(0.131)	(0.127)	(0.197)	(0.192)	(0.184)
Stop at Night		0.463***	0.460***		0.023	0.017
		(0.171)	(0.157)		(0.161)	(0.162)
Young Driver (Age<26)		-0.220*	-0.186		0.354**	0.347**
		(0.118)	(0.115)		(0.176)	(0.176)
Male Driver		0.006	0.012		0.191	0.181
		(0.119)	(0.111)		(0.194)	(0.185)
In-State Driver		0.545**	0.560**		-0.079	-0.082
		(0.219)	(0.229)		(0.257)	(0.259)
In-Town Driver		-0.261**	-0.277**		0.220	0.271*
		(0.113)	(0.125)		(0.153)	(0.160)
Accident		0.163	0.133		0.037	0.058
		(0.168)	(0.163)		(0.276)	(0.291)
Neighborhood Controls	NO	NO	YES	NO	NO	YES
Observations	33,249	33,249	33,249	37,403	37,403	37,094

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Note: There are fewer observations in the last column because there is one neighborhood, Jamaica Plain, in which no search was conducted by an officer with more than 10 years of experience.

Table 8: Racial Composition of Police Districts

	Census Benchmark Population 18 and Older				Citation-Level Data Racial Breakdown of Officers by District			
District	White	Black	Hispanic	Asian	White	Black	Hispanic	Asian
A-1	76.7%	3.3%	3.2%	15.1%	62.8%	24.8%	8.0%	4.4%
A-7	53.0%	2.4%	36.6%	3.7%	72.1%	16.4%	9.8%	1.6%
B-2	22.1%	47.8%	17.0%	4.7%	51.7%	35.7%	10.9%	1.7%
B-3	3.8%	78.9%	10.8%	1.1%	55.2%	37.3%	6.6%	0.9%
C-6	87.5%	1.8%	5.2%	4.0%	76.5%	14.8%	7.4%	1.3%
C-11	41.4%	28.7%	9.0%	12.5%	70.4%	17.3%	8.7%	3.6%
D-4	66.7%	9.9%	8.8%	11.5%	69.8%	21.2%	7.3%	1.7%
D-14	71.3%	3.9%	8.0%	13.1%	71.1%	16.3%	9.6%	3.0%
E-5	80.7%	6.2%	7.8%	3.0%	71.1%	22.2%	5.9%	0.7%
E-13	54.3%	15.2%	25.0%	2.3%	60.5%	21.6%	17.2%	0.8%
E-18	47.9%	33.7%	11.9%	3.0%	59.1%	30.7%	10.2%	0.0%

Source: The Census numbers are taken from the "Massachusetts Racial and Gender Profiling Project: Preliminary Tabulations," prepared by Northeastern University, Institute on Race and Justice. The citation-level numbers were derived from our data.

Table 9: Probability of Search, Drivers Stopped in Their Own Neighborhood Weighted Probits

	(1)	(2)	(3)
Black Driver	-0.017	-0.050	0.242
	(0.190)	(0.191)	(0.259)
Hispanic Driver	0.251	0.198	0.259
	(0.242)	(0.244)	(0.256)
Black Officer	-0.219	-0.210	-0.146
	(0.159)	(0.159)	(0.154)
Hispanic Officer	-1.161***	-1.136***	-1.119***
	(0.292)	(0.296)	(0.318)
Mismatch	0.499***	0.499***	0.491***
	(0.189)	(0.184)	(0.173)
Stop at Night		-0.065	-0.039
		(0.161)	(0.156)
Young Driver (Age<26)		0.244	0.275*
		(0.174)	(0.165)
Male Driver		-0.014	-0.046
		(0.179)	(0.170)
Accident		0.290	0.255
		(0.280)	(0.279)
Neighborhood Controls	NO	NO	YES
Observations	11,234	11,234	10,944

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Note: There are fewer observations in the last column because there is one neighborhood, Jamaica Plain, in which residents of that neighborhood were never searched.

Table 10: Probability of Search, Diverse vs. Homogeneous Neighborhoods Weighted Probits

	Diverse Neighborhoods			Homogeneous Neighborhoods		
	(1)	(2)	(3)	(4)	(5)	(6)
Black Driver	-0.364*	-0.399*	-0.313	0.445***	0.450***	0.450***
	(0.206)	(0.226)	(0.220)	(0.159)	(0.161)	(0.160)
Hispanic Driver	-0.147	-0.157	-0.172	-0.083	-0.055	-0.042
	(0.226)	(0.234)	(0.232)	(0.232)	(0.241)	(0.239)
Black Officer	-0.382*	-0.356	-0.327	0.101	0.130	0.142
	(0.226)	(0.219)	(0.211)	(0.167)	(0.168)	(0.166)
Hispanic Officer	-0.360	-0.355	-0.313	-0.918***	-0.908***	-0.873***
	(0.293)	(0.274)	(0.256)	(0.245)	(0.246)	(0.242)
Mismatch	0.386*	0.360*	0.355*	0.233	0.261	0.249
	(0.203)	(0.196)	(0.192)	(0.161)	(0.167)	(0.163)
Stop at Night		0.415**	0.469***		0.123	0.098
		(0.168)	(0.158)		(0.176)	(0.169)
Young Driver (Age<26)		-0.155	-0.152		0.094	0.112
		(0.169)	(0.174)		(0.186)	(0.181)
Male Driver		0.123	0.139		-0.209	-0.223
		(0.184)	(0.171)		(0.162)	(0.157)
In-State Driver		0.272	0.221		0.410*	0.428*
		(0.290)	(0.289)		(0.241)	(0.251)
In-Town Driver		-0.017	-0.001		-0.101	-0.076
		(0.193)	(0.191)		(0.145)	(0.147)
Accident		0.035	0.002		0.443	0.434
		(0.223)	(0.253)		(0.300)	(0.298)
Neighborhood Controls	NO	NO	YES	NO	NO	YES
Observations	19,247	19,247	19,247	37,234	37,234	37,234

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Note: Neighborhoods are categorized as either diverse or homogeneous based on the numbers in the left-hand side of Table 8, which give the racial composition of residents in different police agencies. The mapping between agencies and neighborhoods was based on the name of the agency and the name of the neighborhood. The neighborhoods categorized as diverse are East Boston (including Charlestown), Roxbury/Mission Hill, Jamaica Plain and Hyde Park. The neighborhoods categorized as homogeneous are Boston Central, South Boston, Allston/Brighton and West Roxbury/Roslindale.