ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Racial Profiling or Racist Policing?: Bounds Tests in Aggregate Data

Authors	Rubén Hernández-Murillo, and John Knowles
Working Paper Number	2004-012A
Citable Link	https://doi.org/10.20955/wp.2004.012
Suggested Citation	Hernandez-Murillo, R., Knowles, J.; Racial Profiling or Racist Policing?: Bounds Tests in Aggregate Data, Federal Reserve Bank of St. Louis Working Paper 2004-012. URL https://doi.org/10.20955/wp.2004.012

Published In	International Economic Review
Publisher Link	https://doi.org/10.1111/j.0020-6598.2004.00293.x

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

The views expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment.

Racial Profiling or Racist Policing?: Bounds Tests in Aggregate Data.

Rubén Hernández-Murillo[1] John Knowles[1]
Federal Reserve Bank of St. Louis[2] University of Pennsylvania[3]

April 2004

Abstract

State-wide reports on police traffic stops and searches summarize very large populations, making them potentially powerful tools for identifying racial bias, particularly when statistics on search outcomes are included. But when the reported statistics conflate searches involving different levels of police discretion, standard tests for racial bias are not applicable. This paper develops a model of police search decisions that allows for non-discretionary searches and derives tests for racial bias in data that mixes different search types. Our tests reject unbiased policing as an explanation of the disparate impact of motor-vehicle searches on minorities in Missouri.

Subject Area: Racial Profiling; Discrimination; Bounds Tests; Public Policy.

JEL Keywords: K42, D10, C14.

[1]We thank Kate Antonovics, Scott Decker, Jeff Dominitz, Gregory Kordas, Nicola Persico, and Petra Todd for helpful discussions. The views expressed are those of the authors and do not reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

[2]E-mail: `Ruben.Hernandez@stls.frb.org`

[3]E-mail: `jknowles@econ.sas.upenn.edu`

1 Introduction

A recent wave of public concern about the disparate impact of police behavior on racial and ethnic minorities has led to a significant increase in the publication of statistical reports regarding the impact by race of police actions in the United States. For instance, according to the U.S. Department of Justice (2001), 16 state police agencies required collection of race and ethnicity data for all traffic stops as of March 2001, compared to only seven states in 1999. An additional 23 state agencies required collection of such data when arrests or violence ensued. The objective of publishing these reports is to allow public monitoring of racial or ethnic inequities in the burden of police activities, but it is not clear how or even whether such aggregate data can be useful in measuring discrimination against minorities by the police.

In regards to traffic stops and vehicle searches, these data typically confirm the hypothesis of disparate impact by race: motorists of some minority groups, such as African-Americans and Hispanics, are much more likely than other motorists to be subjected to police searches of their vehicles. For instance, Gross and Barnes (2002) find that on a given stretch of I-95 in Maryland over a 5-year period, black drivers were twice as likely as white motorists to be stopped and five times as likely to be searched. Many commentators agree that this situation is unfair to members of racial minorities, but tend to assume that there is a trade-off between fairness and efficient policing. Indeed, this is implicit in the term *racial profiling*, which refers to the standard police practice of basing search decisions on the fit between the motorist and descriptions (profiles) of typical law-breakers.[4]

The racial profiling story is thus an example of *statistical discrimination* in which the police are racially unbiased, in the sense that they do not care about race directly, but only as an instrument to predict criminality. However, police might also be racially biased in the sense that they care about race directly; for instance they may derive utility from searching minority motorists, in which case statistical discrimination will lead them to search minorities at a higher rate than if they were unbiased. If this is the case, then the public controversy over the apparent dilemma between effective policing and racial profiling may actually harm minorities by camouflaging instances of racial bias in the sense defined above. Disparate impact caused by racist policing is probably not the whole story behind the racial profiling controversy, but if such a bias explains even a part of the problem, then disparate impact can be reduced in a non-controversial way by testing for and eliminating bias that does not fit the racial profiling model. In this perspective, the essential question is: how can the aggregate profiling statistics published at the state level be used to distinguish racial profiling from racial bias?

In this paper, we examine ways to apply a well-known test, the comparison of average search-success rates by race, to summary statistics by race where different kinds of search are aggregated together. The key properties that we require to analyze a report is that it include, for each police force, reports of search-success rates by race, as well as the proportion of searches that were

[4] The ACLU, on its racial-profiling web site <http://www.aclu.org/profiling>, gives the following definition of racial profiling: "this practice of substituting skin color for evidence as a grounds for suspicion by law enforcement officials". The Department of Justice, in its Fact Sheet of March 17th, 2003, says "Racial profiling rests on the erroneous assumption that any particular individual of one race or ethnicity is more likely to engage in misconduct than any particular individual of other races or ethnicities".

discretionary (as opposed to mandatory, or non-discretionary) because we presume that police agents have the ability to exercise racial bias only when searches are not compulsory. To the best of our knowledge Missouri is the only state with a public-access traffic-stop report, published by the Missouri Attorney General's Office (henceforth AGO), that fulfills these criteria. A recent survey of state traffic-stop data sets by the U.S. Department of Justice (2001) does not even mention types of search or search outcomes in their report on data collection policies. The results of our paper suggest that this is a serious oversight that limits the ability to monitor the racial bias of police forces.[5]

The tests that we develop here are based on an equilibrium argument by Knowles, Persico and Todd (2001) (henceforth KPT). The key assumptions are: 1) police choose search rates by race to maximize overall find rates, 2) the race of motorists is related to their probability of carrying contraband at a given search probability, and 3) motorists respond to higher risks of search by reducing their probability of carrying. In the equilibrium with unbiased policing, the success rate for discretionary searches will be equal across all observable groups that are searched with interior probabilities. If however the police are biased against a given group of motorists, this will be revealed by a lower success rate in such searches. KPT found that success rates for the Maryland State Police did not vary significantly by race, supporting racial profiling, in the sense of unbiased policing, as an explanation for higher search rates of black motorists.

In contrast to the Maryland data, search success rates in the Missouri data are often much lower for blacks and Hispanics than for white motorists. However the 2001 report stated that the reason lower success rates for minorities did not imply biased policing was that minorities are more often arrested in traffic stops, and police are required to search the motorist and vehicle as part of the arrest procedure.[6] Since the success rates that we observe aggregate over both types of search of a given race by a given police force, it is possible for unbiased policing to result in lower overall success rates for minorities. Therefore the KPT test is not applicable when find rates on non-discretionary searches are lower than on discretionary searches.

We extend the KPT model to incorporate non-discretionary searches, and propose a way to test this explanation of racial disparities. We use basic properties of probabilities to infer non-parametric bounds on success rates for discretionary searches, and derive confidence intervals for these bounds. We then explore two ways to sharpen these bounds: using additional information, and making auxiliary assumptions on the distribution of the find rates of non-discretionary searches. As additional information, we impose the restriction implied by the AGO statement, that non-discretionary searches have lower success rates. As auxiliary assumptions, we impose parameterized restrictions on the disparity between the success rates of the two types of searches. We compare our results with these tests to the results under the KPT test, which is based on the assumption that all searches are discretionary.

Our main empirical results are (statistics below refer to tests with 95-percent confidence inter-

[5] During the final revision of this paper, we learned that a Minnesota state report for 2002 does in fact contain statistics on success rates by each type of search.

[6] The exact quote is "The contraband hit rate for whites was 22 percent, compared with 15 percent for Blacks and 11 percent for Hispanics. This means that, on average, searches of Blacks and Hispanics produce less contraband than do those of whites. This difference is most likely attributable to the higher arrest rates for Blacks and Hispanics, circumstances that compel a search."

vals):

1. Under the KPT test, we reject equality of search rates for police forces representing at least 46 percent of all discretionary searches.

2. Police forces that are found to be biased under our strictly non-parametric bounds account for at most 2 percent of all discretionary searches. Despite the apparently large differences in success rates by race, the proportions of non-discretionary searches are too large to rule out equality by race of find rates for discretionary searches.

3. Our test based on the AGO restriction finds that police forces biased against minorities account for at least 30 percent of all discretionary searches; while this is much more than under the basic non-parametric bounds, it is significantly less than under the KPT test. Hence we find that a large fraction of searches in Missouri are carried out by police forces for which we do not reject unbiased policing.

4. Most of the police bias identified by the AGO test is directed against Hispanics rather than blacks: the share of searches by police biased against blacks is only 7 percent, compared to 40 percent for police biased against Hispanics.

5. Black motorists are most at risk from biased police when travelling through areas where blacks have a relatively small share of population. Similarly, Hispanics are most at risk in neighborhoods with a small minority share of the population.

6. Standard measures of disparity between racial composition of the resident population and the population searched or stopped by police are not useful predictors of racial bias according to our tests.[7]

Our non-parametric bounds are based on Horowitz and Manski (1995). Previous work by Kreider and Pepper (2003) and Dominitz and Sherman (2003) shows the usefulness of additional assumptions to sharpen bounds. Our use of the AGO statement most closely resembles the "ordered outcomes" assumption described in Manski (1995), and used by Pepper (2000). The confidence intervals we use to implement our test are derived from Horowitz and Manski (2000) and Imbens and Manski (2003). Our problem is somewhat simpler than those described above, however, as we do not have to make inferences about the unconditional distribution of find rates, which in our case is identified by the data, but only about the find rate conditional on discretionary search.

The rest of the paper is organized as follows. In the next section we extend the model of Knowles, Persico and Todd (2001) to the case where some searches are non-discretionary and derive formally the various tests described above. Section 3 describes the data for Missouri. In Section 4 we describe our main results from applying the bounds tests to these data. In Section 5 we present a statistical analysis of the differences between the tests, and in Section 6 we characterize the police forces we find to be biased. We conclude with a summary and some caveats.

[7]This type of index is intended to reflect a more traditional measure of disparate impact, based on empirical studies of driver population, which is often used in courts as an indicator of racial bias, for instance in the seminal case *New Jersey v. Soto,* 734 A.2d, Superior Court of New Jersey, 1996, and Zingraff et al. (2000).

2 The Racial Profiling Model

We suppose that there is a population of motorists who decide whether or not to carry contraband. There are three types of motorists; non-felons ($z = 0$), type-1 felons ($z = 1$) and type-2 felons ($z = 2$). Motorists also differ according to race, $r \in \{b, w\}$ and other characteristics, c. They are randomly stopped by the police with an exogenous probability π_s that is independent of motorist characteristics. When the police agent makes the stop, she learns whether the motorist is a felon; if so, she arrests the motorist and searches his vehicle, otherwise the police agent observes the motorist's characteristics and decides whether or not to search his vehicle for contraband. This gives rise to two types of search; in the first case, these are non-discretionary searches ($d = 0$), and the second these are discretionary $(d = 1)$. All searches of motorists who carry contraband are assumed to be successful, in which case the motorist is subjected to a punishment that is independent of his characteristics, but might depend on his type. We also assume that all searches reported to be successful were of motorists who were in fact carrying contraband.

We assume that police agents get utility from finding contraband. In order to reflect the possibility of racial bias, we allow the cost of searching motorists, t_r, to depend on the race of the motorist. We also assume that police know the function $p(g|c, r, z)$ that gives the probability of a motorist carrying contraband. Let the density function $\mu(c, r, z)$ represent the distribution of c across motorists who are not felons. Let the payoff to police from finding contraband be normalized to 1. Then we can write the police problem as choosing a search probability function $\gamma(c, r)$ to solve:

$$\max_{\gamma(c,r)} \left\{ \int \left[p(g|c, r, d = 1) - t_r\right] \gamma(c, r)\, \mu(c, r, z = 0)\, dc \right\}. \tag{1}$$

The solution implies that $\gamma(c, r) \in (0, 1)$ only if the expected payoff from a search is zero.

We assume that whether the motorist is a felon does not affect the utility $\nu(c, r)$ from carrying contraband, but may affect the disutility $j(c, r, z)$ from being subjected to a successful search. There are no costs of being subjected to unsuccessful search. If motorists carry contraband with some positive probability $p(g|c, r, z) \in (0, 1)$, then they must be indifferent between carrying and not carrying. For non-felon motorists, this implies the unconditional search probability satisfies:

$$\gamma(c, r) = \frac{\nu(c, r)}{\pi_s\left[\nu(c, r) + j(c, r, 0)\right]}. \tag{2}$$

Felon motorists on the other hand must take into account that they face a higher probability of being searched than non-felon motorists. This is because a greater number of motorists are stopped than are eventually searched, but all felons who are stopped are searched. Since non-felons are indifferent at a lower search probability, the carrying probability for felon motorists would be zero if the payoffs were identical; hence to model positive carrying rates for felons, we allow the utility gain from carrying contraband to be higher for some felons. We assume that for type-1 felons, arrest does not affect the net utility gain from carrying contraband, but for type 2 felons the utility loss of being found with contraband is smaller. This may arise for instance if minor offenses like carrying contraband tend to be ignored when sentencing such felons.

Taken together, conditions (1) and (2) imply an equilibrium in which the probability of carrying equals the cost per search and police impose higher search rates on motorists with higher utility of carrying contraband (relative to the punishment utility). The main result of the KPT model is that search rates may differ by the race or other characteristics of the motorist, but if the cost of search is independent of motorist characteristics, and if carrying probabilities are interior, then success rates of discretionary searches are equalized across all groups of motorists that are distinguishable to the police. The economic force that drives this result is that if some observable group of motorists were to carry contraband at a higher rate than the rest, then the police would allocate more resources to searching that group, and the observed find rate would be higher for that group. This would lead motorists of that group to reduce their probability of carrying contraband.

In addition to the implications for discretionary searches, the current model implies carrying rates for felons that depend on the size of the punishment $j(c, r, 0)$ relative to $j(c, r, 2)$, and the distribution of felons across felony types. Let $\pi^{ri} = \int \mu(c, r, i)\, dc$ be the proportion of motorists of race r who are type-i felons. If $j(c, r, 2) < j(c, r, 0)$ then for each groups (c, r), if the non-felons carry with positive probability, then type-1 felons do not carry, while type-2 felons carry contraband for sure. The find rate on non-discretionary search is thus given by the relative frequency of type-2 felons:

$$p(g|r, d = 0) = \frac{\int \mu(c, r, 2)\, dc}{\int [\mu(c, r, 1) + \mu(c, r, 2)]\, dc} = \frac{\pi^{r2}}{\pi^{r1} + \pi^{r2}}.$$

A particular racial group r may therefore be subjected to discretionary search at a higher rate than other groups for two possible reasons. First, non-felon motorists of race r may experience a higher utility gain $\nu(c, r) - j(c, r, 0)$ from carrying contraband, and second, the police agent may have a lower cost of search t_r for that race. The first explanation corresponds to *statistical discrimination*, and the second to *racial bias*. It is clear that in a legal context, such as the sources cited in the introduction, the term 'racial profiling' is intended to refer to the first type of discrimination, although public usage may confound the two.

KPT discuss two related issues that may arise with the interpretation of the model. First, the above results for discretionary search apply only to groups of motorists who, in equilibrium, carry contraband and are searched with interior probabilities. KPT show that it is trivial to extend the model to allow for types that are searched with probabilities zero or one.[8] Second, the reader may object, on the grounds of implausibility, to the result that motorists choose a mixed strategy. However KPT show there is an alternative interpretation of the model, in which the equilibrium properties are identical, such that motorists do not choose mixed strategies, but instead each group of motorists distinguishable to the police contains types who carry at the equilibrium search rate and types who do not.

[8] In our model, the *ex ante* search probability is bounded above by the stop probability, π_s, while in KPT the upper bound is one. Thus the possibility of motorists who are undeterred even by the maximum search rate may seem to be more relevant in our model. As Becker (1968) points out, however, when punishments are sufficiently severe, even a low probability of detection will have a deterrence effect.

2.1 Testing for unbiased policing

In this section, we develop non-parametric bounds on the success rates of discretionary searches, and then test whether confidence intervals around these bounds overlap. We sharpen these bounds in two ways: by using additional information, and by making auxiliary assumptions on the distribution of the find rates of non-discretionary searches.

A key implication of the equilibrium is that, for motorists subject to discretionary searches with some interior probability, the average search success rate $p(g|c, r, z = 0)$ depends only on the search cost. Thus in the absence of racial bias, the success rate will not depend on the observable characteristics of the motorists searched. On the other hand, if racial bias leads to higher search rates for a given class (c, r), then it must be the case that the success rate of searching that class will be lower, as motorists respond by reducing their carrying probability.

2.1.1 The KPT Test: Equality of Success Rates

The essence of the test by KPT is therefore to check whether the observed search rates differ by race of the motorist or by other characteristics. It is important to note that this test does not require identifying the marginal motorist of each race, as is the case with standard applications of outcomes-based testing. This makes the test particularly well suited to analyzing summary statistics, as in the current paper, where motorist characteristics are not available at an individual level. When individual data are available, then the KPT test applies directly; it suffices to compute the mean success rates of the discretionary searches on each race of motorist, and test whether the differences among these rates are statistically significant. In the absence of individual data, the same direct test can be applied, provided that the statistics were generated from discretionary searches only.

When the test is to be applied to summary statistics that combine both discretionary and non-discretionary data however, then this test is no longer valid. While in the original KPT model lower success rates indicate taste-based discrimination, in our model, different success rates of non-discretionary searches may also be an explanation of different observed success rates. Before concluding that differences in success rates indicate biased policing, it is therefore necessary to ask whether it is likely or even possible that the rate of non-discretionary search can explain these differences.

2.1.2 The HM Test: Non-parametric bounds

Since probabilities are bounded and must sum to one, it is theoretically possible to reject unbiased policing at the level of the individual police force by establishing upper and lower bounds for the success rates of searches. To clarify our approach, let P_d^r be the success rate for type-d searches. The idea, derived from Horowitz and Manski (1995), is to compute the success rates P_1^r for each race r for each police force under extreme assumptions about the success rates in non-discretionary searches, P_0^r. This defines the *identification region*, the range of possible success rates for each racial group. If the range of success rates of searches on blacks overlaps with that of white motorists, then no inference is possible. On the other hand, if there is a sufficiently high probability that the lower bound on the success rate for white motorists is higher than the upper bound on the success rate for black motorists, then we can reject unbiased policing.

According to the model, the success rate for discretionary searches in equilibrium is given by:

$$P_1^r = \frac{\int p\,(g|c,r,0)\,\gamma\,(c,r)\,\mu\,(c,r,0)\,dc}{\int \gamma\,(c,r)\,\mu\,(c,r,0)\,dc} = t_r.$$

The success rate of non-discretionary searches of felons is P_0^r. The model implies this is given by:

$$P_0^r = \frac{\pi^{r2}}{\pi^{r1} + \pi^{r2}}.$$

Suppose the success rate we observe, Q^r, is the mean over searches of both types. Then letting $\pi^r = \pi^{r1} + \pi^{r2}$, this is given by:

$$Q^r = (1 - \pi^r)\,P_1^r + \pi^r P_0^r. \tag{3}$$

Even if we observe π^r, the problem is that the data does not identify P_1^r : we would need to know both π^r and P_0^r. By solving equation (3) for P_1^r, given that $P_0^r \in [0,1]$, we can establish bounds on P_1^r:[9]

$$P_1^r \in \left[\frac{Q^r - \pi^r}{1 - \pi^r}, \frac{Q^r}{1 - \pi^r}\right]. \tag{4}$$

We derive confidence intervals that contain the parameter P_1^r with probability α. These extend beyond the bounds by a distance $D\,(\alpha)$, which we derive in the appendix. Our HM test rejects unbiased policing if the confidence interval on the search success rate for searches of white motorists lies entirely above that of black motorists.[10]

2.1.3 The AGO test: Sharpening the Bounds with Additional Information

These bounds can be made sharper using additional information or assumptions. For instance, the Missouri AGO report implies that non-discretionary searches have lower success rates.[11] In terms of our model, we take this to mean $P_0^r \leq P_1^r$. This is an example of an "ordered outcome" assumption, as described by Manski (1995); imposing this information on equation (3) serves to sharpen the lower bound on P_1^r:

$$P_1^r \in \left[Q^r, \frac{Q^r}{1 - \pi^r}\right]. \tag{5}$$

In this case, the bounds for P_1^r are tighter than the previous bounds for P_1^r in expression (4), as the lower bound corresponds to the probability of the observable search success Q^r, and the upper

[9] These bounds are equivalent to those derived in Horowitz and Manski (1995), Corollary 1.2.

[10] Since probabilities are bounded in $[0,1]$, for the purposes of our tests we effectively consider the interval $[\max\{0, \hat{P}_L - D\}, \min\{1, \hat{P}_H + D\}]$, where $\hat{P}_L$ and $\hat{P}_H$ are the sample lower and upper bounds.

[11] This view was confirmed by one of the consultants who prepared the report, professor Scott Decker of the University of Missouri.

bound corresponds to the upper bound in expression (4).[12]

To implement this test on finite samples, we derive confidence intervals for the identification region defined by the bounds; the derivation is shown in the appendix. We reject unbiased policing whenever the confidence interval around the identification region for the success rate on searches of white motorists lies above that of black motorists.[13] Thus the condition, $P_0^r \leq P_1^r$, which may raise the lower bounds, but has no effect on the upper bounds, will tend to result in a greater rate of rejection than the non-parametric bounds in (4). The opposite restriction, $P_0^r \geq P_1^r$, would lower the upper bounds, but have no effect on the lower bounds. This in general would also increase the rate of rejection. Hence, in general, the bounds would be tightened no matter what the sign of the restriction, so long as the same condition holds for both black and white motorists.[14]

3 The Data

We analyze statistics on traffic stops and searches from the 2001 edition of the *Annual Report on Missouri Traffic Stops,* published on the web by the Missouri Attorney General's Office.[15] According to the Missouri state law section 590.650, RSMo (2000), enacted August 28, 2000, every law enforcement agency in the state is required to compile a report containing summary information of all traffic stops made by the agency and to submit it to the office of the state Attorney General by March 1st of each year. The report for 2001 includes 609 law enforcement agencies out of a total of 668 in the state. Traffic stops in the report are made only for alleged violations of motor vehicle statutes or ordinances, and the reports do not include other stops, such as those of suspicious vehicles. The information reported in the forms includes a number of other variables, such as the category of moving violation that motivated the stop. The number of searches reported includes searches of drivers and of the property of the driver, but not those where only a passenger is searched. The report summarizes 1,389,947 traffic stops, of which 99,860 resulted in searches and 76,567 in arrests.

The report provides counts of stops and searches by racial category for each law enforcement agency in the state that reported data to the AGO. But the individual records for each traffic stop are not provided in the *Annual Report*, and were not available for the current paper. Thus the only tabulations possible are by race and by police agency, not by any other variables present in the report.

The AGO requires police officers to classify the driver's race or minority status on the basis of visual observation. The idea is that while this may not be altogether objective, the object of interest is the officer's impression of the driver's status; police are not required to ask motorists their race. It is not clear from the documentation how visual observation is meant to distinguish

[12] In equation (3) let P_0^r vary from 0 to P_1^r and solve for P_1^r.

[13] Note that while the confidence interval in the HM test contains the parameter P_1^r with probability α, the confidence interval in our other tests contains the identification region with probability α.Imbens and Manski (2003) show that the latter contains the former, so our tests tend to reject less often than would tests based on the first type of confidence interval. We use the second type of interval for reasons of tractability.

[14] Clearly the bounds can be tightened only if the AGO model is correct, and only the HM bounds provide a *true* lower bound because they are derived from the least restrictive model.

[15] URL: <http://www.ago.state.mo.us>.

Hispanics from other motorists.

The reports also list the numbers of searches motivated by the different types of probable cause used as pretext to conduct the search. These are Consent, Inventory, Drug/Alcohol odor, Incident to the Arrest, Plain-view contraband, Reasonable suspicion of weapon, Drug-dog alert, and Other. In particular, a search is classified as *Incident to the Arrest* (henceforth ITA) if the search follows an arrest made solely on information derived from the traffic stop.

Arrests made on information based solely on the traffic stop result in the police officer being forced to search the driver and the vehicle. In the 2000 report, the Attorney General, Jay Nixon, states: "A motorist with an outstanding arrest warrant will, and should, be arrested and will, and should, be searched pursuant to that arrest." The 2001 report states that when arrests are made on the basis of the stop only, searches are "almost always performed," and that arrests "compel a search." Thus ITA searches are non-discretionary; we assume that all other reported searches are discretionary.[16] In terms of our model, ITA searches correspond to police stops of felons of either type, while discretionary searches are of non-felons.

We have restricted our analysis to the three main types of law enforcement agencies: police departments, sheriff departments, and state troopers. The excluded law enforcement agencies include minor agencies such as park rangers, college police departments, and airport security. We have also restricted the study to the three largest racial or ethnic groups: white, African-American, and Hispanic drivers, excluding three smaller categories: Asian, American-Indian and Other.

Because our analysis focuses on discretionary searches, we have additionally excluded from the sample agencies where *all* searches of minority motorists were non-discretionary. Finally, for our tests we have selected those agencies that reported 5 or more searches for both white and African-American drivers and those that reported 5 or more searches for both white and Hispanic motorists. Despite these restrictions, our sample accounts for almost all searches in the report, as the police forces we exclude tend to conduct very few motorist searches. Our final sample of 191 law enforcement agencies represents about 83 percent of all stops, and 85 percent of all searches in the report.

[Table 1 about here.]

In Table 1 we present summary statistics of traffic stops and searches from the report, broken down by racial groups. The unit of observation in this table is the individual police force. From this table we learn that the total number of traffic stops in our sample is 1,158,922, of which 963,074 correspond to white drivers, 174,193 correspond to black drivers, and 21,655 to Hispanic drivers. The total number of searches is 84,652, of which 61,209 correspond to white drivers; 20,574 were searches of black drivers, and 2,869 were searches of Hispanic drivers. Of the total number of searches, 44,475 were non-discretionary, or ITA. On average, each police force accounted for 6,067 stops, of which 5,042 stops were of white drivers, 912 stops were of black drivers, and 113 stops were of Hispanic drivers. On average, each law enforcement agency accounted for 443 searches, leading

[16] This view was also confirmed by professor Scott Decker. He said: "ITA hit rates are lower in general because they are mandatory, required in all circumstances. Often this means that there is no other objective reason to search, such as visible contraband, occupants of the car displaying signs of intoxication (drug or alcohol), plain view, etc."

to 320 searches of white drivers, 107 searches of black motorists, and 15 searches of Hispanics. There is large variation across agencies in the number of stops and searches. The table indicates that about half of the agencies conducted less than 23 searches of black drivers and less than 5 searches of Hispanic drivers. In contrast, the median number of searches of white drivers was 131.[17]

The table includes the three measures of search success we use in our analysis: (1) searches where drugs were found, (2) searches where contraband in general was found, and (3) searches that resulted in an arrest. Hence 14,434 searches resulted in drug finds, and 17,515 resulted in general contraband finds, while 11,630 searches resulted in arrests.[18] This table additionally presents the counts of traffic stops that indicate other characteristics of the drivers, such as age and gender, but these were not used in the analysis.

The median values for search success among black motorists are equal to 3 and equal to 0 for Hispanic drivers. This is because many agencies reported no finds among minorities.

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

Figures 1-3 show the distribution of the fraction of non-discretionary, or ITA, searches by police force and race. In each case, the distribution only includes agencies that reported at least 5 searches for each racial group. There were 27 police forces that reported no ITA searches of whites, 6 agencies that reported no ITA searches of black drivers, and 6 agencies that reported no ITA searches of Hispanics. When the ITA frequency is zero, the observed searches for the category of motorist in question are discretionary. The charts indicate that most police forces, however, had positive ITA frequencies, and therefore it is important to take into account the effect of non-discretionary searches.

[Table 2 about here.]

In Table 2 we present another view of the numbers in Table 1, now expressed as fractions of total stops or searches. We learn from this table that in the aggregate, about 7 percent of stops result in searches, and that the probability of being searched given a traffic stop is much higher for Hispanics (about 0.13), and for blacks (about 0.11), compared to whites (about 0.06). About half of the searches are non-discretionary or ITA. Again, this rate is higher for Hispanics (about 0.53), and even higher for blacks (about 0.59), than for white drivers (about 0.50). Of our measures of search success, all three rates are much higher for white drivers. For Hispanic drivers, the probability that a search will result in drug finds, contraband finds, or arrests are, respectively, 0.08, 0.11, and

[17] The summary statistics of Tables 1 and 2 were computed using data from all these 191 agencies to compute correctly the sum of searches, but the comparison tests for black and Hispanic drivers were carried out separately with only those agencies which reported at least 5 searches of both white *and* minority drivers.

[18] The measure of contraband includes drugs, alcohol, currency, weapons, or stolen property. The measure of searches resulting in arrests was made available in the data provided to us by the AGO, but it is not included in the *Annual Report*. The measure of arrests included in the report does not distinguish arrests resulting from stops only or from the search outcome.

0.08. For black drivers, the probabilities are 0.12, 0.15, and 0.13. For white drivers, these numbers are 0.19, 0.22, and 0.14. The fraction of traffic stops with drivers age 29 or younger is about 50 percent for blacks and whites, but it is higher for Hispanics (about 58 percent). The fraction of traffic stops with male drivers is about 66 percent for blacks and whites, but it is also higher for Hispanics (about 83 percent).

In Table 2 we present a measure of racial disparity that indicates the incidence of traffic stops and searches among racial groups relative to the group's size in the population. The disparity index in terms of traffic stops is included for each law enforcement agency in the *Annual Report* and forms the basis for the analysis presented by the Missouri Attorney General of the problem of disparate impact and racial bias of law enforcement. This measure equals the share of traffic stops of a given group divided by the group's population share among the total population of driving age.[19] We also computed the disparity index in terms of searches. The interpretation of these disparity measures is as follows. If the disparity index exceeds 1 the group is over-represented relative to its share of the population. The table shows that whites and Hispanics are neither over-represented nor under-represented in terms of traffic stops, relative to their share of population, but blacks are over-represented and are being stopped at a rate 48 percent higher than their share of the population. In terms of the share of searches, the disparity index for blacks is 2.40 and for Hispanics 1.86, indicating that both minorities are strongly over-represented in the search statistics, relative to the population composition of the locations where the police forces are based.

4 Results

[Table 3 about here.]

Table 3 summarizes the results for the different types of bounds we constructed according to the analysis in Section 2. The numbers in the table give the percent of all discretionary, or non-ITA, searches carried out by police forces for which we can reject the null hypothesis of equal search-success rates among racial groups. The first column gives the percentage of all discretionary searches accounted for by these forces, while the next two columns give the percent of discretionary searches of blacks or Hispanics accounted for by these police forces.

We report these percentages for three motorist types: the rows labelled *Black* refer to police forces for which we rejected the null for black motorists, *Hispanic* if we rejected for Hispanics, and *Minority* for rejection of the null for either Hispanics or blacks. Each of the three sections of the table refers to a different measure of search success. The tests are labelled KPT for the direct comparison of observed search success rates (which do not distinguish between discretionary and non-discretionary searches), HM for the widest bounds, given in equation (4), and AGO for the bounds that are narrowed using the inequality $P_0^r \leq P_1^r$. For the bounds tests, we reject unbiased policing if and only if the upper limit of the confidence interval around the bounds on the minority rate does not exceed the lower limit of the confidence interval around the bounds of the white

[19] For our analysis, we obtained the population of individuals age 18 or older in each city or county corresponding to each police or sheriff department in the sample, and the state population for the State Police. Ideally, the population of motorists should be used, but the census information is more readily available.

rate. We report results for the 95-percent and 99 percent confidence intervals. Our discussion of table 3 below focuses on the column corresponding to the 95-percent confidence intervals under the heading *non-ITA Searches* and the rows corresponding to *Minority*,

According to the KPT test, police forces biased against minorities account for 67 percent of total searches, if the success measure is the drug find rate. In regards to general contraband finds, the police forces found to be biased under the KPT test account for 55 percent of discretionary searches. For searches resulting in arrest, biased police forces account for 46 percent of discretionary searches according to the KPT test.

We find that police forces that are found to be biased under the HM bounds account for a small share of total searches. The table shows that using the HM bounds, it is possible to reject unbiased policing against minorities for at most 2.3 percent of searches carried out, looking at the bottom panel of the table labelled *Guilt 3: Arrests from Searches*. This is because, despite the apparently large differences in success rates by race, the proportions of ITA searches are too large to prevent overlap of the bounds for the discretionary search success rates.

However the AGO test suggests a picture much closer to that of the KPT test: biased police forces account for a large share of discretionary searches. Using Drug Finds as the success measure, police forces biased against minorities account for 42 percent of total searches according to the AGO test, compared to 67 percent for the KPT test. The other definitions of success yield similar results. In regards to contraband finds, biased police forces account for 35 percent of discretionary searches under the AGO test, while for searches resulting in arrest, biased police forces account for 29 percent of discretionary searches. Thus while the information that ITA success rates are lower than for discretionary searches does reduce the rate of rejection of unbiased policing when compared with the KPT test, it raises the rejection rate when compared with the HM bounds.

Most of the police bias that we can detect is directed against Hispanics rather than blacks. According to the KPT test for contraband, police biased against blacks account for 35 percent of discretionary searches, whereas those biased against Hispanics account for 38 percent. According to the AGO test however, the share of searches by police biased against blacks is only 6 percent, compared to 32 percent for police biased against Hispanics. It is clear from the larger difference in the case of the AGO test that the rate of non-discretionary searches is a more important factor for black searches than for Hispanics.

According to the AGO test, police biased against blacks account for a much smaller share of searches of blacks than they do of total searches. The opposite is true for Hispanics, though the effect is less pronounced. According to the AGO test for drug finds, police biased against blacks account for about 1 percent of searches of blacks, compared to 7 percent of total discretionary searches. For the other success measures, the pattern is similar. Police biased against Hispanics, on the other hand, account for large shares of searches of Hispanics as well as large shares of total discretionary searches.

5 Discussion

The tests presented above are based on contrasting assumptions about the difference in find rates for ITA searches. The KPT test implicitly assumes that ITA and discretionary search-sucess rates are identical, while the AGO test assumes that the minority ITA find rate can be as low as zero, but no greater than the non-ITA find rate. Overall, the AGO assumption, that success rates for non-discretionary searches are no greater than those for discretionary search, results in a significant tightening of the bounds. Except in the case of arrests from searches, it is clear that far fewer police forces appear biased against blacks under the AGO test than under the KPT test. This is because the AGO test is more conservative than the KPT test, in that it attributes differences in find rates by race to disparities in the find rates of non-discretionary search. Thus the difference between the KPT and AGO rejection rates measures the extent to which separate tabulation of hit rates for discretionary and non-discretionary search would be helpful in identifying biased policing, over and above the information already contained in the statement that ITA find rates are lower.

If find rates on minority motorists are lower than for whites, and the rate of ITA search is non-zero in both cases, the AGO test will assume that the ITA hit rate for whites was equal to the average for whites, and that for minorities it was zero. The HM test on the other hand, would assume in this case that the ITA find rate for whites was one, further reducing the probability of rejection of the unbiased policing hypothesis. Note that to the extent that ITA hit rates for minorities are greater than zero or those for whites less than the discretionary success rate, our estimates under the AGO test of the share of searches by biased police will under-estimate the actual share.

Since the difference in rejection rates among the tests is quite large, it would be useful to know whether ITA find rates have to be implausibly lower than those for non-ITA searches for the AGO results to obtain. We do not have a measure of plausibility, but we can carry out the tests under the assumption that ITA find rates are bounded below by some proportion $\gamma \in [0, 1]$ of the find rates for non-ITA searches, and above by P_1^r, that is, $P_0^r \in [\gamma P_1^r, P_1^r]$. Note that $\gamma = 1$ corresponds to the KPT test, since in this case $P_0^r = P_1^r = Q^r$, while $\gamma = 0$ corresponds to the AGO test. We can ask, for instance, to what extent do the differences in rejection rates depend on the fact that the AGO test allows ITA hit rates to be as low as zero, rather than bounding them below by 10 percent or 25 percent of the discretionary find rate?

For this parameterization, the bounds are given by:

$$P_1^r \in \left[Q^r, \frac{Q^r}{(1 - \pi^r) + \pi^r \gamma} \right].$$

In this case, we derived confidence intervals as in the AGO test. See the appendix for details.

[Table 4 about here.]

[Table 5 about here.]

In Table 4 we report results for five values of γ. We find that the rejection rate for drug finds on blacks remains essentially at the same level as under the AGO test until γ is above 0.75, while for Hispanics, nearly half of the difference between the two tests is gone by the time γ reaches 0.5. In fact, if we look at the 95-percent confidence intervals for Hispanics, setting γ=0.25 increases the share of searches by rejected police forces by about 13 percentage points. With regards to black drivers, these results mean that even slightly lower success rates for ITA searches will result in much lower rejection rates, and hence it is important, as the AGO report claims, to take into account the role of ITA searches of black motorists before concluding in favor of racial bias. This does not appear to be the case in regards to Hispanics however; even success rates of zero on ITA search cannot explain away the large share of searches accounted for by police forces who fail the KPT test, and success rates on ITA search would have to be very low relative to non-ITA search in order for the share of searches not to increase significantly above the shares implied by our AGO test.

A similar question arises in explaining the gap between AGO and the HM results; how critical is it that ITA find rates be no greater than the discretionary find rate? We can ask, for instance, to what extent do the differences in rejection rates depend on the fact that ITA hit rates are bounded above by some number 10 percent or 25 percent greater than the discretionary find rate? We can carry out the tests under the assumption that ITA find rates are bounded above by some number between P_1^r and 1, that is $P_0^r \in \left[0, \overline{P}_0^r\right]$, where $\overline{P}_0^r = \lambda P_1^r + (1-\lambda)$, for $\lambda \in [0,1]$. Note that $\lambda = 1$ corresponds to the AGO test, while $\lambda = 0$ corresponds to the HM test. From this parameterization, the bounds are given by:

$$P_1^r \in \left[\frac{Q^r - \pi^r(1-\lambda)}{(1-\pi^r) + \pi^r\lambda}, \frac{Q^r}{(1-\pi^r)}\right].$$

In Table 5 we show the results from allowing $\lambda \in [0.5, 1.0]$. We find that for black motorists, the share of searches by biased police forces falls below 1 percent as λ falls to 0.5 when finds are measured as drugs or contraband, but that under the arrest rate measure, almost all the difference between the HM and the AGO results disappears once λ falls below 0.75. For Hispanics, the test gives results very close to the HM test for $\lambda < 0.5$ (not shown) for all three measures of search success. In general, as the minority row shows, the percent of searches by biased police is visibly greater than for the HM test for $\lambda \geq 0.75$, but that the greatest share of additional rejections occurs when λ is increased from 0.875 to 1. Hence what is critical for the AGO test is that the ITA find rate be not much greater than the discretionary find rate, not that the ITA find rate is actually lower.[20]

6 Other Predictors of Racial Bias

In general, it may be useful to know what sort of police forces tend to be biased. In particular, are these police forces the same ones that are identified as biased by more conventional tests of racial profiling? Are our results driven by a few large police forces? In this section we use census data for

[20]We also conducted the bounds test with the opposite assumption, that ITA find rates were uniformly higher than discretionary find rates. The results of this test were closer to the HM test. The reason that this test is not as powerful as the AGO test is that, as Table 2 indicates, many police forces reported zero success rates for minorities. In these cases it is far more useful to have information that bounds the white success rate above zero than to have restrictions on the black find rates, which are point-identified when the observed success rate is zero.

the counties in which these police are based in order to compute racial-composition benchmarks. We also report the results of our test for the three of the largest police forces in the state.

[Table 6 about here.]

In Table 6 we organize our results according to the traffic stop disparity index score of the police forces used in the *Annual Report*. This table shows searches ranked by the stop-disparity index and for each quintile of the stop-disparity distribution reports the fraction of non-ITA searches accounted for by law enforcement agencies that fail the AGO and KPT tests with 95-percent confidence intervals. The sum of the rows for each quintile under the AGO column add up to the corresponding row for the AGO test reported in Table 3. Similarly for the KPT column.

We find that the relationship between racial disparity and indicators of biased policing is quite weak. First, at the top quintile of the disparity index, the probability of rejecting unbiased policing for a police force is quite low. For instance, using drug finds as search success, we find that police forces that fail the AGO test of bias against black or Hispanic drivers and that fall in the top quintile of the stop-disparity index account for only about 2 to 3 percent of all non-ITA searches. Agencies that fail the KPT test in the top quintile represent about 3 to 6 percent of all non-ITA searches. Using contraband finds as search success, the fraction of non-ITA searches represented by police forces that fall in the top stop-disparity quintile and fail the AGO test of racial bias against blacks is about 3 percent, and about 3 percent for those agencies that fail the AGO test of bias against Hispanics. Agencies that fail the KPT tests of bias against blacks account for 6 percent, and about 3 percent for the test on searches of Hispanic drivers.

Such a low rejection rate might seem to be a consequence of the weakness of the AGO test, but our second result shows this not so: we reject unbiased policing for police forces at the bottom quintile of the index at rates that are not systematically lower than for the top quintile. Indeed, for Hispanics, the rejection rate is much higher at the bottom than the top of the distribution of the disparity index. For instance, the AGO test for drug finds rejects unbiased policing for agencies representing about 14 percent of all non-ITA searches by agencies in the bottom quintile of the search-disparity index. For blacks, the fraction of searches represented by police forces in the bottom quintile who fail the AGO test is about 0.6 percent. Hence, according to our tests, not only does the disparity-index approach identify too many police forces as racially biased, it also identifies the wrong ones. Of course, since our AGO test is quite conservative, in the sense that it allows ITA hit rates to be zero, it is not surprising that we reject at a low rate overall, but the fact that the correlation with the disparity index appears so low indicates that concerns with racial bias are not well served by focussing on the disparity index as a performance measure.[21]

[Table 7 about here.]

Three of the largest police forces in the state are the Missouri State Police, the St. Louis City Police Department, and the Kansas City Police Department. Table 7 shows that these account for about, respectively, 13 percent, 5 percent and 3 percent of all non-ITA searches in our data set.

[21]We conducted this exercise with the alternative search-disparity index described earlier, and we reached similar conclusions.

The city departments account for much larger shares of searches of blacks: about 18 percent (St. Louis City) and 10 percent (Kansas City). The State Police accounts for 24 percent of all non-ITA searches of Hispanics, and the city departments account only for about 2 percent (St. Louis City) and 8 percent (Kansas City). Under the AGO test for drug finds, we reject unbiased policing with respect to Hispanics for the State Police with all three measures of search success, and we reject unbiased policing with respect to Hispanics for the St. Louis City police department only when using Drug Finds as the measure of search success. Hence we conclude that both of these agencies are biased against Hispanic motorists. However with respect to searches of black motorists we do not reject unbiased policing for any of the three police forces. While the state police accounts for a large share of Hispanic searches, the fraction of Hispanic searches of police forces that fail the AGO test is much larger than this share, and even in the absence of this agency, we would still reject unbiased policing for forces representing about 10 to 15 percent of the searches of Hispanics.

[Table 8 about here.]

[Table 9 about here.]

In Table 8 we present some statistics on demographic characteristics of police forces that fail the AGO test of unbiased policing behavior, and similarly for police forces that do not fail in Table 9. When using drug or contraband finds as measures of search success, we learn that, on average, areas with police forces that are biased have a larger fraction of white population. These findings confirm the explanation we offered earlier for the fact that searches by police biased against minorities account for a much larger share of all discretionary searches than of searches of minorities.

We also learn that areas with police forces that fail the tests of unbiased policing against Hispanic drivers appear to be more affluent than areas where police forces do not fail the AGO test. This is indicated by the average and median values for the median household income, which are higher in areas with biased police forces, and by the average and median values for the poverty rate, which are lower in areas with biased police forces. Areas with police forces biased against black drivers appear to be less affluent that areas where police forces do not fail the AGO test. The fraction of sheriff departments appears to be higher among police forces that seem biased against minorities. Finally, police forces that failed the tests of unbiased policing against black drivers conducted fewer stops and searches than polices forces which did not fail the tests, but police forces that failed the test of unbiased policing against Hispanic drivers conducted more stops and searches than police agencies which did not fail the tests.

The overall impression that emerges from this analysis is that black and Hispanic motorists are most at risk from biased police forces when travelling through areas where minorities represent a relatively small population share. It is not clear from the data why the patterns should differ, but it does suggest the possibility that blacks or Hispanics are searched because their skin color makes them seem more out of place in small white towns. This would be consistent with the opinion expressed by the court in State v. Dean, that good police work requires officers to be more suspicious of motorists who do not fit in with the local community.[22]

[22]In the 1975 case *State v. Dean*, the Arizona Supreme Court stated that "the fact that a person is obviously out of place in a particular neighborhood is one of several factors that may be considered by an officer". Regardless of

7 Conclusion

In this paper we extended the KPT test for biased policing to cover cases where data were available only in the form of summary statistics. This allows the test to be applied to data sets covering a much larger number of motorists than was the case for previous applications, which were based on individual records, data which are rarely in the public domain. However when such aggregate data do not distinguish between the results of discretionary and non-discretionary searches, the test cannot be applied directly.

The 2001 Annual Traffic-Stop Report of the Missouri Attorney General's Office (AGO) covers 1,389,947 traffic stops. The report claims that differences in search-success rates by race should not be interpreted as evidence of racial bias because the statistics reflect mandatory as well as discretionary searches. We used bounds methods to develop tests for racial bias at the level of the individual police force that allow us to examine the plausibility of this explanation. We applied our bounds tests at the level of individual police forces, and found evidence in support of racial bias against African-American and Hispanic motorists, even when allowing for lower success rates of non-discretionary searches.

Our first test involved non-parametric bounds. These were quite wide, and did not permit rejection of unbiased policing except for a few police forces accounting for a very small fraction, about 2 percent, of all discretionary searches. By contrast, the KPT test resulted in rejection of unbiased policing for police forces accounting for 40 to 67 percent of searches when success was measured according to drug finds, and 40 to 65 percent for other measures. Since the KPT test does not allow for non-discretionary search, the larger rejection rates are not surprising. Thus these results are consistent with the statement in the AGO report.

We therefore refined our test by imposing a restriction that we inferred from the AGO statement: that success rates of non-discretionary searches are lower than for discretionary searches. We showed how this would theoretically tighten the bounds, as well as the confidence intervals around our non-parametric bounds. Our basic finding was that a large share of searches in Missouri are carried out by police forces who appear biased against minority motorists. Most of the bias appears to be directed against Hispanic rather then black motorists. The police forces that appear biased against black motorists tend to be found in areas with a relatively large white population.

We also found that restricting the lower bound on non-discretionary searches of minorities did not have a strong effect on the results for blacks until the lower bound rose above 75 percent of the find rate for discretionary searches, while for Hispanics, the measure of biased policing increased significantly for even relatively low values of this lower bound. Thus the AGO explanation is more persuasive with respect to blacks than Hispanics. We also found that allowing the non-discretionary find rate to rise as much as 25 percent above that for discretionary searches would still result in rejection rates substantially above those of the non-parametric bounds.

The AGO report suggests measuring the disparity between the population compositions of motorists searched and of locations residents as a rough guide to racial profiling. We find the

whether this view is still acceptable, the point is that the court at that time felt that this was part of standard police practice. Our test results suggest that this practice is associated with biased policing that is not justified by the search outcomes.

relation between this profiling measure and our measure of racial bias to be fairly weak. While for black motorists we are more likely to reject biased policing for agencies in the top quintile of the disparity index distribution, we only reject police forces in the top quintile that account for about 3 percent of all discretionary searches. For Hispanic motorists, we are much more likely to reject police forces in the bottom quintile of the disparity index than at the top, indicating that the disparity index is not useful for identifying racial bias.

Finally, our result that a significant number of searches are carried out by racially biased police forces is not the result of a few large, biased, police forces. For the three largest police forces in the state, we do not reject unbiased searches with respect to blacks. However we do find that the searches by the State Police are biased against Hispanics. While the State Police accounts for a large share of searches of Hispanics, we rejected unbiased policing with respect to a much larger share.

State data collection efforts could make testing easier with some improvements in the way they present the traffic statistics to the public and researchers. First, states should emphasize the collection of outcome variables, such as the nature and quantity of contraband found, as well as citations and arrests resulting from the search. In addition, it would be extremely helpful if the statistics reported tabulated the outcomes of discretionary and non-discretionary searches separately. While our tests are able to extract some information from the combined statistics currently issued, the actual share of searches accounted for by biased police may be significantly higher than what our results suggest.

The model that gives rise to our tests is based on a few key economic assumptions that could be the focus of future research. For instance, we assume that police act as though they were maximizing the number of arrests, and that they are aware of the probability of guilt conditional on observables. Likewise, we assume that motorists are aware of the probability of being searched, and that the risk of search deters them from carrying contraband. While standard for economic modelling, these assumptions are not known to be true, and our results should be viewed as illustrating the potential of the economic approach rather than as definitive judgments. [23]

Subject to the above proviso, our results suggest that police bias or animus against minorities is a plausible explanation of at least some of the disparate impact faced by minorities in Missouri, and that therefore the problem of disparate impact by race could be significantly reduced without reducing police effectiveness. According to Kennedy (1997), the mainstream view of the courts, at least until recently, was that "race can be appropriately used as a factor of suspicion... so long as this use of race is reasonably related to law enforcement and not a mere pretext for racial harassment". In the case of Missouri, to the extent that the objective of discretionary motor-vehicle searches is to detect the transport of contraband, it appears that a significant share of the excess burden faced by minorities is unrelated to law enforcement.

[23] Alternative economic interpretations of racial patterns in police search data are offered in Dharmapala and Ross (2004) and Bunzel and Marcoul (2003); these argue that the KPT test is in fact too weak, and that equal find rates can arise even when the police are racially biased.

Appendix: Confidence Intervals for Search Success Rates

For each of the different cases that we examine, we follow Horowitz and Manski (2000) and construct $100\times\alpha$ *percent* confidence intervals for the identification bounds of the form $\left[\hat{P}_L - D, \hat{P}_H + D\right]$, solving for a level of D to ensure that, conditional on the probability of non-discretionary search, the probability of coverage of the identification region $[P_L, P_H]$ is at least α :

$$\Pr\left(\hat{P}_L - D \leq P_L, P_H \leq \hat{P}_H + D | \hat{\pi}\right) \geq \alpha.$$

In the following sections, we derive an expression containing D that can be easily solved numerically. For our tests we compute values of D for each race-agency cell, and check whether the results imply that the confidence interval for white motorists lies above that for minority motorists.

A HM Bounds

Let $\hat{Q}$ indicate the sample search success rate in a race-agency cell with N searches. Since successes in each search are independent of success in other searches carried out by the agency, $\hat{Q}$ follows a binomial distribution, with expectation Q and variance $\sigma^2 = Q(1-Q)$. Standard asymptotic results imply convergence to the standard normal distribution:

$$\frac{\sqrt{N}}{\sigma}\left(\hat{Q} - Q\right) \xrightarrow{d} \mathcal{N}(0,1).$$

According to equation (4), after suppressing the race superscript, the discretionary find rate is somewhere inside the identification region:

$$P_1 \in \left[\frac{Q-\pi}{1-\pi}, \frac{Q}{1-\pi}\right]. \tag{6}$$

The sample analog of this interval is

$$\left[\frac{\hat{Q}-\hat{\pi}}{1-\hat{\pi}}, \frac{\hat{Q}}{1-\hat{\pi}}\right] \equiv \left[\hat{P}_L, \hat{P}_H\right] \tag{7}$$

Following Imbens and Manski (2003) we construct a confidence interval for the parameter P_1 of the form $\left[\hat{P}_L - D, \hat{P}_H + D\right]$ with the distance D constructed to guarantee that

$$\Pr\left(\hat{P}_L - D \leq P_1 \leq \hat{P}_H + D\right) = \alpha.$$

Using the formula for the true value of P_1, derived from equation (3),

$$\Pr\left(\hat{P}_L - D \leq \frac{Q-\pi P_0}{1-\pi} \leq \hat{P}_H + D\right) = \alpha.$$

Now we apply the formulas for the estimated upper and lower bounds:

$$\Pr\left(\frac{\hat{Q}-\hat{\pi}}{1-\hat{\pi}} - D \leq \frac{Q-\pi P_0}{1-\pi} \leq \frac{\hat{Q}}{1-\hat{\pi}} + D\right) = \alpha.$$

Conditioning on $\hat{\pi}$ and rearranging we obtain:

$$\Pr\left(-\frac{[\pi(1-P_0)+D(1-\pi)]}{\sigma/\sqrt{N}} \leq \frac{Q-\hat{Q}}{\sigma/\sqrt{N}} \leq \frac{[\pi P_0 + D(1-\pi)]}{\sigma/\sqrt{N}}\middle|\hat{\pi}\right) = \alpha.$$

To implement this, we set the parameters σ and π to their observed values; thus D must solve:

$$\Phi\left(\frac{[\hat{\pi}P_0 + D(1-\hat{\pi})]}{\hat{\sigma}/\sqrt{N}}\right) - \Phi\left(-\frac{[\hat{\pi}(1-P_0)+D(1-\hat{\pi})]}{\hat{\sigma}/\sqrt{N}}\right) = \alpha.$$

Using the fact that this expression is concave in P_0, the probability is minimized at each of the end points $P_0 \in \{0,1\}$. So letting $P_0 = 0$ (which yields the same value as $P_0 = 1$) we must set D so that

$$\Phi\left(\frac{D(1-\hat{\pi})}{\hat{\sigma}/\sqrt{N}}\right) - \Phi\left(-\frac{[\hat{\pi}+D(1-\hat{\pi})]}{\hat{\sigma}/\sqrt{N}}\right) = \alpha.$$

B AGO Bounds

For this and the remaining cases we follow Horowitz and Manski (2000) and construct $100\times\alpha\%$ confidence intervals for the identification bounds with the form $\left[\hat{P}_L - D, \hat{P}_H + D\right]$. The bounds assume that the search success rate for non-discretionary searches is lower than for discretionary searches, $P_0^r \leq P_1^r$, as in equation (5). Suppressing again the race superscript, the bounds are

$$\hat{P}_L = \hat{Q} \quad \text{and} \quad \hat{P}_H = \frac{\hat{Q}}{1-\pi}.$$

The distance D in the confidence intervals is constructed to ensure that, conditional on the probability of a non-discretionary search, the probability of coverage of the identification region $[P_L, P_H]$ is at least α :

$$\Pr\left(\hat{P}_L - D \leq P_L, P_H \leq \hat{P}_H + D\right) = \alpha.$$

Replacing the expressions for the bounds and after some manipulations, we can express the above as

$$\Pr\left(-\frac{D}{\sigma/\sqrt{N}} \leq \frac{Q-\hat{Q}}{\sigma/\sqrt{N}} \leq \frac{D(1-\pi)}{\sigma/\sqrt{N}}|\hat{\pi}\right) = \alpha.$$

We determine D by replacing the parameters σ and π with their sample values:

$$\Phi\left(\frac{D(1-\hat{\pi})}{\hat{\sigma}/\sqrt{N}}\right) - \Phi\left(-\frac{D}{\hat{\sigma}/\sqrt{N}}\right) = \alpha.$$

B.1 Parameterized Bounds

B.1.1 AGO and KPT

In this case, the bounds are constructed under the assumption that $P_0^r \in [\gamma P_1^r, P_1^r]$. Notice that setting $\gamma = 0$ is equivalent to the AGO bounds defined above, and setting $\gamma = 1$ is equivalent to assuming that $P_0^r = P_1^r = Q^r$, which yields the test where we do not account for the effect of non-discretionary searches. The bounds are given by

$$\hat{P}_L = \hat{Q} \quad \text{and} \quad \hat{P}_H = \frac{\hat{Q}}{(1-\pi)+\pi\gamma}.$$

The coverage probability is given by

$$\Pr\left(-\frac{D}{\sigma/\sqrt{N}} \leq \frac{Q-\hat{Q}}{\sigma/\sqrt{N}} \leq \frac{D\left[(1-\pi)+\pi\gamma\right]}{\sigma/\sqrt{N}}|\hat{\pi}\right) = \alpha.$$

Replacing the sample values of σ and π, we have that D must solve the following expression:

$$\Phi\left(\frac{D\left[(1-\hat{\pi})+\hat{\pi}\gamma\right]}{\hat{\sigma}/\sqrt{N}}\right) - \Phi\left(-\frac{D}{\hat{\sigma}/\sqrt{N}}\right) = \alpha.$$

Notice that setting $\gamma = 1$, in the KPT case, the constant reduces to

$$D = \Phi^{-1}\left(\frac{1+\alpha}{2}\right)\frac{\hat{\sigma}}{\sqrt{N}},$$

which yields the standard confidence intervals for the sample success rate $\hat{Q}$.[24]

B.1.2 AGO and HM

In this case the bounds are constructed assuming that the search success rate of non-discretionary searches is bounded above by a number that may exceed the search success rate of discretionary searches. To implement this, the upper bound is set at the value of P_1 corresponding to $P_0 = 0$, i.e., $P_H = \frac{Q}{1-\pi}$. We then parameterize the lower bound to cover values of $P_0 \in [P_1, 1]$ by setting $P_0 = \lambda P_1 + (1-\lambda)$, and letting $\lambda \in [0,1]$. Thus when $\lambda = 0$ we have $P_0 = 1$, which gives us the same lower bound as in the HM bounds, $P_L = \frac{Q-\pi}{1-\pi}$, and when $\lambda = 1$, we have $P_0 = P_1$, which gives us the lower bound as in the AGO bounds, $P_L = Q$. The expressions for the bounds are then

$$\hat{P}_L = \frac{\hat{Q}-\pi(1-\lambda)}{(1-\pi)+\pi\lambda} \quad \text{and} \quad \hat{P}_H = \frac{\hat{Q}}{1-\pi}.$$

The expression for the coverage probability reduces in this case to

$$\Pr\left(-\frac{D\left[(1-\pi)+\pi\lambda\right]}{\sigma/\sqrt{N}} \leq \frac{Q-\hat{Q}}{\sigma/\sqrt{N}} \leq \frac{D(1-\pi)}{\sigma/\sqrt{N}}|\hat{\pi}\right) = \alpha.$$

[24]Notice also that all of our confidence intervals coincide with the standard confidence intervals for the sample success rate $\hat{Q}$ when the search success rate is point-identified, that is when $\pi = 0$.

Therefore D must solve the following expression.

$$\Phi\left(\frac{D\,(1-\hat{\pi})}{\hat{\sigma}/\sqrt{N}}\right)-\Phi\left(-\frac{D\,[(1-\hat{\pi})+\hat{\pi}\lambda]}{\hat{\sigma}/\sqrt{N}}\right)=\alpha.$$

References

Becker, G. S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76 (1968), 169–217.

Bunzel, H. and P. Marcoul, "Can Racially Unbiased Police Perpetuate Long-Run Discrimination?" Manuscript, Iowa State University, 2003.

Dharmapala, D. and S. L. Ross, "Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence," Manuscript, University of Connecticut, 2004.

Dominitz, J. and R. P. Sherman, "Nonparametric Analysis of Mixture Models with Verification with an Application to Test Score Data," Manuscript, Carnegie Mellon University, 2003.

Gross, S. R. and K. Y. Barnes, "Road Work: Racial Profiling and Drug Interdiction on the Highway," *Michigan Law Review* 101 (2002), 651–754.

Horowitz, J. L. and C. F. Manski, "Identification and Robustness with Contaminated and Corrupted Data," *Econometrica* 63 (1995), 281–302.

—, "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data," *Journal of the American Statistical Association* 95 (2000), 77–84.

Imbens, G. W. and C. F. Manski, "Confidence Intervals for Partially Identified Parameters," Manuscript, Northwestern University, 2003.

Kennedy, R., *Race, Crime and the Law* (New York: Pantheon Books, 1997).

Knowles, J., N. Persico and P. Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy* 109 (2001), 203–229.

Kreider, B. and J. Pepper, "Inferring Disability Status from Corrupt Data," Manuscript, Iowa State University, 2003.

Manski, C. F., *Identification Problems in the Social Sciences* (Cambridge, MA: Harvard University Press, 1995).

Pepper, J., "The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis," *Review of Economics and Statistics* 82 (2000), 472–488.

U.S. Department of Justice, "Traffic Stop Data Collection Policies for State Police, 2001," Bureau of Justice Statistics: Fact Sheet, 2001.

Zingraff, M. T., H. M. Mason, W. R. Smith, D. Tomaskovic-Devey, P. Warren, H. L. McMurray and C. R. Fenlon, "Evaluating North Carolina State Highway Patrol Data: Citations, Warnings, and Searches in 1998," North Carolina Department of Crime Control and Public Safety, 2000.

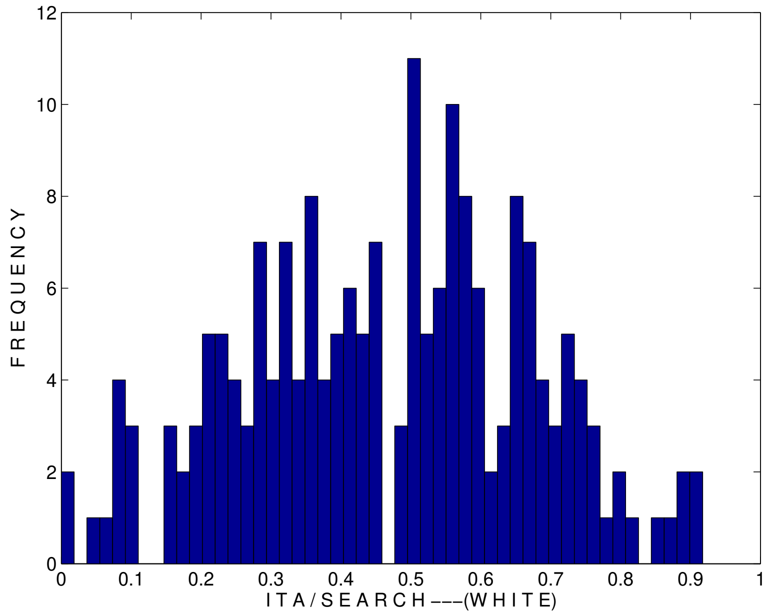

Figure 1: ITA frequency White Drivers.

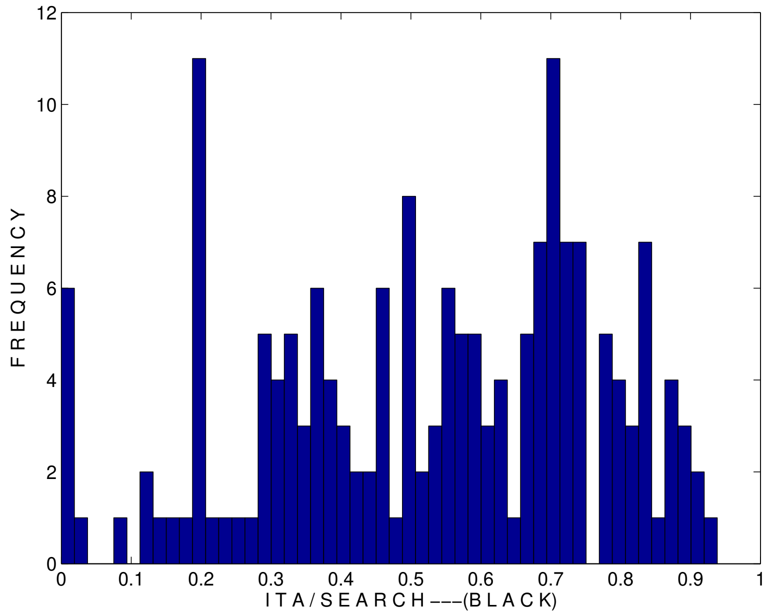

Figure 2: ITA frequency Black Drivers.

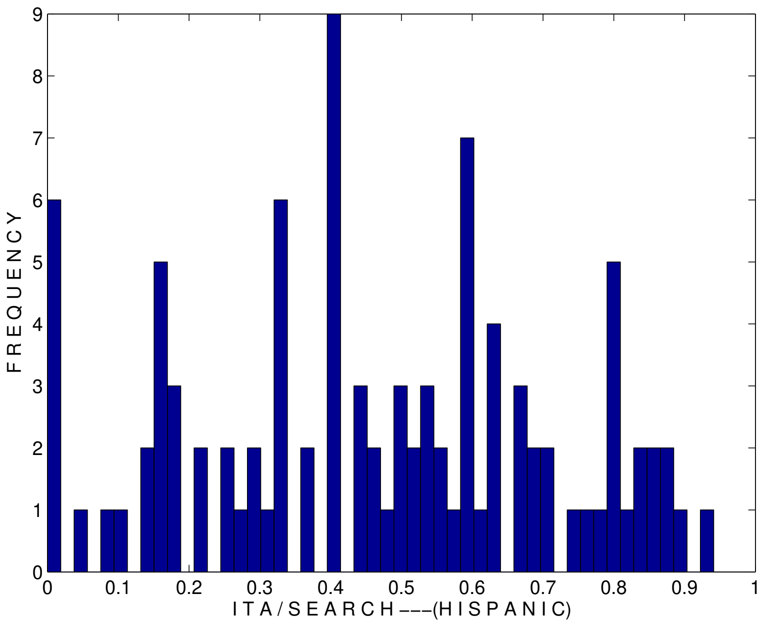

Figure 3: ITA frequency Hispanic Drivers.

Table 1: Characteristics of law enforcement agencies with 5 or more searches for both white and African American drivers or for both white and Hispanic drivers.

Race of Motorist	Statistic[a]	Stops	Searches	ITA searches	Drug Finds	Contraband Finds	Search Arrests	Stops of Young[b] Drivers	Stops of Male Drivers
All[c]	sum	1,158,922	84,652	44,475	14,434	17,515	11,630	572,842	764,215
	mean	6,067.65	443.20	232.85	75.57	91.70	60.89	2,999.17	4,001.13
	median	2,243.00	168.00	78.00	25.00	34.00	28.00	1,124.00	1,475.00
	st. dev.	24,611	1,125	635	328	362	116	11,130	16,925
White	sum	963,074	61,209	30,781	11,663	13,909	8,677	470,108	631,004
	mean	5,042.27	320.47	161.16	61.06	72.82	45.43	2,461.30	3,303.69
	median	1,853.00	131.00	50.00	20.00	25.00	22.00	948.00	1,173.00
	st. dev.	22,100	868	489	292	320	88	9,885	15,132
Black	sum	174,193	20,574	12,170	2,524	3,267	2,716	90,141	115,266
	mean	912.01	107.72	63.72	13.21	17.10	14.22	471.94	603.49
	median	160.00	23.00	11.00	3.00	3.00	3.00	74.00	107.00
	st. dev.	3,173	355	219	45	58	44	1,654	2,164
Hispanics	sum	21,655	2,869	1,524	247	339	237	12,593	17,945
	mean	113.38	15.02	7.98	1.29	1.77	1.24	65.93	93.95
	median	30.00	5.00	2.00	0.00	0.00	0.00	18.00	25.00
	st. dev.	513	49	25	7	7	3	290	430

[a]The mean and standard deviations were computed using equal weights.

[b]*Young* refers to drivers of age 29 or younger.

[c]Number of agencies = 191

Table 2: Characteristics of search and find rates in agencies with 5 or more searches for both white and African American drivers or for both white and Hispanic drivers.

Race of Motorist	Statistic[a]	Searches /Stops	ITA Searches /Searches	Drug Finds /Searches	Contraband Finds /Searches	Searchs Arrests /Searches	Stops of Young [b] Drivers /Stops	Stops of Male Drivers /Stops
All[c]	proportion	0.0730	0.5254	0.1705	0.2069	0.1374	0.4943	0.6594
	st. dev.	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
White	proportion	0.0636	0.5029	0.1905	0.2272	0.1418	0.4881	0.6552
	st. dev.	(0.001)	(0.003)	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)
Black	proportion	0.1181	0.5915	0.1227	0.1588	0.1320	0.5175	0.6617
	st. dev.	(0.002)	(0.004)	(0.007)	(0.006)	(0.006)	(0.002)	(0.001)
Hispanics	proportion	0.1325	0.5312	0.0861	0.1182	0.0826	0.5815	0.8287
	st. dev.	(0.006)	(0.013)	(0.018)	(0.018)	(0.018)	(0.004)	(0.003)

Disparity Indices

		Population Share	Stops Share	Search Share	Stops share /Population share	Searches share /Population share
	White	0.85	0.83	0.72	0.97	0.85
	Black	0.10	0.15	0.24	1.48	2.40
	Hispanics	0.02	0.02	0.03	1.03	1.86

[a]The standard deviation for the proportions were computed assumming a binomial distribution.
[b]*Young* refers to drivers of age 29 or younger.
[c]Number of agencies = 191

Table 3: Share of Searches by Biased Police Forces. KPT and Bounds Tests.

	Bounds	non-ITA Searches		Searches of Blacks		Searches of Hisp.	
		99%CI	95% CI	99%CI	95% CI	99%CI	95% CI
Guilt 1: Drug Finds							
Black	KPT	0.3123	0.3408	0.2242	0.2631	0.3197	0.3398
	HM	0.0005	0.0015	0.0006	0.0012	0.0007	0.0007
	AGO	0.0658	0.0710	0.0074	0.0113	0.0312	0.0342
Hispanic	KPT	0.4654	0.6227	0.4110	0.6103	0.4974	0.6015
	HM	0.0000	0.0015	0.0000	0.0001	0.0000	0.0045
	AGO	0.3791	0.3988	0.3364	0.3588	0.3717	0.3926
Minority	KPT	0.5985	0.6673	0.5704	0.6418	0.5546	0.6297
	HM	0.0005	0.0029	0.0006	0.0013	0.0007	0.0052
	AGO	0.3995	0.4243	0.3425	0.3688	0.3792	0.4030
Guilt 2: Contraband Finds							
Black	KPT	0.3304	0.3574	0.2649	0.2934	0.3249	0.3487
	HM	0.0005	0.0005	0.0006	0.0006	0.0007	0.0007
	AGO	0.0597	0.0628	0.0064	0.0104	0.0260	0.0268
Hispanic	KPT	0.3733	0.3883	0.2197	0.2250	0.4253	0.4513
	HM	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	AGO	0.2901	0.3244	0.1539	0.1787	0.3219	0.3539
Minority	KPT	0.5194	0.5525	0.4022	0.4294	0.4900	0.5323
	HM	0.0005	0.0005	0.0006	0.0006	0.0007	0.0007
	AGO	0.3093	0.3466	0.1592	0.1880	0.3286	0.3613
Guilt 3: Arrests from Searches							
Black	KPT	0.0887	0.2241	0.0499	0.2262	0.0662	0.1316
	HM	0.0235	0.0235	0.0006	0.0006	0.0126	0.0126
	AGO	0.0459	0.0548	0.0049	0.0099	0.0230	0.0283
Hispanic	KPT	0.2947	0.4200	0.1872	0.3644	0.3881	0.4565
	HM	0.0235	0.0235	0.0006	0.0006	0.0126	0.0126
	AGO	0.1148	0.2693	0.0683	0.1611	0.0892	0.3472
Minority	KPT	0.3313	0.4600	0.2183	0.3816	0.4030	0.4803
	HM	0.0235	0.0235	0.0006	0.0006	0.0126	0.0126
	AGO	0.1324	0.2926	0.0723	0.1688	0.0952	0.3546

Table 4: Rejection of Unbiased Policing. Parameterized AGO/KPT Bounds.

	γ	Drug Finds		Contraband Finds		Search Arrests	
		99%CI	95% CI	99%CI	95% CI	99%CI	95% CI
Black	0.00	0.0658	0.0710	0.0597	0.0628	0.0459	0.0548
	0.25	0.0686	0.0714	0.0605	0.0632	0.0494	0.0552
	0.50	0.0693	0.0714	0.0625	0.0656	0.0514	0.0573
	0.75	0.0714	0.2071	0.1695	0.1747	0.0534	0.0950
	1.00	0.3123	0.3408	0.3304	0.3574	0.0887	0.2241
Hispanic	0.00	0.3791	0.3988	0.2901	0.3244	0.1148	0.2693
	0.25	0.3928	0.5270	0.3184	0.3315	0.2639	0.2927
	0.50	0.4150	0.5659	0.3248	0.3655	0.2911	0.2947
	0.75	0.4249	0.6064	0.3557	0.3883	0.2947	0.3130
	1.00	0.4654	0.6227	0.3733	0.3883	0.2947	0.4200
Minority	0.00	0.3995	0.4243	0.3093	0.3466	0.1324	0.2926
	0.25	0.4160	0.5530	0.3384	0.3541	0.2838	0.3160
	0.50	0.4389	0.5919	0.3468	0.3905	0.3110	0.3200
	0.75	0.4509	0.6342	0.4847	0.5225	0.3161	0.3407
	1.00	0.5985	0.6673	0.5194	0.5525	0.3313	0.4600

Table 5: Rejection of Unbiased Policing. Parameterized AGO/HM Bounds.

	λ	Drug Finds		Contraband Finds		Search Arrests	
		99%CI	95% CI	99%CI	95% CI	99%CI	95% CI
	0.500	0.0027	0.0085	0.0052	0.0081	0.0257	0.0257
Black	0.625	0.0052	0.0096	0.0287	0.0362	0.0257	0.0269
	0.750	0.0313	0.0369	0.0348	0.0389	0.0257	0.0360
	0.875	0.0635	0.0658	0.0581	0.0620	0.0360	0.0393
	1.000	0.0658	0.0710	0.0597	0.0628	0.0459	0.0548
	0.500	0.0015	0.0015	0.0000	0.0000	0.0310	0.0374
Hispanic	0.625	0.0325	0.0548	0.0362	0.0575	0.0374	0.0540
	0.750	0.0976	0.1199	0.0832	0.0960	0.0419	0.0540
	0.875	0.1582	0.2323	0.1254	0.2990	0.0540	0.0747
	1.000	0.3791	0.3988	0.2901	0.3244	0.1148	0.2693
	0.500	0.0042	0.0099	0.0052	0.0081	0.0332	0.0396
Minority	0.625	0.0377	0.0644	0.0414	0.0702	0.0396	0.0574
	0.750	0.1054	0.1333	0.0945	0.1114	0.0442	0.0665
	0.875	0.1763	0.2527	0.1429	0.3205	0.0665	0.0905
	1.000	0.3995	0.4243	0.3093	0.3466	0.1324	0.2926

Table 6: Rejection of Unbiased Policing and the Stop-Disparity Index.

	Quintile	non-ITA Searches		Searches of Blacks		Searches of Hisp.	
		AGO	KPT	AGO	KPT	AGO	KPT
Guilt 1: Drug Finds							
Black	0-20	0.0068	0.1407	0.0015	0.0650	0.0007	0.2394
	20-40	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	40-60	0.0010	0.1096	0.0006	0.1508	0.0000	0.0468
	60-80	0.0310	0.0310	0.0056	0.0056	0.0149	0.0149
	80-100	0.0322	0.0594	0.0036	0.0418	0.0186	0.0387
Hispanic	0-20	0.1425	0.2888	0.2394	0.4403	0.0572	0.1398
	20-40	0.0692	0.0795	0.0289	0.0346	0.0461	0.0617
	40-60	0.1354	0.1606	0.0635	0.0839	0.2431	0.2833
	60-80	0.0162	0.0404	0.0130	0.0243	0.0126	0.0506
	80-100	0.0354	0.0533	0.0139	0.0272	0.0335	0.0662
Minority	0-20	0.2350	0.2375	0.2462	0.2468	0.2840	0.2877
	20-40	0.0183	0.0360	0.0030	0.0109	0.0037	0.0312
	40-60	0.0160	0.1369	0.0143	0.1680	0.0208	0.0818
	60-80	0.0866	0.1288	0.0614	0.0977	0.0572	0.1056
	80-100	0.0684	0.1280	0.0439	0.1184	0.0372	0.1234
Guilt 2: Contraband Finds							
Black	0-20	0.0074	0.1414	0.0024	0.0658	0.0015	0.2401
	20-40	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	40-60	0.0000	0.1175	0.0000	0.1568	0.0000	0.0543
	60-80	0.0244	0.0401	0.0048	0.0293	0.0097	0.0171
	80-100	0.0310	0.0585	0.0032	0.0415	0.0156	0.0372
Hispanic	0-20	0.0888	0.1009	0.0754	0.0851	0.0476	0.0654
	20-40	0.0655	0.0758	0.0232	0.0289	0.0349	0.0506
	40-60	0.1340	0.1591	0.0634	0.0838	0.2387	0.2788
	60-80	0.0088	0.0201	0.0067	0.0099	0.0074	0.0164
	80-100	0.0273	0.0323	0.0100	0.0174	0.0253	0.0401
Minority	0-20	0.1867	0.1893	0.0671	0.0677	0.2639	0.2677
	20-40	0.0000	0.0080	0.0000	0.0046	0.0000	0.0126
	40-60	0.0088	0.1298	0.0067	0.1604	0.0074	0.0684
	60-80	0.0826	0.1117	0.0591	0.0904	0.0550	0.0848
	80-100	0.0684	0.1138	0.0551	0.1063	0.0349	0.0989
Guilt 3: Arrests from Searches							
Black	0-20	0.0065	0.0083	0.0015	0.0025	0.0037	0.0037
	20-40	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	40-60	0.0020	0.1107	0.0008	0.1506	0.0007	0.0476
	60-80	0.0154	0.0410	0.0045	0.0337	0.0089	0.0320
	80-100	0.0308	0.0641	0.0030	0.0394	0.0149	0.0483
Hispanic	0-20	0.0430	0.1652	0.0564	0.2288	0.0275	0.0818
	20-40	0.0395	0.0395	0.0152	0.0152	0.0253	0.0253
	40-60	0.1387	0.1608	0.0658	0.0851	0.2424	0.2758
	60-80	0.0197	0.0228	0.0133	0.0181	0.0186	0.0327
	80-100	0.0285	0.0318	0.0104	0.0171	0.0335	0.0409
Minority	0-20	0.1822	0.1840	0.0690	0.0700	0.2669	0.2669
	20-40	0.0012	0.0012	0.0005	0.0005	0.0030	0.0030
	40-60	0.0179	0.1266	0.0081	0.1579	0.0149	0.0617
	60-80	0.0405	0.0667	0.0396	0.0689	0.0357	0.0595
	80-100	0.0508	0.0815	0.0516	0.0844	0.0342	0.0892

Table 7: Largest Police Departments.

	Missouri State Police		St. Louis City P.D.		Kansas City P.D.	
Race	non-ITA Searches	Share of Total	non-ITA Searches	Share of Total	non-ITA Searches	Share of Total
All	5,382	0.1340	1,907	0.0475	1,379	0.0343
Whites	4,528	0.1488	374	0.0123	448	0.0147
Blacks	533	0.0634	1,511	0.1798	828	0.0985
Hispanic	321	0.2387	22	0.0164	103	0.0766

Rejection of the Statistical Discrimination Model

Bounds	Blacks	Hispanics	Blacks	Hispanics	Blacks	Hispanics
Guilt 1: Drug Finds						
KPT	yes	yes	no	yes	no	no
HM	no	no	no	no	no	no
AGO	no	yes	no	yes	no	no
Guilt 2: Contraband Finds						
KPT	yes	yes	no	no	no	no
HM	no	no	no	no	no	no
AGO	no	yes	no	no	no	no
Guilt 3: Arrests from Searches						
KPT	no	yes	no	no	no	no
HM	no	no	no	no	no	no
AGO	no	yes	no	no	no	no

Table 8: Characteristics of Police Forces that Fail the AGO Test of Unbiased Policing with 95% Confidence Intervals.

Variable	Statistic	Drug Finds		Contraband Finds		Arrests from Searches	
		Minority					
		Black	Hispanic	Black	Hispanic	Black	Hispanic
Median	mean	34,117.79	45,748.69	33,052.00	43,476.17	32,750.13	40,563.36
Household	median	32,772.00	40,322.00	32,118.00	40,322.00	32,772.00	35,647.00
Income (1999)	st. dev.	8,280.16	22,222.23	7,360.80	14,045.16	5,559.74	22,937.21
Share of	mean	0.14	0.10	0.14	0.09	0.14	0.11
Poor	median	0.12	0.10	0.12	0.08	0.12	0.12
Population	st. dev.	0.08	0.06	0.08	0.05	0.07	0.05
Share of	mean	0.93	0.89	0.92	0.90	0.93	0.88
White	median	0.96	0.94	0.96	0.94	0.96	0.90
Population	st. dev.	0.08	0.14	0.09	0.11	0.08	0.11
Share of	mean	0.05	0.07	0.06	0.06	0.04	0.08
Black	median	0.02	0.02	0.02	0.02	0.02	0.05
Population	st. dev.	0.08	0.13	0.09	0.11	0.08	0.11
Share of	mean	0.01	0.02	0.01	0.02	0.02	0.02
Hispanic	median	0.01	0.02	0.01	0.01	0.01	0.02
Population	st. dev.	0.01	0.01	0.01	0.01	0.02	0.01
Total	mean	1,814.79	5,812.44	1,637.00	5,137.25	1,558.00	3,570.16
Stops	median	1,365.00	3,415.00	1,012.00	4,193.50	1,622.00	2,875.00
	st. dev.	2,062.63	7,317.12	2,117.41	3,743.29	909.55	2,828.93
	sum	34,481	185,998	26,192	123,294	23,370	89,254
Total	mean	241.21	578.91	241.88	573.25	206.13	402.96
Searches	median	126.00	413.50	97.50	515.50	163.00	305.00
	st. dev.	365.07	562.14	397.22	381.41	239.46	301.60
	sum	4,583	18,525	3,870	13,758	3,092	10,074
Number of agencies[a]	sum	19	32	16	24	15	25
Sheriff Depts.	percentage	0.32	0.25	0.38	0.25	0.27	0.16

[a]The State Police was excluded to compare the demographic indicators.

Table 9: Characteristics of Police Forces that Do Not Fail the AGO Test of Unbiased Policing with 95% Confidence Intervals.

Variable	Statistic	Drug Finds		Contraband Finds		Arrests from Searches	
		Minority					
		Black	Hispanic	Black	Hispanic	Black	Hispanic
Median	mean	41,559.94	39,816.64	41,529.63	40,431.08	41,507.06	40,853.96
Household	median	36,117.00	35,352.00	36,240.00	35,177.00	36,363.00	36,111.00
Income (1999)	st. dev.	20,208.94	18,778.69	20,078.40	20,132.57	20,074.18	18,964.15
Share of	mean	0.12	0.12	0.12	0.13	0.12	0.12
Poor	median	0.11	0.12	0.11	0.12	0.11	0.12
Population	st. dev.	0.07	0.07	0.07	0.07	0.07	0.08
Share of	mean	0.84	0.84	0.84	0.84	0.84	0.84
White	median	0.91	0.91	0.92	0.91	0.91	0.92
Population	st. dev.	0.20	0.20	0.20	0.20	0.20	0.20
Share of	mean	0.12	0.12	0.12	0.12	0.12	0.12
Black	median	0.04	0.04	0.04	0.04	0.04	0.03
Population	st. dev.	0.20	0.21	0.20	0.20	0.20	0.21
Share of	mean	0.02	0.02	0.02	0.02	0.02	0.02
Hispanic	median	0.01	0.01	0.01	0.01	0.01	0.01
Population	st. dev.	0.03	0.03	0.03	0.03	0.03	0.03
Total	mean	4,685.39	4,111.92	4,652.24	4,291.49	4,641.78	4,523.81
Stops	median	2,343.00	2,135.00	2,353.50	2,135.00	2,343.00	2,150.00
	st. dev.	8,988.62	8,820.93	8,915.86	9,083.92	8,908.77	9,153.46
	sum	801,201	649,684	809,490	712,388	812,312	746,428
Total	mean	396.10	340.45	393.37	352.76	395.57	377.22
Searches	median	173.00	148.50	172.50	150.50	171.00	151.00
	st. dev.	749.97	744.91	743.85	755.03	747.13	766.01
	sum	67,733	53,791	68,446	58,558	69,224	62,242
Number of agencies[a]	sum	171	158	174	166	175	165
Sheriff Depts.	percentage	0.17	0.17	0.17	0.17	0.18	0.19

[a]The State Police was excluded to compare the demographic indicators.