

Rational Choice Foundations of Equal Protection in Selective Enforcement: Theory and Evidence*

Nicola Persico

1	Introduction	3
1.1	Related Literature	8
2	The Law	11
2.1	Fourth Amendment	11
2.2	Title VI.....	13
2.3	Title VI Regulations	15
2.4	Equal Protection	15
2.4.1	Supreme Court Backdrop: Armstrong and Davis/McCleskey	17
2.4.2	<i>Chavez v. Illinois State Police</i> Approach to Discriminatory Intent in Selective Enforcement	21
2.4.3	<i>Anderson v Cornejo</i> Approach to Discriminatory Intent in Selective Enforcement.....	22
2.5	Recap of the law	25
3	A Rational Choice Model of Crime and Police Prejudice in Searches	25
3.1	Illustrative Model.....	26
3.2	Hydraulic Intuition for the Workings of the Model, and Derivation of the Hit Rates Test	28
3.3	Hit Rates Test Superior to “Large Disparate Impact” Test	33
3.4	Judge Easterbrook’s “Inframarginality” Concern.....	35
3.5	Simplifying Assumptions that Can Be Relaxed	36
3.5.1	Richness of observable characteristics.....	36
3.5.2	Unobservable Characteristics (Omitted Variables).....	37
3.5.3	Number and Effectiveness of Searches.	38
3.5.4	Motorists from all groups need not carry drugs with equal probability	38
3.5.5	Quantity of Drugs Transported.....	39
4	The Maryland Case Study.....	39
4.1	The Evidence.....	40
4.2	Interpreting the Evidence.....	45
4.2.1	Validity of Data Collected by MSP	45
4.2.2	Validity of Modeling Assumptions: Does the data tell us we missed something Important in the Model?.....	46
4.2.2.1	The Competing Hypothesis: Unresponsive Motorists	46
4.2.2.2	Police Incentives to Go After Large Finds.....	50

4.2.2.3	Differential Search Intensity	51
4.2.2.4	Costless Focusing of Interdiction.....	51
4.3	Inference Possible about Bias in Searches Only, not Bias in Stops	52
4.4	Gross and Barnes: Similar Data, Different Interpretation	54
5	Selection of Appropriate Remedies.....	55
5.1	Eliminating Disparate Impact.....	56
5.2	Eliminating Intent to Discriminate	58
5.2.1	Rooting Out Intent to Discriminate Does not Guarantee Crime Reduction.....	59
5.3	Other Remedies	60
5.4	What is the Best Remedy?	61
6	Conclusion.....	62

1 Introduction

Blacks and Hispanics tend to be overrepresented in police stops and searches, and numerous lawsuits have been brought alleging racially biased law enforcement practices. In his first address to Congress, President George W. Bush said that he asked his Attorney General "to develop specific recommendations to end racial profiling. It's wrong, and we will end it in America."¹ By the term "racial profiling," he was referring to a presumed unlawful use of race or ethnicity in police interdiction. These questions have taken on a new hue in light of the events surrounding the September 11 bombings,² which have brought into relief the critical importance of efficient police work. As a result, the issue of how and when race and ethnicity can be used in interdiction has been at the front and center of the public debate.

The current state of affairs raises a basic question. In a selective enforcement case, given the information that is generally available, is it possible to tell whether the observed racial and ethnic disparities reflect police bias or whether they are the inevitable byproduct of goal-oriented enforcement? We will show that this is possible using aggregate data on policing; specifically, we will present a "bright-line" test for police bias, which is based on statistics on the percentage of successes in enforcement. This test arises from a rational choice model of crime and policing which has been developed in the economics literature. Having argued that our test is able to flag biased policing, we then ask whether the current legal approach to the use of race or ethnicity in policing is consistent with the use of our test. We find that, when properly understood, the current legal approach is not only consistent with, but indeed calls for the type of test we propose. As a "case study," we present evidence

¹ Address delivered on February 27, 2001.

² Albert W. Alschuler, *Racial Profiling and the Constitution*, 2002 U. Chi. Legal F. 163, 163 (2002)

from litigation involving the Maryland State Police. While the state police agreed to settle the pending litigation, our bright-line test shows no evidence of intent to discriminate on the part of the police.³ Finally, we ask which remedies would be most appropriate in case that impermissible behavior is ascertained.

³ NAACP v. Maryland, 72 F.Supp.2d (D.Md. 1999). The NAACP used data from the consent decree in Wilkins v.

Maryland to bring a second case, alleging that racial profiling was continuing. Maryland settled both cases.

The Constitution gives clear directions as to the elements that are to be proved in a selective interdiction suit. The theory has been codified in the McCleskey standard, which applies to all of Title VI and 14th amendment litigation, and that requires that the plaintiff prove (i) disparate impact and (ii) intent to discriminate. On how specifically to apply the two requirements in a selective enforcement case, the case law is not overly clear. The key difficulty is in the

6 See David A. Harris, Racial Profiling Revisited: "Just Common Sense" in the Fight Against Terror?, 17 American Bar Association Criminal Justice Magazine Summer 2002 Issue 2, available at <http://www.abanet.org/crimjust/cjmag/17-2/profiling.html>. ("The few successful private lawsuits regarding these practices, such as those in New Jersey and Maryland, are not nearly so numerous as those that have failed. Thus, traditional litigation strategies, although certainly a tool that should not be ignored, do not represent the best method for attacking these issues.") (internal citations removed); Samuel R. Gross & Katherine Y. Barnes, Road Work: Racial Profiling And Drug Interdiction On The Highway, 101 Mich. L. Rev 651, 741 (December 2002) ("In practice, the value of the Equal Protection Clause as a remedy for discrimination in criminal investigations is deeply compromised by the near impossibility of proof. As a result few cases are litigated and the legal doctrine remains undeveloped.")

proof of intentional discrimination: the courts are unsure about how to apply the constitutional principles to the specifics of the case. As a result, litigation is generally unsuccessful,⁶ and in the rare case that it is successful⁷, it is because of out-of-court settlements which are rooted in political pressure rather than in the law.

Faced with a general difficulty of plaintiffs to obtain relief in the courts, legal scholars have concluded that the McCleskey standards must be relaxed or reinterpreted in the case of racial profiling, so as to increase the chance of successful prosecution.⁸ These commentators would prefer the intentional discrimination requirement to be eliminated, or replaced by different requirements. Because the proposed alterations would result in major and far-reaching changes to the way that Civil Rights legislation is applied, such changes in legal standards are unlikely to be embraced by the courts. The courts, instead, have most recently reaffirmed its commitment to the existing legal standards (in *Anderson v. Cornejo*).⁹ In this paper, we will argue that the existing legal standard is perfectly applicable to racial profiling litigation, and when properly applied, it is able to identify and successfully prosecute unlawful use of race in policing.

To understand the key features of our analysis, consider a simple model with two groups of motorists. If two groups are differently disposed to carry drugs, and police search members of both groups with the same probability, then the

⁷ Gross & Barnes, *supra* note 3 at 743

⁸ See, e.g., *id.*; Steven Graines & Justin Wyatt, *The Rehnquist Court, Legal Process Theory, and McCleskey V. Kemp*, 28 *Am. J. Crim. L.* 1, 3 (2000) (“Many a Supreme-Court watcher has regarded McCleskey with the same contempt customarily accorded to the notorious cases of *Plessy v. Ferguson*.”)

⁹ *Anderson v Cornejo* 355 F3d 1021 (7th Cir 2004).

group with the greater criminal disposition (Group A) will more likely transport drugs than the other group (Group B). Thus, searches of Group A members will be more likely to yield contraband than those of Group B. In this case, we say that the success rate, or hit rate,¹⁰ is higher in Group A than in Group B. The police therefore will invest more resources in the higher probability searches of Group A, thus deterring additional criminal activity among Group A members but failing to deter Group B. This process will cause criminal activity in Group A to fall until the hit rates for both groups are equal, unless the police are biased. If the police are biased then they will search the group against which they are biased more, causing a disparity in the hit rates. Thus, the proper way to determine the existence of bias, or whether there is discriminatory intent, is to see if the hit rates are statistically different.

The insight described above provides a simple way to connect legal requirements and data, and thus yields a simple empirical test of whether the McCleskey¹¹ legal requirements are met. As such, our analysis provides a bright-line test that faithfully interprets the spirit of the current application of the McCleskey standard and dovetails with the most recent judicial approach (Anderson). We find, therefore, that the guidance that the Constitution and existing legal standards provide is eminently applicable to racial profiling litigation, and there is no reason for radical redesigns of legal standards that have been advocated by legal commentators.

Although we argue that the McCleskey standard, when properly applied, is well suited to selective enforcement cases, we do not imply that it necessarily furthers the broader social goal of crime minimization. Indeed, some

¹⁰ Samuel R. Gross and Katherine Y. Barnes, *Road Work: Racial Profiling and Drug Interdiction on the Highway*, 101

Mich L Rev 651, 667 (2002).

¹¹ *McCleskey v. Kemp*, 481 U.S. 279 (1987)

commentators have used this observation to advocate for rejecting the McCleskey standard in selective enforcement cases.¹² We doubt that such a radical step is realistic, but we observe that this end might be achieved by different means, that is, by appropriately fashioning remedies. So in Section 5 we ask whether it might be possible for the courts to achieve crime minimization by appropriate choice of remedy in selective enforcement suits. Our conclusion is that while theoretically remedies could be used to that end, in practice the type of information that would be required is unavailable to the courts. We conclude that the courts should avoid using their powers under the civil rights legislation to achieve crime minimization.

1.1 Related Literature

Much of this literature focuses on the legal standards to be used in racial profiling litigation. For reasons that we will discuss in Section 2.5 , most racial profiling lawsuits are brought under the Equal Protection Clause of the 14th amendment. The currently accepted procedure in the 14th amendment is the McCleskey standard, whereby the plaintiff must establish (a) disparate impact, and (b) intent to discriminate.

Gross and Barnes discuss a narrow vs. a broad definition of racial profiling. In brief, the narrow definition prohibits only the use of race as the sole factor in determining a candidate for a search.¹³ The broader definition of racial profiling is when race is used either solely, or in conjunction with other

¹² See, e.g., Bernard E. Harcourt, Rethinking Racial Profiling: A Critique of the Economics, Civil Liberties, and

Constitutional Literature, and of Criminal Profiling More Generally, 71 U. Chi. L. Rev 1275, 1347 (2004)

¹³ Id. at 739.

factors.¹⁴ The authors advocate the correctness of applying the broad approach and assert that using race at all is impermissible constitutionally¹⁵. However, the case law is not fully on their side. Regarding the 4th amendment, decisions go both ways, but *Whren* seems to rule the field and only requires that a police officer have probable cause to stop a vehicle even if partially motivated by racial animus. Regarding the 14th amendment the established criteria is disparate impact and intent to discriminate. Use of race is not per se illegal (see below for dispositive opinion in *Anderson*). Implementing a blanket exclusion of race as a criterion seems impractical and would overturn a large body of case law. We return to this point in Section 4.4.

Alschuler¹⁶ proposes that systemic disparate racial impact should be a violation of the fourth amendment unless it is appropriately tailored to meet a significant state interest. Alschuler couches his argument in the language of *Whren*, noting that the Court only relegated the analysis of subjective intentions to the realm of Equal Protection analysis. Alschuler's approach does not deal with intentional discrimination or individual incidents and indeed his reliance on the fourth amendment has an intuitive feel as systemic disparate impact might in fact be indicative of unreasonable searches. However, this does not seem to be the mainstream interpretation of *Whren*, which most believe has eliminated the fourth amendment for racial profiling cases.¹⁷

¹⁴ Id. at 738.

¹⁵ Gross and Barnes, *supra* note XX at 744.

¹⁶ Alschuler, *supra* note 1 at 196-96

¹⁷ One of Alschuler's suggestions is relevant to the Equal Protection realm as well. Alschuler argues that the proper baseline to determine disparate impact should be criminality rate by race, not strictly by racial proportion in the population. This is one step closer to the economic model but it leaves out consideration of detectability and deterrence and suffers from perhaps even greater information gathering problems. A prospective defendant must gather data not

Harcourt's position¹⁸ is different and innovative. He points out that the police may have no intent to discriminate and yet police behavior may be inefficient. This can happen if police behavior serves the purpose of catching criminals and not of minimizing crime. From this observation, Harcourt concludes that we should abandon the legal test of intent to discriminate, and thus the McCleskey standard used in 14th amendment litigation. Instead, Harcourt proposes for the plaintiff to be able to establish a prima facie case based on disparate impact alone, similar to the approach in Title VII litigation and *Batson v. Kentucky*, 476 U.S. 79 (1986). After that, the defendant would be able to rebut if at least two conditions are met: (i) the use of race can be shown to reduce crime over what it would be if the police did not use race; and (ii) the use of race does not result in one group of citizens being overrepresented in the prison population relative to that group's representation in the criminal population.¹⁹ In our opinion, Harcourt is right in identifying the potential concerns that police behavior may be inefficient and unbiased, at the same time. The solution, however, is much too radical in our view, in that we doubt that the courts would be willing to abandon the McCleskey standard. In Section 5 we take up the same question in a different guise, and ask whether remedies in selective enforcement cases can be fashioned in a way that takes crime minimization into account. After carefully weighing competing

only on search rates but also on background criminality. The chances of confounded data are significant and not easily overcome.

18 Bernard E. Harcourt, *Rethinking Racial Profiling: A Critique of the Economics, Civil Liberties, and Constitutional Literature, and of Criminal Profiling More Generally*, 71 U. Chi. L. Rev 1275, 1347 (2004)

19 See Harcourt, p. 1279-80. At times, Harcourt also seems to require that, for race to be used as a factor in policing, the race of a citizen must be proved to be a significant predictor of crime (see p. 1347). This would represent an additional condition beyond the two identified above.

considerations, we conclude that in practice the courts are not well placed to make judgments involving crime minimization.

2 The Law

Private Plaintiffs have attempted to bring actions against law enforcement agencies under a variety of statutory and constitutional theories. These include Title VI, Fourth Amendment and Equal Protection Claims.²⁰ We will discuss these theories in turn. After Whren, Fourth amendment claims stand little chance of success, and for the two remaining theories the legal standards for a successful claim by a private party have merged, and now courts require proof of intentional discrimination for both theories. There is also more relaxed standard whereby a federal agency can bring a Title VI claim against a local law enforcement agency and only require a showing of disparate impact.

2.1 Fourth Amendment

²⁰ Additionally, for an interesting article proposing that the Thirteenth Amendment could be a viable for source of law for a racial profiling case, see William M. Carter, Jr., A Thirteenth Amendment Framework For Combating Racial Profiling, 39 Harv. C.R.-C.L. L. Rev 17, 17-89 (Winter 2004).

²² U.S. CONST. amend. IV. ("The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.")

The Fourth Amendment²² would seem to be a reasonable starting place for a charge of racial profiling against local law enforcement. The Fourth Amendment could come into play with respect to either unreasonable stops or unreasonable searches. In 1914, the Supreme Court adopted the “exclusionary rule” stating that a defendant could not be prosecuted with material which was obtained through a search by federal actors that was in violation of the Fourth Amendment.²³ The Court later made the “commonsense” argument that this protection should be extended against State action as well²⁵. However, the Supreme Court has recently greatly restricted the use of Fourth Amendment for racial profiling cases. In *Whren v. Unites States*, the Court ruled that as long as an officer had probable cause to believe that a traffic violation was being committed, the Fourth Amendment’s prohibition against unreasonable stops was not violated even if the officer may have been motivated by racial animus.²⁶ As the vast majority of drivers are frequently speed or violate some other traffic code at any given time, this means that a police officer can stop almost any car at almost any time and fall safely within the Fourth Amendment standards.²⁷

Theoretically the Fourth Amendment could still be used for unreasonable searches. However, once a car is stopped, the police officer may permissibly search either with probable cause, e.g. seeing drugs in plain view in the car (which is clearly Constitutional from the plain language of the amendment), or by consent. The Court has ruled that as long as consent to a search is freely

23 *Weeks v. United States*, 232 U.S. 383 (1914) (ruling explicitly that this protection only reached the actions of federal officers, and not state or local actors).

25 *Mapp v. Ohio*, 367 U.S. 643, 657 (1961)

26 517 U.S. 806, 819 (1996)

27 Samuel R. Gross & Katherine Y. Barnes, *Road Work: Racial Profiling And Drug Interdiction On The Highway*, 101 *Mich. L. Rev* 651, 666 (December 2002).

given, it does not violate the Fourth Amendment.²⁸ These consent searches seem to be pervasive. (For example, Gross and Barnes, in their analysis of the Maryland data, report that greater than 95% of people will consent to a search of their vehicle, even if they are carrying drugs.²⁹ ³⁰) This means that except in cases of blatant abuse, the usefulness of the Fourth Amendment in racial profiling cases has been largely gutted.³¹

2.2 Title VI

Title VI prohibits recipients of federal funding from discriminating on the basis of race.³³ Private plaintiffs seeking injunctive or monetary relief must

28 *Schneckloth v. Bustamonte*, 412 U.S. 218, 248 (1973)

29 Gross and Barnes, *supra* note XX.

30 It is mysterious why motorists who have drugs overwhelmingly give consent to searches. Rudovsky notes that police may question the driver without issuing a Miranda warnings, and that the driver need not be told that consent is not required or that they are free to go. See David Rudovsky, *Law Enforcement by Stereotypes and Serendipity: Racial Profiling and Searches Without Cause*, 3 U. PA. J. Const. L., 296, 318 (2001). It is certainly possible that drivers don't know that they really do not have to consent to a search and believe that they are better off cooperating.

31 But see, Albert W. Alschuler, *Racial Profiling and the Constitution*, 2002 U. CHI. LEGAL F. 163, 192-207 (2002) (proposing that the Court should hold racial discrimination by the police unreasonable in violation of the Fourth Amendment).

33 42 USC §§ 2000(d)—(d)(7) (2000) (“No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance.”)

prove actions that amount to a violation of their constitutional rights equal to a violation of the Equal Protection Clause.³⁴ The exact requirements for proving a violation of Title VI will be discussed in the Equal Protection section below.

³⁴ See, e.g., *Alexander v. Sandoval*, 532 U.S. 275, 279 (2001) (stating that it is “beyond dispute” that Title VI itself only reaches intentional discrimination); *Alexander v. Choate*, 469 US 287, 293 (1985) (“Title VI itself directly reached only instances of intentional discrimination.”); *Regents of University of California v. Bakke*, 438 U.S. 265, 284-85 (1978) (noting that the Congressional Intent of Title VI was to “insure that Federal funds are spent in accordance with the Constitution”). This may be somewhat surprising when compared to the burden shifting framework of Title VII that permits a plaintiff to make a prima facie case on a showing of disparate impact in employment cases. In *Griggs*, the Supreme Court ruled that in a Title VII (42 U.S.C 2000e-2) suit, once a plaintiff has proven that an employer’s policy has a disparate impact on a racial group, the employer bears the burden of justifying the policy. *Griggs v. Duke Power Company*, 401 U.S. 424 (1971). Cf., *McDonnell Douglas Corp. v. Green*, 411 US 792 (1973). The key difference is that Title VII is based on Congress’s commerce power, whereas Title VI is based on Congress’s spending power. See, e.g. *Barnes v. Gorman*, 536 U.S. 181, 185 (2002) (noting that Title VI was enacted as a exercise of Congress’s Spending Power); *Johnson v. Transportation Agency, Santa Clara County, Cal.* 480 U.S. 616, 628 n.6 (1987) (explaining that the exercise of the commerce power allows Congress to prohibit actions by an employer greater than the Constitution, in contrast to Title VI, which can only prohibit Constitutional Violations).

³⁸ 28 C.F.R. § 42.104(b)(2) (“A recipient, in determining the type of disposition, services, financial aid, benefits, or facilities which will be provided under any such program, . . . may not utilize criteria or methods of administration which have the effect of subjecting individuals to discrimination because of their race, color, or national origin. . . .”); U.S. Department of Justice, Title VI Legal Manual, (September 1998), available at <http://www.usdoj.gov/crt/cor/coord/vimannual.htm> (last visited Aug. 25, 2005) (laying out a general overview of who is considered a Title VI recipient and what their obligations are).

2.3 Title VI Regulations

Title VI has a broader reach than prohibiting only constitutional violations, however. The Departments of Justice and Transportation have passed internal regulations conditioning receipt of federal monies by organization on a prohibition of even discrimination that only rises to disparate impact.³⁸ The Supreme Court has thus far upheld these hooks on the regulations.³⁹ However, the Court has ruled that individuals may not sue to enforce the regulations, but must sue on Title VI itself.⁴⁰ Individuals are able to ask a federal investigatory body to review a complaint.⁴¹ In summary, for a private party to bring a suit alleging violations of Title VI they must prove intentional discrimination to the same standard as Equal Protection. It bears noting however, that local law enforcement agencies which receive federal funding face being held to a disparate impact standard.

2.4 Equal Protection

39 The Court has not yet ruled directly on the issue. In *Guardians Ass'n v. Civil Service Com'n of City of New York*, five justices in at least alternate theories agree that the regulations should be upheld. 463 U.S. 582 (1983). In *Alexander v. Sandoval*, the Court assumed for purposes of deciding, but without actually ruling, that the regulations were valid. 532 U.S. 275 (2001).

40 *Alexander v. Sandoval*, 532 U.S. 275, 293 (2001)

41 See Department of Justice: Civil Rights Division: Coordination and Review Section, *Investigation Procedures Manual for the Investigation and Resolution of Complaints Alleging Violations of Title VI and Other Nondiscrimination Statutes*, (September 1998) available at: <http://www.usdoj.gov/crt/cor/Pubs/manuals/receive> (last visited Aug. 25, 2005) (laying out the procedures for an investigating agency to take when a complaint is received).

Whren significantly closed the door on racial profiling suits by pointing to the Equal Protection Clause of the Fourteenth Amendment as the proper clause.⁴² There are two basic types of racial discrimination cases that could fall under Equal protection, facial and as applied. If a local law enforcement agency admitted openly that they use race as a factor a plaintiff could bring a suit challenging the policy on its face.⁴³ In this paper we will not focus on this type of case. The second, and more common type of case involves claims of systematic racial animus in policing efforts that are not a part of an official policy, but real nonetheless. Here the plaintiff does not allege that a policy is facially discriminatory, but rather that the policy, as applied, results in discrimination. This is much more difficult to identify and deal with and is the subject of this paper.

42 517 US 806, 813 (1996) (“We of course agree with petitioners that the Constitution prohibits selective enforcement of the law based on considerations such as race. But the constitutional basis for objecting to intentionally discriminatory application of laws is the Equal Protection Clause, not the Fourth Amendment.”).

43 What is means to use race as a factor is an open question. Gross and Barnes distinguish between a narrow vs. a broad definition. The authors note that some local law enforcement agencies and courts only consider use of race to be improper when police conduct is based exclusively on race; others consider any use of race in policing to be unlawful (Gross and Barnes, *supra* note XX at 738.) To bring the issue into relief, consider the following thought experiment. In Yakuza culture, those who have failed to carry out their duty in some way may be forced to atone by (voluntarily or not) by slicing off parts of their pinkies. A police officer might use the combination of Japanese ancestry and a missing finger as a proxy for gang membership. This would violate the broad definition of racial profiling, but it would not violate the narrow definition. The narrow definition would only be violated if the police officer used Japanese ancestry itself as a proxy for gang membership.

In theory either or these uses of race might violate the Equal Protection Clause and because race was used explicitly, the local agency’s policy would be subject to strict scrutiny (see, e.g., *Grutter v. Bollinger* 539 U.S. 306, 326 (2003) noting that all use of racial classifications are held up a standard of strict scrutiny). The agency would thus have to demonstrate that that program was narrowly tailored to a compelling government interest. *Id.* This analysis is complicated, however. by Supreme Court rulings on searches at border crossings. The Court upheld the use of appearance of Mexican ancestry as a factor in determining secondary inspections at border crossings (*United States v. Martinez-Fuerte*, 428 U.S. 543 (1976); *United States v. Brignoni-Ponce*, 422 U.S. 873 (1975)). Some courts have used these cases to allow evidence obtained in a search when race was a factor in the search criteria (Gross and Barnes, *supra* note XXX, 744).

51 *McCleskey v. Kemp*, 481 U.S. 279 (1987) (internal quotes removed).

2.4.1 Supreme Court Backdrop: Armstrong and Davis/McCleskey

The Court in McCleskey noted that when a party alleges an Equal Protection violation, he has the “burden of proving the existence of purposeful discrimination” and that “the purposeful discrimination had a discriminatory effect on him.”⁵¹ The Court was echoing its sentiment from Washington v. Davis, where the Court held the Equal Protection Clause only reaches intentional discrimination with respect to statutes that are facially neutral.⁵² In other words, a plaintiff must prove discriminatory effect and intent to discriminate.⁵³ Disparate impact alone,⁵⁴ a showing that a policy or procedure

52 Washington v. Davis, 428 U.S. 229, 241 (1976) (“A statute, otherwise neutral on its face, must not be applied so as invidiously to discriminate on the basis of race.”).

53 See, e.g., Chavez v Illinois State Police (“To show a violation of the Equal Protection Clause, plaintiffs must prove that the defendants' actions had a discriminatory effect and were motivated by a discriminatory purpose.”) 251 F.3d 612, 635-36 (7th Cir 2001); Flowers v. Fiore 239 F.Supp.2d 173 D.R.I.,2003. (“In order to prevail on this claim, Flowers must present evidence that he was treated differently from similarly situated white motorists and that the action taken against him was motivated, at least in part, by his race”; Bradley v. United States 299 F.3d 197 C.A.3 (N.J.),2002 (“To make an equal protection claim in the profiling context, Bradley was required to prove that the actions of customs officials (1) had a discriminatory effect and (2) were motivated by a discriminatory purpose“; Christman v. Kick 342 F.Supp.2d 82 D.Conn.,2004 (In order to establish a violation of equal protection based on selective enforcement, the plaintiff must ordinarily show (1) the person, compared with others similarly situated, was selectively treated; and (2) that such selective treatment was based on impermissible considerations,” such as race, religion, intent to inhibit or punish the exercise of constitutional rights, or malicious or bad faith intent to injure a person”).

54 A note on language conventions: The courts often interchangeably use disparate impact and discriminatory impact. For the purposes of statistical analysis they mean essentially the same thing. However, discriminatory impact can also include the effect of discrimination on one particular individual, rather than a group as a whole. For this paper we are more focused on group effects, and thus often use the term “disparate impact” to describe a disparity in treatment

impacts one racial group disproportionately, is not enough itself to prove a constitutional violation of Equal Protection Rights.⁵⁵

In order to prove discriminatory effect, the plaintiff must show “that they are members of a protected class, that they are otherwise similarly situated to members of the unprotected class, and that plaintiffs were treated differently from members of the unprotected class.”⁵⁶ Proving discriminatory effect can either be accomplished by naming such a similarly placed individual who was not subjected to scrutiny or by the use of statistics to prove disparate impact.⁵⁷ In other words, plaintiffs can either point to a specific incident that demonstrates that they were singled out as a result of their race, or they can attempt to prove a larger scale, disparate impact on their race.

More challenging than proving discriminatory effect is proving intent to discriminate. Discriminatory intent requires that plaintiff prove that the actions of the law enforcement agency were motivated by racial animus “at

across racial groups. Also, discriminatory purpose is generally synonymous with intent to discriminate. We will use the intent language.

55 *Adm'r of Mass. v. Feeney*, 442 U.S. 256, 272-74, (1979) (Disparate impact, even when the effects are predictable is not an Equal protection violation without intent). *Washington v. Davis* 426 U.S. 229, 238 (1976) (“But our cases have not embraced the proposition that a law or other official act, without regard to whether it reflects a racially discriminatory purpose, is unconstitutional solely because it has a racially disproportionate impact”).

56 *Chavez*, 251 F.3d. at 636.

57 See, e.g., *Hazelwood School Dist. v. United States*, 433 U.S. 299, 310 (1977) (ruling that the government had met its burden by showing overwhelming statistical evidence to prove disparate impact in an employment setting).

59 *McCleskey*, 481 U.S. 279, 298 (1987); *Christman v. Kick*, 342 F.Supp.2d 82, 93-4 (2004) (“Disparate treatment by itself, not resulting from an impermissible consideration or malicious or bad faith intent to injure, is an insufficient basis for an equal protection claim.”).

least in part because of . . . its adverse effects upon an identifiable group.”⁵⁹ McCleskey v. Kemp is considered to be highly restrictive, as the Court stated that a plaintiff must show that “the decisionmakers in *his* case acted with discriminatory purpose”.⁶⁰ The Court’s emphasis on “his” indicates that proof of intent to discriminate must be present in each case brought under the Equal Protection Clause of the Fourteenth Amendment (a causation requirement). Not only must systemic discrimination be proven, but also a plaintiff must prove discrimination in her particular case.

Additionally, even if the plaintiff has met her burden, the defense still has a chance to rebut. The Court has recognized that in some cases, the State’s actions arise from a number of different motivations. “Rarely can it be said that a legislature or administrative body operating under a broad mandate made a decision motivated solely by a single concern, or even that a particular purpose was the ‘dominant’ or ‘primary’.”⁶¹ The plaintiff does not need to show that the State took its action solely because of the discriminatory purpose, only that it was part of the motivation.⁶² The defendant can still attempt to rebut by showing by the preponderance of the evidence that they would have done the

⁶⁰ McCleskey v. Kemp, 481 U.S. 279, 292 (1987). See, e.g., Gross and Barnes, *supra* note xx at 723, (“The holding in McCleskey has been widely criticized, and rightly so. As Justice Brennan points out in dissent, there was no real doubt that race did influence capital sentencing in Georgia; everybody who dealt with the issue in practice knew it and acted on that knowledge. The Court denies the obvious.”)

⁶¹ Arlington Heights v. Metropolitan Housing Development Corp. 429 U.S. 252, 265 (1977).

⁶² *Id.*

same thing absent the discriminatory purpose.⁶³ The McCleskey Court did not even reach the rebuttal stage.⁶⁴

The Supreme Court has not directly dealt with allegations of racial profiling as a violation of Equal Protection, but most cases follow familiar pattern in the district courts, largely patterned after a Supreme Court case involving selective prosecution, United States v. Armstrong⁶⁵. The Armstrong Court ruled that the “requirements for a selective-prosecution claim draw on ordinary equal protection standards. The claimant must demonstrate that the federal prosecutorial policy had a discriminatory effect and that it was motivated by a discriminatory purpose”.⁶⁶ In the absence of a Supreme Court ruling directly on point, the lower courts have differed in their approach to proving intent to discriminate.

63 Id. at 271, n.21 (“Proof that the decision by the Village was motivated in part by a racially discriminatory purpose would not necessarily have required invalidation of the challenged decision. Such proof would, however, have shifted to the Village the burden of establishing that the same decision would have resulted even had the impermissible purpose not been considered. If this were established, the complaining party in a case of this kind no longer fairly could attribute the injury complained of to improper consideration of a discriminatory purpose. In such circumstances, there would be no justification for judicial interference with the challenged decision.”). See also *Mt. Healthy City School Dist. Bd. of Educ. v. Doyle*, 429 U.S. 274, 286 (1977) (explaining that this is basically a causation test, if the plaintiff was not actually injured by defendants, then she cannot complain).

64 *McCleskey*, 481 U.S. 296-97 (“Moreover, absent far stronger proof, it is unnecessary to seek such a rebuttal, because a legitimate and unchallenged explanation for the decision is apparent from the record: *McCleskey* committed an act for which the United States Constitution and Georgia laws permit imposition of the death penalty.”)

65 517 U.S. 456 (1996)

66 Id. at 465 (internals citations and quotes removed).

2.4.2 *Chavez v. Illinois State Police* Approach to Discriminatory Intent in Selective Enforcement

In a Seventh Circuit case, *Chavez*, Hispanic plaintiffs attempted to prove a claim of racial discrimination with respect to highway stops in Illinois. Plaintiffs were able to prove discriminatory effect (*supra*) but the court rejected their attempts to prove discriminatory intent through both statistical and non-statistical evidence. Their non-statistical evidence included the officer stating “that one can never tell with “you people”, which was considered evidence of racial animus, but was not considered itself enough evidence to prove a constitutional violation.⁶⁷ The statistical data included stops, searches and contraband and includes the driver’s race.⁶⁸ Plaintiff’s experts testified that the data showed dramatically higher search rate for African-American and Hispanic motorist. ⁶⁹ *Chavez* noted that statistics about discriminatory impact alone are generally not enough to prove a constitutional violation.⁷⁰ *Chavez*

⁶⁷ *Chavez*, 251 F3d at 625.

⁶⁸ *Id.*

⁶⁹ See, e.g., *Hazelwood School Dist. v. U.S.* 433 U.S. 299 (1977) (ruling “for purpose of determining whether a prima facie case of discrimination was made, percentage of black teachers in the district should have been compared to the percentage of blacks in the school teacher population in the relevant labor market rather than with the percentage of black students in the district”); *Anderson v. Cornejo*, 355 F.3d 1021 (7th Cir. 2004) (holding that aggregate statistics at the national level are not dispositive of local animus). *Chavez* also presented an interesting piece of evidence that overlaps the statistical and non-statistical realms. The officer recorded *Peso Chavez*’s ethnicity as white, rather than Hispanic. *Chavez*, 251 F3d at 624. This certainly raises the implication that Illinois officers may be tweaking their statistics by misreporting data.

⁷⁰ *Chavez*, 251 F3d at 647 (“Only in “rare cases [has] a statistical pattern of discriminatory impact demonstrated a constitutional violation”); *Yick Wo v. Hopkins*, 118 U.S. 356, 374 (1886) (reversing criminal charges as discriminatory for violation of a local ordinance that required a license to operate a laundry in a wooden building when data showed that Permits for laundries granted to 79/80 whites but 0/200 Chinese in similar situations.); *Hunter v. Underwood*,

uses a categorical exclusion model, modeled after McClesky,⁷¹ noting the rare occasions that the Supreme Court has allowed statistics to prove intentional discrimination and ruling that this is not one of them. In so noting, the court was referring to statistics that are more closely analogous to search rates rather than hit rates, which we argue below are not informative about intent to discriminate.⁷² Thus, the court in Chavez correctly (in our view) rejected an attempt to prove intentional discrimination by means of statistics that were probative of disparate impact alone.

2.4.3 *Anderson v Cornejo* Approach to Discriminatory Intent in Selective Enforcement

Anderson v. Cornejo also a Seventh Circuit case, although in front of a completely different panel, takes a seemingly different approach, in that *Anderson* actually uses statistics to prove a lack of discriminatory intent. Crucially, these are statistics on hit rates, not search rates as was the case in *Chavez*. This is consistent with the theory developed in Section 3 below, as well as with our interpretation of the intent of the *Chavez* court in limiting the use of “statistics.”

471 U.S. 222, (1985) (affirming that a finding that a disenfranchisement of blacks at a rate of 1.7 times more than whites constituted disparate impact); *International Broth. of Teamsters v. U.S.* 431 U.S. 324 (1977).

⁷¹ 481 U.S. at 294 note.12.

⁷² For example the statistics in *Yick Wo* and *Teamsters* involve binary data aggregated to ask how many individuals in their racial group receive a certain treatment. This is closely analogous to asking how many motorists of a given race were stopped. Hit rate, by contrast, measures how successful police are at identifying drugs. The hit rate equivalent in a *Yick Wo* would be if the government went door to door looking for people with qualifications to run laundries and then compiled statistics that determined fitness grouped by race. Presumably the government would be motivated by finding laundry sites, and therefore you would expect to find equivalent hit rates among racial groups.

In *Anderson*,⁷³ a group of African-American women challenged the constitutionality of customs searches performed on them, pointing to a General Accounting Office (“GAO”) study that African-American women were by far the most likely racial group to be searched at O’Hare Airport in Chicago.⁷⁴ The study showed that x-ray search rates for black women were almost twelve times those of white men and more than eight times those of white women.⁷⁵ Judge Easterbrook held that “[t]hese and similar data from the GAO’s report do not support any constitutional claim . . . for at least four reasons.”⁷⁶ The most relevant to our model, and to future cases, is the fourth objection,⁷⁷ made on the grounds that “these statistics show disparate impact, not disparate treatment (intent to discriminate), and the equal protection guarantee is concerned only with the latter.”⁷⁸

The study also showed hit rates of 27.6 percent for black women, 25.1 percent for white men, 19.5 percent for white women, 61.6 percent for black

73 355 F3d 1021 (7th Cir 2004).

74 *Id* at 1023.

75 United States General Accounting Office, U.S. Customs Service: Better Targeting of Airline Passengers for Personal Searches Could Produce Better Results, GAO/GGD-00-38, 13 (March 2003), available at <http://www.gao.gov/new.items/gg00038.pdf> (visited February 22, 2005). We do not pass judgment on the methodology used in this specific study, but rather focus on how the court viewed the importance of search rates and hit rates in determining whether the plaintiffs were intentionally discriminated against.

76 *Anderson*, 355 F3d at 1023.

77 The first three objections relate to the quality of the data collected in that specific study, *id* at 1023-24, and are not of broader relevance.

78 *Id* at 1024 (emphasis added) (explaining that *Personnel Administrator of Massachusetts v Feeney*, 442 US 256 (1979), requires a showing of disparate treatment to demonstrate a violation of the equal protection clause).

80 *Anderson*, 355 F3d at 1023.

men, 58.8 percent for Hispanic men, and 45.7 percent for Hispanic women.⁸⁰ Judge Easterbrook noted that these data imply that “black women seem to have been treated similarly to both white men and white women.”⁸¹ He held that “the success rate of strip searches . . . show that Customs officials search black women with (on average) the same degree of suspicion that leads them to search white women or white men.”⁸²

The court’s analysis is consistent with our model. First, Judge Easterbrook declined to use search rates to infer intent to discriminate.⁸³ Second, he deduced the absence of disparate treatment between different groups from roughly equal hit rates. Indeed, our analysis (detailed below) further suggests that although African-American women received no disparate treatment relative to white men and white women, all three groups received disparate unfavorable treatment relative to black men and Hispanic men and Hispanic women.⁸⁴

81 Id. (emphasis added).

82 Id. at 1024-25. Judge Easterbrook later added that “[i]f about 0.1% of black women returning from foreign travel are smuggling, and the agents select so carefully that 28% of those searched are caught with contraband, where’s the beef?” Id. at 1025. However, a high success rate for one group is not relevant to an equal protection claim.

83 Other courts have also taken the approach that statistical analyses examining search rates establish only disparate impact, and not different treatment. See, e.g., Chavez, 251 F3d 612 at 641 (questioning the sufficiency of statistics that would show that minority “motorists are stopped at a significantly higher rate than are white motorists”); Hurn v United States, 221 F Supp 2d 493, 501 (D NJ 2002) (finding that general statistics demonstrating higher search rates for African American women did not demonstrate specific discriminatory intent in the case at hand).

84 The court did not find discrimination for two additional reasons. The first is simply that the methodological problems with the statistical evidence were sufficient for Judge Easterbrook to refuse to find discrimination based on it. Anderson, 355 F3d at 1023-24. The second, which he discusses briefly, is the inframarginality problem caused by the relative ease with which different groups can be caught. Thus, equal marginal success rates, which are not measured, will lead to different average success rates, which are measured. Id at 1025. In Section 3.4 we address this concern and show that, as a matter of pure logic, it is not internally consistent. In any case, even though the court did

2.5 Recap of the law

The law on racial profiling litigation is modeled on the McCleskey standard. A plaintiff must prove disparate impact and intent to discriminate. In practice, disparate impact is more easily proved than intent to discriminate. In Chavez and in Anderson, the courts have declined to infer intent from “extreme disparate impact.” In Anderson, the courts have looked at statistics on the success rate of searches as informative regarding intent to discriminate.

In what follows we develop a rational choice model of policing in which statistics on search rates are informative about disparate impact, and statistics on the success rates of searches are informative about intent to discriminate.

3 A Rational Choice Model of Crime and Police Prejudice in Searches

The model is adapted from the rational choice model of policing originally developed by Knowles, Persico, and Todd (“KPT”), and generalized in Persico and Todd (2005).⁸⁵ The goal of the modeling exercise is to provide a simple test that would detect the presence of a discriminatory intent in searches based on statistical evidence of police behavior. The test must be able to distinguish search disparities reflecting discriminatory intent from those that might be generated in the bona fide pursuit of crime. The analysis will show that intent to discriminate in searches can be deduced by focusing on hit rates. An

not find discrimination on this basis, it still noted that “[t]he GAO recommended that the Customs Service increase the number of searches in high-success-rate categories.” *Id.*

85 Knowles et al, 109 J. Pol. Econ. 203; Persico and Todd, Papers and Proceedings of the American Economic Association, 2005.

appendix⁸⁶ obtains the same result in the context of a much more general model.

3.1 Illustrative Model

There are many motorists, each of whom makes a dichotomous choice: whether or not to carry drugs. These motorists are outwardly distinguishable only by race, A and W, but they are inwardly heterogeneous in their propensity to carry drugs: some motorists (the majority, perhaps) are not intent on carrying drugs; others may be willing to carry if their likelihood of being searched does not exceed some threshold.⁸⁷

There are many police officers, and each of them chooses which type of motorists to search, A or W. All searches have the same intensity, and each officer has exactly 1 search to allocate. We assume that all police officers seek to find drugs, and we allow for the possibility that police have intent to discriminate.

We assume that an unbiased police officer will choose whom to search in the pursuit of successful searches only; thus, unbiased officers will focus their searches on whichever group presents a high likelihood of success. A biased police, on the other hand, will also take some pleasure from the mere act of searching A's; it is in this way that we translate the 14th amendment's language that "[d]iscriminatory purpose' . . . implies more than intent as volition or intent as awareness of consequences. It implies that the decisionmaker, in this case a state legislature, selected or reaffirmed a

86 Available online at <http://www.econ.upenn.edu/~persico/research/Papers/papers.htm>

87 In a more accurate and realistic interpretation of the model, motorists are not divided into categories by race, but rather by a characteristic that is correlated with race. We will return to this point momentarily.

particular course of action at least in part ‘because of,’ not merely ‘in spite of,’ its adverse effects upon an identifiable group.”⁸⁸ Operationally, our definition means that the biased officer will favor searching A’s even when this group presents a somewhat lower likelihood of success relative to W’s. The extent to which a biased officer is willing to trade off a lower likelihood of success against the pleasure of searching A’s measures the intensity of the officer’s bias.⁸⁹

A comment is in order. The reader might argue that, in using race to choose whom to search, the police are behaving impermissibly even if they are not motivated by bias. This point has some force, but only in the stark model described where for simplicity of exposition the labels “African American” and “White” are the only observable characteristic. In a more accurate and realistic interpretation of the model, two groups of motorists are never identical except for race. Typically, the two groups will differ along many characteristics---dress code, body ornaments, vehicle accessories, demeanor---some of which the police will use in choosing whom to search, and some of which are correlated with race. So, instead of “African American” and “White,” a more realistic labeling of the variables in our model should be “individuals who share a characteristic (or set of characteristics) that is more frequently found in the African American (resp., White) population.” When the model is properly understood, then, the police is not behaving impermissibly by conditioning its search decision on group identity, because the groups are not identified by race but rather by a set of characteristics correlated with race.⁹⁰

88 *McCleskey v. Kemp*, 481 U.S. 279, 298 (quoting *Personnel Administrator of Massachusetts v. Feeney*, 442 U.S. 256, 279 (1979)).

89 Besides lowering the cost of searching A’s, bias could also be modeled as increasing the (psychic) benefit to the officer of finding a guilty member of group A. This is immaterial for our analysis. Our conclusions are unchanged if we admit this additional channel for bias.

90 We will return to this point in Section 4.4.

The reader familiar with Game Theory will have noticed that our simple model is in fact a game. In what follows, we shall explore the workings of this simple model; more formally, we shall characterize the Nash Equilibrium of this game. Our goal in so doing will be to develop an intuition for the features of the data that we should be looking for if we are checking for police bias. It is important to note that this model is particularly simplified---it assumes away many realistic details that may result in “frictions,” that is, they may change the workings of the model. The frictions we choose to ignore may, in some cases, affect the conclusions we seek to draw from the model. In other cases, these conclusions may be robust. We will comment later on the import of certain key simplifying assumptions.

3.2 Hydraulic Intuition for the Workings of the Model, and Derivation of the Hit Rates Test

To obtain some intuition for the workings of this model, it may be helpful to introduce an analogy between searching motorists for drugs and sampling from an urn. Let’s imagine that a racial group corresponds to an urn, which is filled with red balls representing drug-carrying motorists, and green balls representing honest motorists. The fraction of red to green balls in an urn represents the crime rate in a group. The unbiased police officer’s search decision is analogous to the problem of a statistician who must choose which of the two urns to draw a ball from in hopes of drawing a red ball. If the proportions of red to green balls are as depicted in Figure 1, the statistician will choose to draw from the W urn, because the proportion of red balls is greater in that urn.

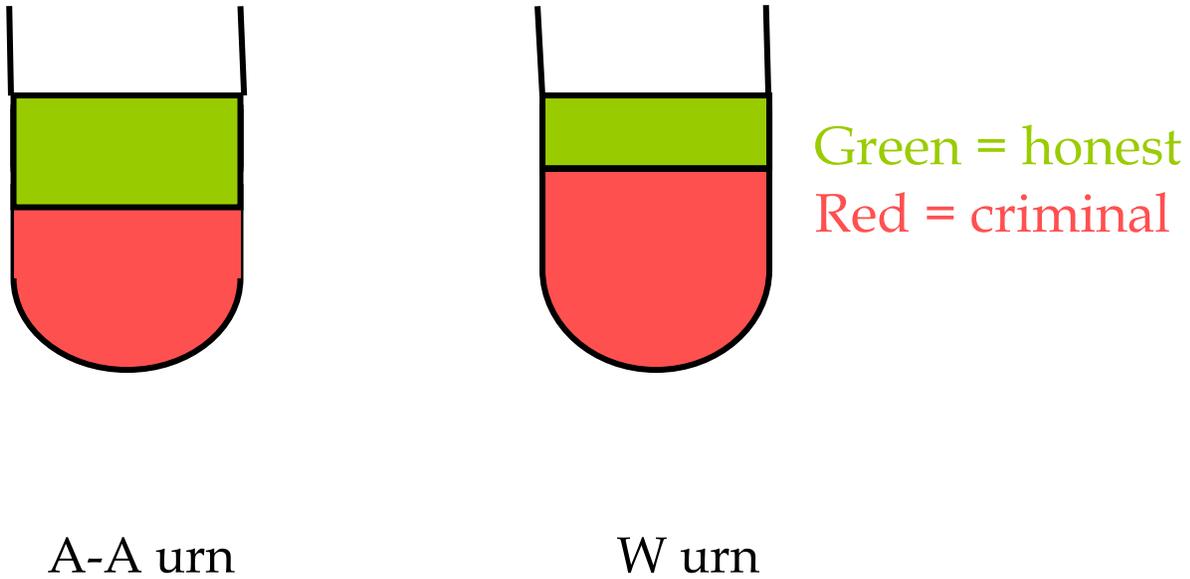


Figure 1: The Urn Model

Our immediate goal is to predict what type of police and motorist behavior is “stable” in this environment. By stable, we mean that no motorists or police officer wants to change their behavior. Technically, we are looking for the Nash Equilibrium of this game.⁹¹

To this end, let’s start by imagining that all police officers only searched from the W urn. This is the situation depicted in Figure 2, where the thick arrow on top of urn W indicates that all police interdiction is focused over that group (there is no arrow over urn A-A). Given that the police choose this particular interdiction pattern, motorists in group A will be less deterred from carrying drugs relative to those in group W. We assume that, because of this lopsided interdiction pattern, proportionally more motorists choose to carry drugs in

⁹¹ Formally, a game is in a Nash equilibrium when no player of the game must find it profitable to change his/her action given the behavior of the other players. Nash equilibrium is the standard solution concept used in game theory to predict how a game will be played by rational agents.

group A than in group W. This is reflected in Figure 2, where the proportion of red balls is greater in urn A-A than in urn W.

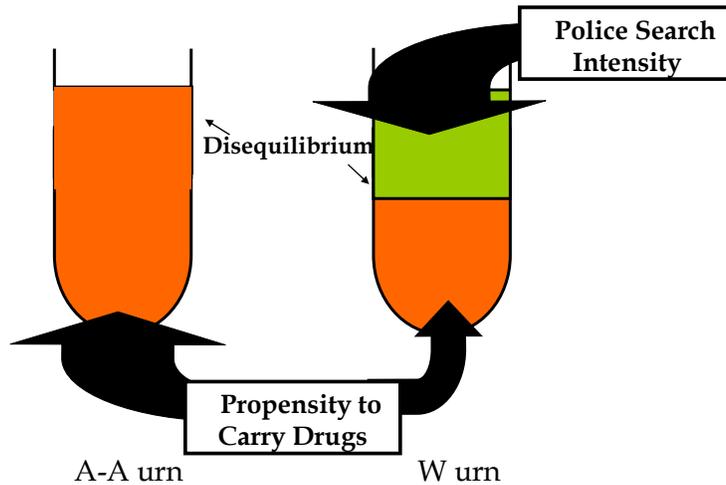


Figure 2: Disequilibrium

Is the situation depicted in Figure 2 a Nash equilibrium? No, because it is not stable: the police officers want to change their search behavior. Indeed because just like sampling from urn A maximizes the chances of drawing a red ball, the police officers who is motivated to find drugs will want to switch to searching members of group A.

Let us then follow the natural adjustment process, and let us examine whether shifting some search intensity towards group A, as depicted in Figures 3a,b, puts the game in equilibrium. As the police progressively shift to searching more members of group A, the difference in crime rate between group A and W falls. (This is because A motorists realize that they are being searched with increasing intensity, so fewer will carry drugs). Thus, the proportion of red

balls is closer to equal across urns in Figure 3a than in Figure 2, and it becomes perfectly equal in Figure 3b.

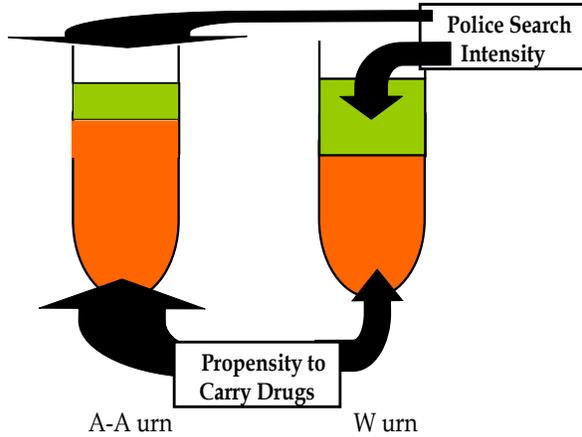


Figure 3a: Towards Equilibrium

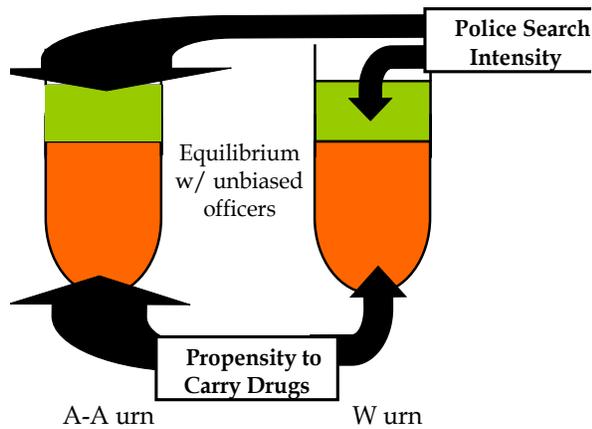


Figure 3b: Equilibrium with Unbiased Police

Since in Figure 3b the two urns have exactly the same proportion of red and green balls, a statistician who maximized the chance of drawing a red ball would be indifferent as to which urn to sample from. Similarly, at the level of police search intensity represented by the thick arrows in Figure 3b, a police officer has exactly the same chance of finding drugs when searching from either group. Thus, an unbiased police officer is indifferent between searching

A or W motorists. This is therefore a stable situation: the game is at a Nash equilibrium.

Observe that, for the game to be in Nash equilibrium, it is necessary for the probability of finding drugs to be equal in the two groups. For, if that probability were lower in one of the two groups, then unbiased police officers would want to switch away from searching that group, and so we would not be in equilibrium.

Consider now the case in which all police officers are to some degree biased, in the sense that they derive some additional pleasure from searching (or from finding drugs on) a member of the A group. This means that, if the two groups had the same crime rate, all officers would strictly prefer to search group A motorists over group W motorists. This means that the situation depicted in Figure 3b is no longer stable: instead, biased officers would continue switching to searching group A motorists until the crime rate in group A is sufficiently lower than in group W so that the police bias is exactly balanced. When the police are biased against A motorists, then, Nash equilibrium is achieved only when the crime rate among A motorists is lower than among W motorists.

This simple discussion has identified a powerful conclusion: if a group is discriminated against, then within our model police searching its members should be less likely to find drugs. The “hit rates” of police searches should be lower. If, conversely, police are equally likely to find drugs upon searching a member of either group---the hit rates are equalized---then disparities in search behavior, while possible,⁹² are not the result of police bias. This is a

⁹² In light of our model, search disparities (in some direction) are almost inevitable in equilibrium; it is very unlikely that the police, be they biased or not, would search both groups with the same intensity. The only way that this could happen is if the difference in the average propensities to carry drugs between the two groups is exactly offset by the police bias.

remarkable finding because we have been able to relate the existence of police bias, which is very difficult and subjective to document, with disparities in the success rate (“hit rates”) of police searches, which are relatively straightforward to document.

3.3 Hit Rates Test Superior to “Large Disparate Impact” Test

In an effort to identify intent to discriminate, the courts often accept a test that relies on disparate impact. In the context of racial profiling, this test is based on search rates. In this test for discrimination, one would first need to determine which motorist characteristics can correctly be used by the police in choosing whom to search. Race would, signally, not be among those “valid” characteristics. Then, one would estimate the statistical relationship that links the motorists’ “valid” characteristics to the probability that they are searched. If adding the characteristic “Race” to these set of variables helps improve the model’s predictive power, then the conclusion would follow that the police are using race to decide whom to search. If race is found to have a large predictive power, in particular, then the test is taken to “suggest” the presence of intent to discriminate.

This test has several major shortcomings as a test for intent to discriminate. First, it is sensitive to the so-called “omitted variables” bias. If we fail to include in our statistical model a variable that is legitimately being used by the police in deciding whom to search, and if the omitted variable is correlated with race, then the variable “Race” will improve the predictive power of our (misspecified) statistical model. This is not, however, evidence that the police are using race as a factor in deciding whom to stop. It only means that the police are using a variable that is correlated with race and that is missing from our statistical specification. Since we are unlikely to have all (or even many) of the

characteristics that the police can legitimately use to decide whom to search, the omitted variable bias is most likely a serious concern.

The second major shortcoming of the conventional test is that it is not clear which motorist characteristics are to be deemed “valid.” For example, the police might argue that a certain very specific dress style is a suspicious indicator of gang activity. This may be true. But, if the dress style is highly correlated with race, it may also be true that the police use the dress style as an excuse to cover up racially biased policing. If we were to admit “specific dress style” among the conditioning variables in our statistical model, it might be that subsequently adding the characteristic “Race” would not significantly improve the statistical model’s predictive power. In this case, the conventional test would fail to detect a discriminatory intent when in fact it is present.

Finally, the fact that police may appear to be using race as a factor to decide whom to stop does not prove intent to discriminate. To see this, let’s go back to the simple example in Figure 3a. Let’s imagine that the arrow on top of urn A is bigger than that on top of urn W, meaning that A’s are searched more intensely than W’s. Nevertheless, in that picture searches of A’s are more likely to result in finds than searches of W’s. If we observed the pattern depicted in this example, then, we would conclude that the police are forgoing more profitable searches of A’s in favor of less profitable searches of W’s. In other words, the police would be displaying bias against W’s even as they are searching A’s more intensely than W’s.

These difficulties illustrate, in a nutshell, the fact that the conventional “search rate” test is not able to detect intent to discriminate. At best, it is a test of disparate impact. In contrast, the “hit rate” test we advocate is immune from all these problems: it is robust to omitted variables (See Section 3.5 above), it does not require us to determine which characteristics are “valid,” and most crucially, it is a proper test of “intent to discriminate.”

3.4 Judge Easterbrook's "Inframarginality" Concern

In Anderson, Judge Easterbrook seems to embrace the hit rates logic that we presented above. Indeed, he uses the similarity in hit rates between black women and other groups to conclude that “black women seem to have been treated similarly to both white men and white women.” However, the opinion in Anderson apparently stops short of fully embracing the hit rates logic because of a concern for inframarginality. The inframarginality argument is developed, if a little obscurely, at page 1025. The concern is that if there are two types of searches, one more intrusive than the other, then differences in hit rates might arise even though Customs officers are unbiased. The opinion suggests that, if low intensity searches catch most of the black men with drugs but few of the black women with drugs, then the more intrusive searches will necessarily show a lower success rate among black men than among black women, without this disparity being evidence of officers bias. However, this argument is flawed. Indeed, if intrusive searches resulted in a lower success rate on black men, then unbiased officers would choose to reduce the number of black men that are subject to intrusive search, and redirect their intrusive searches towards the population that is more likely to yield a find, namely black women. This process will, of course, lead towards the equalization of the hit rates in the two groups, unless for some reason officers are biased against one of the groups. We conclude that the hit rates test is applicable even when there are different types of searches, provided of course that hit rates are compared across groups within the same type of search. The “inframarginality” concern raised in Anderson is thus logically flawed. If the court had recognized the logical fallacy of the inframarginality concern, we suggest that the court might have fully embraced the hit rates test, as we suggest it should have.

3.5 Simplifying Assumptions that Can Be Relaxed

The hit rates test for police bias remains valid in a much more general environment than the one we used to illustrate it. We now list several dimensions along which the model can be generalized, and explain why the hit rates test is preserved. An appendix⁹³ presents a formal derivation of the hit rates test in the fully general model.

3.5.1 Richness of observable characteristics.

We have assumed that the police can only distinguish two groups in the population, A and W. In reality, the police surely can observe more than just race. So, in reality the police will classify individuals in groups characterized by a combination of characteristics: race, sex, age, length of hair, etc. In terms of the model, this means that instead of being able to draw from two urns only, the police is able to pick whom to search from many urns. An urn would represent, for example, the group of white female motorists in their thirties with short hair. In this more sophisticated many-urns model, the same logic applies that we developed in Section 3.2: if the police are unbiased, the hit rate in every urn will be equalized in equilibrium. If not, then the police would switch from searching urns with low hit rates to those with high hit rates. By the same token, if the police are biased against African Americans, we should observe lower hit rates on all urns that contain African-Americans. Thus, the hit rate tests is still valid to detect police bias.

⁹³ Available online at <http://www.econ.upenn.edu/~persico/research/Papers/papers.htm>

3.5.2 Unobservable Characteristics (Omitted Variables).

Characteristics that are unobservable both to the police and to the researcher obviously do not pose a problem for our analysis: they can be folded into the “average propensity to commit a crime within a group.” What might potentially pose a problem are characteristics that are observable to the police but not to the researcher. Suppose for example that the police thought that not being the owner of the vehicle was an indicator that the motorist is more likely to carry drugs. If vehicle ownership is not recorded in our data, then we have a problem of omitted variable. This omission is especially relevant if one believes that in reality an experienced officer uses a variety of subtle clues to decide whether to search a motorist.

Fortunately, our analysis is robust to the presence of omitted variables, and the hit rates test continues to be a valid test of racial bias. To see why, let us think of how we would represent omitted variables within our model. Suppose, for example, we know the race of the driver but we do not know whether the driver owned the car. Then, in the parlance of our model, we are unable to distinguish between the “white driver, own vehicle” urn and the “white driver, third-party vehicle” urn. So, one can think of omitted variable (car ownership, in this case) as the confounding of two or more urns. This possibility is potentially troublesome because the police are able to distinguish the two urns and may treat them differently. The police might, for example, search third-party vehicles more intensely than other cars, and we will never know that because we can’t tell which vehicles are third-party. Fortunately, in equilibrium the hit rate from the third-party and own-vehicle urns ought to be the same (otherwise, the police would switch to searching more motorists from one of the two white urns). Then, under the null hypothesis that the police are unbiased, being unable to distinguish the two urns is not a problem, since their hit rate is the same. It is therefore valid to compare the A and W hit rates even if we

realize that, in so doing, we are aggregating hit rates of vehicle owners and non-owners.

3.5.3 Number and Effectiveness of Searches.

It is not important that each police officer can only search once. The analysis generalizes straightforwardly to the case of multiple searches per officer. Perhaps less obviously, the hit rates test remains valid even if officers are less able to detect the presence of drugs on one race. Say, for example, that officers were on average less able to detect drugs on A's than on W's, owing perhaps to their lack of familiarity with the cultural cues among the A's. *Ceteris paribus*, this will decrease the hit rate of a search conducted on a member of the A group. The abnormally low hit rates on the A group will drive officers to search fewer A's and more W's until the hit rate equality is restored between the two groups.⁹⁴ If the police are unbiased, this hit rate equalization will represent an equilibrium point. So, again, we find that equal hit rates are associated with unbiased police---the hit rates test remains valid.

3.5.4 Motorists from all groups need not carry drugs with equal probability

In our simple model there were only two groups, and so motorists from the two groups carried drugs at the same rate in equilibrium. In a more general incarnation of that model there will be urns corresponding to groups of motorists, such as old ladies, who are likely to exhibit a low propensity to carry drugs. In equilibrium, these groups will need to be searched very lightly, if at all, to bring their crime rates in line with that of more criminally inclined

⁹⁴ Hit rates equalization will require the crime rate to be higher among the A's given the primitives in this thought experiment.

groups. If their propensity to carry drugs is low enough, such groups will not need to be searched at all, and still their crime rate will be below that of the groups that are searched. So, our model does not predict that all groups on the road should have the same crime rate. It only predicts that those groups that are searched---a small minority of all motorists---should have the same crime rate (if the police is unbiased).

3.5.5 Quantity of Drugs Transported.

We have assumed that the motorists' decision is dichotomous---whether or not to carry drugs. In reality, motorists choose how much drugs to carry. In terms of the theoretical development of the model, this generalization does not pose particular problems. We can think of the decision to commit a crime as a continuous variable (how much drug to carry) and we can let motorist choose their preferred quantity taking into account the probability of being searched as well as the penalty schedule that is associated with being discovered with different quantities of drugs. If we assume that the police still maximizes the probability of a "hit," and so that the police receive the same benefit (in terms of psychic and other benefits) from big and small hits, then in equilibrium hit rates must be equalized across groups if and only if the police has no bias. Thus, the hit rates test remains valid. If the police benefit from finding drugs is increasing in the size of the find, then the hit rates test must be amended. We will return to this topic in Section 4.2.2.

4 The Maryland Case Study

To illustrate how the methodology would be applied in a court case, we report the results of statistical analyses carried out on data collected by the police as a result of the settlement agreement in Wilkins v Maryland State

Police (Maryland I)⁹⁵ The settlement entailed the payment of money to the plaintiff, the formulation of a statement by the police renouncing racial profiling, and the collection of the data presented here. Maryland likely would have lost the litigation, as plaintiffs were able to uncover a “smoking gun” document that detailed a plan to stop minority motorists.⁹⁶ The NAACP used this data in a follow up case, alleging that Maryland was still discriminating (Maryland II).⁹⁷ The tables reported in this section are based on the analysis originally performed in KPT, which is based on data from 1590 vehicular searches performed between January 1995 and 1999 on a stretch of I-95 in Maryland. We will also report on the findings of subsequent analyses performed by Gross and Barnes⁹⁸ and by Barnes⁹⁹ that use larger and more up-to-date data sets on Maryland State Police stops and searches, and discuss the interpretation that they offer for their findings.

4.1 The Evidence

A first look at the data reveals a familiar pattern of disparate impact. Table 1 reveals that, of those searched, 63% were African-Americans,

95 *Wilkins v Maryland State Police*, Civil Case No CCB-93-468 (D. Md. 1993).

96 Sean P. Trende, *Why Modest Proposals Offer The Best Solution For Combating Racial Profiling*, 50 *Duke L.J.* 331, 338 (2000) (“[T]he Maryland State Police had, only a week earlier, published a memo asking police officers to be on the lookout for drug couriers, who would be “predominantly black males and black females.”) (quoting David Cole, *No Equal Justice: Race and Class in the American Criminal Justice System* 36 (1999)).

97 *NAACP v. Maryland*, 72 *F.Supp.2d* (D. Md. 1999).

98 Gross and Barnes.

99 Katherine Y. Barnes, *Assessing the Counterfactual: the Efficacy of Drug Interdiction Absent Racial Profiling*, Manuscript 54 *Duke L. J.* 1089 (2005).

unquestionably a much higher percentage than the fraction of black motorists on the road.¹⁰⁰ We also note that men are disproportionately more likely to be searched than women (93% of those searched are men). Other features of the data, while of lesser interest because not directly related to protected categories, are still instructive about police behavior. For example, older vehicles represent 22% of all searches, luxury vehicles 8%, 3rd party vehicles 18%, and 31% of searches were made at night. The variable “Guilty” refers to the fraction of searches that resulted in a find of any drugs or paraphernalia. Many searches resulted in marijuana finds (23%) and 8% resulted in cocaine finds.

¹⁰⁰ The number in parentheses represent standard deviations.

**Summary Statistics
(All Observations)**

Black	0.63 (0.01)
White	0.29 (0.01)
Hispanic	0.06 (0.01)
Female	0.07 (0.01)
Guilty	0.33 (0.01)
Cocaine	0.08 (0.01)
Marijuana	0.23 (0.01)
Crack Cocaine	0.04 (0.005)
Heroin	0.02 (0.003)
Morphine	0.001 (.001)
Other Drugs	0.01 (0.002)
Paraphernalia	0.01 (0.002)
Night (12am-6am)	0.43 (0.01)
Number of Observations	1582

Table 1: Summary Statistics

The first two rows in Table 2 report the hit rates by race. Of the white motorists searched, 32% were found with some illegal drugs. Of the African-American motorists searched, 34% were found with some illegal drugs. Despite the wide disparity in search rates, hit rates are very close. In fact, a Pearson

chi-square test cannot reject the hypothesis that the two hit rates are the same.¹⁰¹ According to our analysis, this suggests that the police have no intent to discriminate against African-Americans.¹⁰² To see why, let us apply the logic of our model to this case. Presumably, police officers are aware that, at the current status quo, the hit rates are approximately the same from searching the two groups (whites and African Americans). Then, a biased officer would know that switching some of her searches to African Americans would allow her to indulge in her prejudice at no cost in terms of success rate. Thus, if officers were biased, they would all switch from the status quo towards searching more African Americans, and this process would continue until the African American hit rate were sufficiently lower than the white hit rate that the bias was offset. But in the data we can't reject the hypothesis that the hit rate is the same for whites and African Americans. Therefore, we can't reject the hypothesis that the police is not biased.

101 The p-value is 0.33.

102 For Hispanics, however, the hit rate (not reported) is lower, possibly indicating a bias against Hispanics. An alternative explanation for the Hispanics finding could be that Hispanics are more likely to be "mules," i.e., to transport high-value shipments of drugs not for personal use. To address this issue, one can perform more sophisticated, though somewhat more subjective, analyses based on the quantities of drugs found. See Knowles et al, 109 J Pol Econ 203 (cited in note 3).

Proportion of Motorists Found to be Carrying Drugs		
By Race	African American	0.34
	White	0.32
	Hispanic	0.11
By Sex	Male	0.32
	Female	0.36
By Time	Day	0.32
	Night	0.33
By Car Characteristic	Luxury	0.25
	Non-Luxury	0.33
	Older	0.32
	Newer	0.33
	Third Party	0.29
	Own	0.33

Table 2: Hit Rates

The next two rows ask the same question with regards to sex. Again, despite the wide disparity in search rates, hit rates are very close and the chi-square test does not reject the hypothesis that the two hit rates are the same.¹⁰³ According to our analysis, this suggests that the police have no intent

¹⁰³ The p-value is 0.37.

to discriminate against women.¹⁰⁴ The remaining rows compare hit rates along other dimensions. Hit rates are similar across these dimensions, too.¹⁰⁵

4.2 Interpreting the Evidence

4.2.1 Validity of Data Collected by MSP

In our analysis, we have assumed that the MSP faithfully recorded all instances of searches and all finds. We must acknowledge the possibility that the police might misreport or underreport certain incidents. If, for example, the police decided to let go a young motorist who was found with a minimal amount of marijuana, the police might not record the search, or might record the outcome as “no drugs found” when in reality some drugs were found. If the police behaved in this fashion, and if these underreportings were skewed by race, then the hit rate computed from our data would not accurately reflect the true hit rate. We have no direct way of disproving this possibility. We note, however, that the police knew that the data collected was mandated by court order, so we suspect that police officers would not take the risk significantly to tamper with the data. Moreover, even though the police knew their behavior was under scrutiny, that scrutiny did not focus on the hit rate. Hence, the police did not have a direct incentive to tweak their statistics on hit rates. We

104 For Hispanics, however, the hit rate (not reported) is lower, possibly indicating a bias against Hispanics. An alternative explanation for the Hispanics finding could be that Hispanics are more likely to be “mules,” i.e., to transport high-value shipments of drugs not for personal use. To address this issue, one can perform more sophisticated, though somewhat more subjective, analyses based on the quantities of drugs found. See Knowles et al, Racial Bias in Motor Vehicle Searches: Theory and Evidence, 109 J. Pol. Econ. 203 (2001).

105 The only time that the Pearson test rejects equality of hit rates is for luxury cars, which are somewhat less likely to result in a find. According to our model, this indicates some preference on the part of the police for searching luxury cars. It is interesting, however, that the disparity disappears when we focus on “large” drug finds. Id.

think, therefore, that the data we received represent a reasonably accurate picture of police behavior in the period under scrutiny.

Because the data only portray police behavior during a period in which the force was under judicial scrutiny, the data cannot speak to the behavior prior to 1995, a period during which the allegations of impermissible use of race were made.

On the positive side, our data are detailed enough that they allow us to exclude non-discretionary searches, such as for example searches incident to arrests for unrelated crimes.¹⁰⁶

4.2.2 Validity of Modeling Assumptions: Does the data tell us we missed something Important in the Model?

In this section we address the validity of some maintained assumptions that are important for the validity of the hit rates test. We do this by looking at the data for indications of the validity (or lack thereof) of the assumptions.

4.2.2.1 The Competing Hypothesis: Unresponsive Motorists

One of the most important assumptions in our model is that motorists react to the probability of being searched by carrying drugs less often. Whether, or perhaps more accurately, the degree to which this assumption is valid is an

¹⁰⁶ When non-discretionary searches cannot be eliminated from the data, it is necessary to resort to sophisticated econometric techniques to provide bounds for the hit rates that would have been observed had the researcher been able to exclude non-discretionary searches. See J. Knowles and R. Hernandez-Murillo, Racial Profiling or Racist Policing?: Testing in Aggregated Data, *International Economic Review* (2003).

empirical question.¹⁰⁷ The fact that hit rates are equalized in our data suggests to me that our model of police behavior accurately approximates the actual behavior of the Maryland State Police, at least the part of it that is represented in our data. It is not direct proof, however.

Indeed, one might look at the same data and draw a very different conclusion. One might infer that the African-American and white motorist populations have the same underlying rate of drug carrying, that this rate is independent of police search intensity, and that the police are indulging their bias in searching African American motorists more intensely than white motorists.¹⁰⁸ This argument is lent some superficial plausibility by the observation that, in the US population at large, there is research to suggest that black and white citizens use drugs in the same proportion.¹⁰⁹

107 For an empirical analysis that assumes no response from the motorists, see Katherine Barnes, *supra* note XX.

108 An important part of this argument is that motorists do not significantly adjust their drug-carrying behavior in response to police interdiction, for otherwise it would be difficult to explain why the hit rates are equalized even as African Americans are searched more intensely than whites. Under the assumption that motorists do not respond to the probability of being searched, however, equalization becomes plausible as a reflection of an underlying similarity in drug-carrying.

109 Rates of current illicit drug use varied significantly among the major racial/ethnic groups in 2004. The rate was highest among persons reporting two or more races (13.3 percent) and American Indians or Alaska Natives (12.3 percent). Rates were 8.1 percent for whites, 7.2 percent for Hispanics, and 8.7 percent for blacks. See Substance Abuse and Mental Health Services Administration. (2005). Results from the 2004 National Survey on Drug Use and Health: National Findings (Office of Applied Studies, NSDUH Series H-28, DHHS Publication No. SMA 05-4062). Rockville, MD, p. 18.

The argument seems not entirely plausible for several reasons. First, we should not use knowledge of the general population to draw inference about the populations that are searched, and vice versa. Drivers who are searched are a highly selected group and they are surely unrepresentative of the entire black and white populations. (For one thing, the general population surely does not carry drugs in their car at the rate of 33%.)¹¹⁰ When we do focus on the populations searched in the Maryland data, we do detect some differences between black and white drivers who are searched in terms of quantity of drugs carried (blacks having bigger quantities on average), suggesting that there are differences in the propensity to carry drugs between the groups of motorists who are searched.

The second reason why the argument is not plausible is that we know that many motorists who carry drugs go to great lengths to conceal it in special compartments of the car. This indicates that motorists are alert to the chance of being searched and are willing to incur considerable effort to reduce the chance of being found with drugs. It stands to reason that there must be a significant fraction of motorists who respond to this risk by not carrying drugs, where they would be carrying otherwise.

¹¹⁰ Further evidence against using data from the general US population can be gleaned by comparing women and men. As in 2002 and 2003, in the general US population males were more likely in 2004 to report current illicit drug use than females (9.9 vs. 6.1 percent, respectively). Males were almost twice as likely to use marijuana as females (8.0 vs. 4.3 percent). See SAMHSA report, p. 17. Yet, in the Maryland data, the likelihood of finding drugs on women who were searched was the same as for men. This is strong evidence in favor of the proposition that we should not use knowledge of behavior in the general population to make inference about the population searched. In connection with this point, it is worth pointing out that the theoretical model we propose does not imply that all groups of drivers on the road carry drugs with probability 33%. The model only requires equality of drug-carrying rates for the groups that are searched in equilibrium. Those groups likely represents a very small percentage of all drivers. The model's only implication for non-searched groups is that their drug-carrying rate be no greater than that of the groups who are selected for search.

In further defense of our model, it must be pointed out that our analysis does not require that the responsiveness of motorists be large. If the response of motorists is sluggish, then it will take a lot of difference in search intensities to equalize hit rates across groups. One may surmise that the wide disparity in search rates that is associated, in our data, with hit rates equalization, may reflect a relatively sluggish response of motorists.

In a sense, it is remarkable that such a simple, perhaps even stark prediction as hit rates equalization, was verified in the data, particularly in light of the wide disparity in search rates. Indeed, we can think of several frictions that we could have introduced that might have interfered with the prediction that an unbiased police force will hit rate equalization.¹¹¹ The fact that hit rate equalization appears to be so pervasive in the Maryland data suggests that, in that instance and for that time period, these frictions did not play a significant role. When applying our analysis to other data sets, however, we can expect

111 Imperfect police updating. Can Racially Unbiased Police Perpetuate Long-Run Discrimination? by Helle Bunzel, Philippe Marcoul, ISU Economics Working Paper #03004, March 5, 2003 More traditional analysis: limited response of motorists Dhammika Dharmapala and Stephen L. Ross, Dharmapala and Ross (2004) Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence Contributions to Economic Analysis & Policy Volume 3, Issue 1 2004 Article 12, Antonovics and Knight (2005) A new look at racial profiling: evidence from the Boston Police Department K Antonovics, BG Knight, MP Page – manuscript, UCSD, 2004, Vani K. Borooah “Racial bias in police stops and searches: an economic analysis,” European Journal of Political Economy Vol. 17 2001, pp. 17–37, David Bjerck “Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate,” manuscript, McMaster University, 2004. Special mention to Anwar and Fang (forthcoming AER)

that the neat equalization of hit rates we observe may not be so pervasive.¹¹² In those cases, it may be necessary to augment the simple model we presented to account for these frictions.

4.2.2.2 Police Incentives to Go After Large Finds

Another reason why the pervasive hit rates equalization is so remarkable has to do with the question of police incentives in relation to the size of the hit. It seems reasonable that the police should receive greater benefits from finding greater quantities of drugs. When we amend along those lines the model of Section 3.1, what we should look for to test bias are not hit rates, but rather some other function of drug finds that takes into account the quantity of drugs recovered. If, for example, we knew that the police received a benefit that was exactly proportional to the weight of marijuana found in a car, we would look for equalization of the expected weight (in grams) of the marijuana found in searches of African-Americans and whites. Unfortunately, in practice we do not know exactly how the size of the find factors in the reward felt by the officer. But the Maryland data allow us to make some preliminary inquiries into this question. When we restrict attention to more significant drug finds, and consider “hits” only those searches that exceed a certain magnitude, we find that the hit rate is higher for African Americans than it is for whites.¹¹³ So, if the police only considered a search successful when it yields large quantities of drugs, then our data would suggest that the Maryland police behavior is biased, if at all, in favor of African Americans. This finding does not seem to me

112 There are some cases in which hit rates equalization is maintained (see "Using Hit Rates to Test for Racial Bias in Law Enforcement: Vehicle Searches in Wichita," with Petra Todd).

113 See KPT, Table 2. As the definition of a hit is made more stringent, so that only very large drug finds are counted as hits, the disparity between the hit rates of African Americans and whites becomes statistically significant (see KPT Table 3).

the most plausible interpretation of the data. Instead, it seems to me that a simpler interpretation of the Maryland data is that the police are actually equalizing hits, defined as any drug recovered. This simple interpretation is also consistent with the equalization of hit rates across other (non-race) dimensions.

4.2.2.3 Differential Search Intensity

Another aspect our model does not do justice to is the question of differential intensity in searches. If the police search one group more intensely than another, comparing hit rates may be less meaningful than our analysis might suggest. In reality, we do not know if there is a systematic correlation between the intensity of searches and race, so it is difficult to gauge how restrictive it is to assume, as we do, that every search has the same intensity. This area must be left for future research.

4.2.2.4 Costless Focusing of Interdiction

Finally, in our model we have assumed that the police can costlessly direct their interdiction power toward any subgroup of the motoring population. This assumption implied that any subgroup whose “hit rate” was higher than the others could, and would, be subject to such intense focus of police interdiction that members of that group would decrease their criminal activity until the “hit rate” in that group was brought back to the population average. In reality, the police might find it difficult, or impractical, to direct a lot of interdiction power on a subgroup, especially if that subgroup is small. Imagine, for example, that a rare subgroup of the population is highly likely to be involved in crime; to fix

ideas, let's imagine these are first-generation Sicilians.¹¹⁴ Imagine that there is only one officer stationed on the road. This officer can stop many drivers (it does not take long to stop someone), but can only search a few of them (a search takes a long time). Moreover, while the officer is searching one car, all other drivers who are passing by cannot be searched. Now, it will be extremely costly for the police to devote himself solely to the task of searching first-generation Sicilians, because that would mean that most of the times the officer will not search anyone, waiting for the off chance of meeting a first-generation Sicilian. Instead, the practical thing to do is for the officer to search a broader variety of motorists. But this means that first-generation Sicilians cannot be searched with probability 1, since as they are driving by, the officer might well be occupied with searching some other motorist. This reasoning shows that it may be difficult for the police to focus a lot of interdiction power on narrowly defined subgroups. We do not know how empirically relevant this potential consideration might be. The fact that hit rates are so well equalized in the Maryland data suggests that, at least in that case, this potential concern might not be very relevant empirically.

4.3 Inference Possible about Bias in Searches Only, not Bias in Stops

The equalization of hit rates is exactly what our model would predict when the police are unbiased and are pursuing a search strategy that maximizes the probability of a successful search. As an important corollary, the evidence suggests that the disparity in search rates between African-Americans and whites does not reflect intent to discriminate on the part of the police.

¹¹⁴ Full disclosure: the author is a first-generation Sicilian.

It must be emphasized that the test we carried out on the Maryland data is a test of discrimination in searches, not in stops. The simple reason is that we have not modeled the officer's decision to stop a driver separately from the search decision. Since in our model the officer could choose freely whom to search, it is as if the entire population had been stopped and stood ready to be searched. In reality, of course, the entire population is not stopped, and so one might be worried that our analysis may have missed some important element that might invalidate the hit rates test. In particular, the concern would be that because in reality the officers are restricted to search only motorists who were stopped, that the pool of stopped motorists would be so small as to restrict the ability of officers to direct their searches at will, and that this friction might somehow invalidate the hit rates test.

While this is a reasonable concern, in the Maryland data that does not seem a particularly relevant one. First, again, the widespread equalization of hit rates we found strongly suggests that if the police were constrained in having to search only motorists who were stopped, it did not seem to have a discernible impact. Second, and consistent with the previous point, Barnes (2004) reports that only a minuscule fraction of motorist stopped were ever searched (less than 0.5%), suggesting that the constraint to only search stopped motorists was not a particularly binding one.

Nevertheless, it is important to point out that our analysis of the Maryland data does not address the question of whether, or how, our conclusions would generalize to discrimination in stops.¹¹⁵

115 If one wanted to look more closely at the issue of discrimination in stops, the first step would be to investigate the link between stops and searches. One might worry that since the population that is subject to search is only the motorists who were stopped and not the entire population, the hit rates analysis on search outcomes might be subject to a selection bias. However, it is important to realize that the hit rates test would continue to be indicative of police bias even if it is applied to the restricted population of those who were stopped.

4.4 Gross and Barnes: Similar Data, Different Interpretation

Gross and Barnes as well as Barnes analyze data obtained from the Maryland State Police. Using larger data sets,¹¹⁶ they report similar findings concerning the relative similarity in hit rates across races.¹¹⁷ Barnes, in particular, reports the same striking absence of disparity in hit rates reported in our data.¹¹⁸ Gross and Barnes, however, come to the conclusion that the behavior of the Maryland State Police behavior is not likely to be found permissible by the courts.¹¹⁹ In our reading of their argument, these authors do not argue that police behavior suggests bias. In fact, these authors point out that the police are motivated by the desire to find drugs.¹²⁰ They argue, however, that in pursuit of drugs, the police choose whom to stop based on race.¹²¹ They argue that relying on race in search decisions is impermissible per se.¹²² In other words, Gross and Barnes argue that it is impermissible for the police to rely on race as one of the elements in establishing whom to search. But this statement quickly leads to paradox. For example, it would be perfectly reasonable for the police to ignore indicators such as tattoos and/or missing finger falanxes in the general population, and yet to view them as highly suspicious indicators when

116 Their data sets subsume that used in KPT.

117 See Table 8 in Gross and Barnes 2002.

118 See Table 4. We note that the hit rates in Barnes (2004) are computed as a fraction of stops and not, as in KPT, of searches. Moreover, Table 4 restricts attention to hit of “hard drugs.”

119 See Gross and Barnes *supra*, at pages 738 and 744. It must be pointed out that Gross and Barnes (2002) could not foresee the decision in *Anderson v. Cornejo*.

120 Cite. See Gross and Barnes *supra*, page 753.

121 Cite. *Id.*, page 753 (again).

122 *Id.* at 744 (“To summarize, it is plainly unconstitutional to use race as a criterion for choosing who to stop or search.”).

exhibited by Japanese males, for whom these signs are indicative of belonging to the Japanese Mafia (Yakuza). According to Gross and Barnes's argument, the police would be violating the law if it stopped individuals with missing fingers and tattoos with higher probability when they are Japanese men. We can find, however, no evidence that such police behavior might be impermissible

In the Yakuza example just discussed, the police are using the variable race to aid their inference about the likelihood that a citizen might be involved in criminal activity. A somewhat different situation might result if the police are not using race in their inference process, and yet race may be correlated with other variables used by the police in deciding whom to search. If that is the case, members of one race may end up being searched more often than others because they possess a characteristic that the police use to decide whom to search. Gross and Barnes do not consider this possibility. In our view, however, this is not implausible: we know of no evidence speaking to this issue. In the absence of evidence on this specific point, and in the presence of evidence that the police are unbiased, we think the prima facie argument for intent to discriminate cannot be successfully established. As we suggested, we believe instead that the argument must be made based on the disparity in hit rates, if any.

5 Selection of Appropriate Remedies

This paper has outlined a procedure to meet the requirements set out in McCleskey. In order to prove impermissible use of race in searches, the plaintiff must first show evidence of disparate impact, which can be done by showing that a protected category was searched disproportionately to their representation in the population. The plaintiff must also show evidence of

intent to discriminate on the part of the police; as we have argued in this paper, this can be done by comparing the hit rates across protected categories. Finally, to show evidence that intent to discriminate had an effect on the plaintiff, the plaintiff can point to the disparity in hit rates of the officer that searched him.

If the plaintiff is able to show these three requirements, then the defendant will, according to the McCleskey standard, be able to rebut the presumption of illegal use of race by showing exculpatory evidence, or by adducing a compelling state interest. It is doubtful whether combating illegal drugs might be considered a compelling state interest, whereas perhaps defense against terrorism might be.

What remedial action should be implemented should the defendant be found guilty? This is a difficult question. Intuitively, the remedy should be to bring the defendant's policy in line with the legal standards, which means eliminating either the disparate impact of the intent to discriminate. The difficulty is that neither of these remedies has any special appeal from the point of view of crime reduction, in the sense that neither eliminating disparate impact nor eliminating intent to discriminate guarantees that crime is minimized. We now discuss these two remedies in turn.

5.1 Eliminating Disparate Impact

Members of a protected group are disparately impacted by a policy if “they are otherwise similarly situated to members of the unprotected class, and that plaintiffs were treated differently from members of the unprotected class.”¹²³ The difficulty with redressing disparate impact is that the

¹²³ Chavez , 251 F.3d at 636.

comparison group is not well specified. The case law talks about “similarly situated,” but it is not clear exactly what that means. At one extreme, a very broad definition would require similarly situated to mean for example “everyone else in another racial group.” So, for example, young minority males should be “similarly situated” to old ladies and treated on average in the same way. At the opposite extreme, one could take a very narrow view of the concept of “similarly situated.” This has been done sometime in litigation, when the ratio of minority motorists among those searched by the police has been compared to the fraction of minorities among the prison population. One might speculate that “similarly situated” would mean “similar in their propensity to commit a crime.” This interpretation, however, is not favored by everyone since it would require making judgments about the (unobserved) propensity of certain subgroups to commit crimes. Clearly, which criterion is used to define “similarly situated” matters a great deal. In Section 3.3 we discussed some of the difficulties involved in making this determination. To drive the point home once more, recall the Yakuza example. Someone who is missing a pinky, has tattoos, and is of Japanese origin, may be considered likely to be a member of Yakuza. It would be misleading, though, to consider this individual as similarly situated to those who miss a pinky but are not of Japanese origin, or to those who have tattoos but do not miss a pinky, etc.

Besides being difficult to implement, requiring that the police treat equally every group and may entail substantial loss in effectiveness of interdiction, as police manpower will be tied up searching the great majority of the population which at current is not searched.

Another order of problems is a long-run effect on the distribution of crime. Suppose we force the police to search fewer African Americans and more whites in an effort to redress disparate impact. Then, under our model the crime rate would increase among African Americans, at least in the long run. If the crime in question affects mostly those of the same racial groups as the perpetrators,

then by directing the police resources towards the whites we may be allowing criminals to encroach among the black population, which may ultimately damage that same population.¹²⁴ In pursuing a more equal impact of policing, then, we might want to be mindful about the distribution of crime across different racial groups.

5.2 Eliminating Intent to Discriminate

Another possible remedy could be to hold the police to an equalization of the hit rates across protected categories. This would replicate the outcome when the police is not biased against any group. As such, this remedy can be thought of as eliminating the intent to discriminate. This policy has the advantage of being relatively easy to implement and monitor.

It could be argued, however, that the policy does treat similarly situated groups equally, in the sense that two groups who had exactly the same propensity to commit a crime would be treated equally. Many people would view this policy as failing to address the issue of disparate impact, however, because equalizing hit rates may result in very disparate impact across broad racial and ethnic categories (this is precisely the situation in the Maryland data, where the police equalize hit rates and yet African Americans are searched disproportionately often relative to whites).

Moreover, even if we find merit in eliminating intent to discriminate, there remains the fact that eliminating intent to discriminate need not reduce crime,

¹²⁴ Whether this consideration applies to the crime of drug trafficking is an open question, though it is probably fair to say that the fabric of ethnic neighborhoods is destroyed by drug traffic. On the other hand, if the crime in question is possession of drugs for personal use, then the social effects of that crime on non-criminals are relatively less salient.

and in fact theoretically may increase it, as we will show next. Thus, eliminating intent to discriminate might theoretically come at a cost in terms of crime fighting.

5.2.1 Rooting Out Intent to Discriminate Does not Guarantee Crime Reduction

To understand why rooting out intent to discriminate may theoretically increase crime, go back to Figure 3b, which represents the equilibrium when the police is unbiased (and so there is no intent to discriminate at all). Suppose we forced the police to reduce the search rate from the A-A urn and to increase that from the W urn (thus nudging the police towards behaving in a manner that is consistent with intent to discriminate against W's). According to the model, the crime rate will increase in the A-A urn and decrease in the W urn. The net effect on the crime rate as a whole will depend on the responsiveness of the two urns to this change in search intensity. If the A-A urn has a relatively sluggish response, the increase in crime in that urn will be smaller than the decrease in crime experienced in the W urn. In that case, forcing the police to behave in a discriminatory way against W's reduces the overall crime rate.¹²⁵ Conversely, in this setup nudging the police towards removing intent to discriminate may well result in an increase in the overall crime rate.

¹²⁵ This point was made in Persico (2002), Persico and Todd (2005) "Passenger Profiling, Imperfect Screening, and Airport Security, Forthcoming, The American Economic Association Papers and Proceedings (2005); Harcourt (2004); Racial Profiling by Michael Alexeev* and Jim Leitzel, Manuscript, University of Chicago, 2001; OPTIMAL SEARCH PROFILING WITH LINEAR DETERRENCE, Charles F. Manski, forthcoming on AEA Paper and Proceedings, May 2005.

This simple observation shows that intent to discriminate is not, in our model, necessarily associated with inefficient interdiction. Once we realize this, we may then question how meaningful is the standard of “no intent to discriminate.” That is, once we realize that “no intent to discriminate” is not necessarily a good proxy for “efficient behavior,” we may be led to question the logic of the remedy that the police eliminate intent to discriminate.¹²⁶

5.3 Other Remedies

Besides rooting out disparate impact and intent to discriminate, which would remove one of the two necessary elements to establish the prima facie case, a third possible type of remedy might focus on the rebuttal phase of the McCleskey standard, i.e., proving that they would have engaged in the same behavior absent the discriminatory purpose. For police practices, they would presumably have to prove that the methods were selected to reduce crime. Thus, a remedy might direct the police to minimize the aggregate crime rate (aggregated across all categories), perhaps under the constraint that disparate impact not be too large across protected subgroups.

A remedy that explicitly took into account crime reduction faces the practical problem that we do not know when that condition is verified. Remember, in the model aggregate crime may increase or decrease when the police shift towards a more racially fair interdiction: that depends on the relative elasticity of crime

¹²⁶ This question is the starting point of Harcourt (2004). Harcourt normatively suggests that we, as a society, should move to an outcome that minimizes aggregate crime rate subject to the constraint that the ratio of black to white inmates be equal to the ratio of black to white criminals in the population at large. See also On the Use of Racial Profiling as a Law Enforcement Tool Author(s): Helle Bunzel, Philippe Marcoul, ISU Economics Working Paper #05021, July 14, 2005

in the two populations. In practice, we are unlikely to be able to measure these elasticities, and so we would not know when the condition is verified under which we would abandon the requirement of hit rate equalization.¹²⁷

Another order of problems arises in connection to the constitutionality of directing police interdiction towards one or the other subgroup of the population. Even if we could measure the relative elasticities in the two groups, and we found that the conditions are met for forcing the police to increase interdiction on whites and reduce it on African Americans, inducing the police to do so will presumably require explicit race-based incentive schemes, such as greater rewards for finding drugs on white motorists. This would trigger a strict scrutiny analysis.¹²⁸ Alternatively, police officers might be given racial quotas requiring them to search at least x percent of whites. Such schemes would, however, be unconstitutional on their face because they explicitly take race into consideration.

A third complicating factor is the long-run effect on the racial and ethnic distribution of victimization, which was discussed in Section 5.1.

5.4 What is the Best Remedy?

127 But see the efforts by Jeff Dominitz and John Knowles, *Crime Minimization and Racial Bias: What Can We Learn From Police Search Data?*, PIER Working Paper 05-019, Manski (op. cit).

128 See, e.g., *Johnson v. California*, 543 U.S. 499 (2005) ("We have held that "all racial classifications [imposed by government] ... must be analyzed by a reviewing court under strict scrutiny." Under strict scrutiny, the government has the burden of proving that racial classifications "are narrowly tailored measures that further compelling governmental interests.")

It seems to us that holding police behavior to the standard of “similar impact on similarly situated” is difficult, due to the difficulty of establishing a practical definition of “similarly situated,” and raises concerns regarding the efficient allocation of resources as well as regarding the distribution of victimization across racial groups.

Directing the police to implement sophisticated interdiction strategies to minimize aggregate crime by trading off crime rates across subgroups of the population is impractical, since it is highly doubtful whether the police have enough information about the relative elasticities of crime to policing among subgroups of the population for them to be able to meaningfully pursue this goal. Even more so, it is hopeless for the courts to monitor whether the police is complying with their orders.

This leaves on the table mandating the equalization of hit rates. This remedy has the advantage of being simple to implement and monitor. It could also be argued that it treats similarly situated groups equally, although large disparities are possible in equilibrium across groups that are not similarly situated. A disadvantage of this remedy is that it does not ensure that crime is minimized in the aggregate. On the other hand, crime minimization is difficult for the police to attain and even more difficult for the courts to monitor. So, on balance our recommendation is that mandating equalization of hit rates is the preferable remedy among the ones considered here.

6 Conclusion

In this paper we have provided a foundation for the current legal approach to selective enforcement cases. We have presented a bright-line test for police bias based on statistical data on the success rate of interdictions. The test finds

bias against a protected group if the success rate of interdiction against that group is *lower* than on another group. Within a rational choice model of crime and policing, we have shown that this “bright-line” test flags police behavior if and only if it is driven by bias against a protected category. This test therefore operationalizes the concept of “intent to discriminate” which is central to the legal approach to selective enforcement.

Our bright-line test could have been applied to the Maryland cases to help prove or disprove intent to discriminate. It is especially relevant in Maryland II, as plaintiffs were armed with direct evidence of discriminatory purpose in Maryland I. In Maryland II, plaintiff would clearly be able to show disparate impact, but using hit rate data would have indicated that there was no intentional discrimination on the part of the Maryland police department. The result would be that the Maryland police would be able to demonstrate that, at least under the McCleskey standard, they were not discriminating.

Other commentators have argued that the current legal approach to selective enforcement is inadequate, particularly because of its reliance on the concept of “intent to discriminate,” and must be refashioned. We show, instead, that the current legal approach is well suited to accommodate the best statistical test available to detect bias (within our model). We view this finding as a strong rationale to preserve the current legal approach.

We agree with some commentators that the current legal approach does not promote the larger social goal of crime minimization. We argued, however, that the courts are not well placed to pursue this goal. Therefore, we do not see the pursuit of crime minimization as a strong argument for challenging the existing legal approach.

Selective prosecution cases are among the most sensitive legal cases, since they touch on matters of public security on the one hand, and on the other on

matters of differential treatments of racial and ethnic groups. Each of these two subjects can be inflammatory, and the combination of the two often makes for a difficult public discourse. These issues should not be decided in response to the pressure of public opinion, whose wind, as we have seen after the events of September 11, can blow in either direction. Instead, it is in the interest of society at large that a coherent and transparent legal approach be devised to deal with such cases. We hope that this work can contribute to this important goal.