Can Racially Unbiased Police Perpetuate Long-Run Discrimination?

Bunzel, H.; Marcoul, P.

Publication date:
2003

No. 2003–16

CAN RACIALLY UNBIASED POLICE PERPETUATE LONG-RUN DISCRIMINATION?

By Helle Bunzel and Philippe Marcoul

February 2003

ISSN 0924-7815

Can Racially Unbiased Police Perpetuate Long-Run Discrimination?*

Helle Bunzel Philippe Marcoul†

January 22, 2003

Abstract

We develop a stylized dynamic model of highway policing in which a non-racist police officer is given incentives to arrest criminals, but faces a per stop cost of stop which increases when the racial mix of the persons he stops differs from the racial mix of the population. We define the fair jail rate to be when the racial composition of the jail population is identical to the racial composition of the criminal population. We study the long-term racial composition of the jail population when the policeman decides whom to stop based only on his last period successes in arresting criminals. The study of this "imperfect recall" case shows, consistent with empirical findings, that the long term racial jail rate is always greater than the fair one and the gap increases when incentives are made more powerful. We then study this rate when policemen are provided with data concerning conviction rates for each race, similar to the data which is now being collected in many states. In this case, we find that although the long term rate is still greater than the fair rate, it is smaller than that obtained in the imperfect recall case. We discuss the desirability of such data collection and dissemination of information among police officers.

*We are grateful to Joydeep Bhattacharya, Subir Bose, Jacques Crémer and Brent Kreider for extensive comments on earlier versions of this paper and to Jeremy Greenwood, Nicola Persico, and Oscar Volij for beneficial feedback. As usual, all remaining errors are our own.

†Corresponding author. Department of Economics, 260 Heady Hall, Iowa State University, Ames IA 50011. Phone (515) 294 6311, Fax: (515) 294 0221, e-mail: marcoul@iastate.edu

1 Introduction

Racial profiling among the police force has been a topic of debate for several years now. It first became a national issue in the media when it was suspected that highway troopers stopping cars on the I-80 corridor made heavy use of racial profiling. In April of 1999, it was confirmed by New Jersey Attorney General Peter Verniero that state troopers routinely used the race of drivers on the New Jersey Turnpike to decide whom to stop and search. Since then, it has become generally accepted that at least some racial profiling is practiced by police.

The two sides of the issue of racial profiling can be briefly stated as follows. Opponents of racial profiling find that it is unreasonable to simply use race, without any additional indication of criminal behavior, as probable cause. Proponents argue that if members of one race are statistically more likely to be involved in a certain crime, racial profiling has to be an important tool when fighting crime. Clearly, both of these points have merit. If there is a type of crime exclusively committed by white Americans, there is little point in forcing police to consider people of all ethnicities as potential criminals. On the other hand, if all ethnicities engage in the crime, common perceptions of fairness would imply that all criminals should be at risk for being caught, even if one of the ethnic groups has a higher crime rate. Complicating matters further is the fact that unless all racial groups are investigated, it is not possible to be certain that a specific type of crime is committed exclusively by one racial group. For example, the police could, by only giving Hispanics breathanalyzer tests, create the impression that only Hispanics drive drunk.

There is no clear picture of the extent to which racial profiling is employed. While it is definitely true that blacks and Hispanics are disproportionately represented in the country's jails and prisons, this may be due to factors other than racial profiling.[1] These other factors include the theoretical possibility that blacks and Hispanics may actually engage in more criminal activity, as well as the reasonable notion that lower income people typically cannot afford good legal representation. These issues complicate matters, and it is clear that before proceeding it is necessary to have a specific definition of what exactly constitutes a fair justice system. In this paper, we choose to define a fair justice system as one in which the racial distribution of the prison population is identical to the racial distribution of the criminal population.[2]

[1] According to 1997 numbers from the US census, the number of incarcerated blacks and whites was virtually the same, although blacks only constituted 12 percent of the total of black and white Americans.

[2] Katherine Kersten implies a similar definition when she argues that "... to determine whether Minnesota has "too many" male inmates, we must compare the proportion of males in the prison population with the proportion of males in the *criminal* population, not the population at large..." in her 2001 article in *The Weekly Standard.* Persico (2002)

In this paper, we examine the issues of racial profiling in the context of "high discretion interdiction" (Persico, 2002). One example of such a situation is highway stops. This is the area of policing in which racial profiling was first brought up, and it is in this context that the phrase 'driving while black' was coined, because many African Americans felt that they were targets of stops only because of their skin color.

In our model, we assume that police officers are not racist.[3] The officers are given incentives to arrest as many guilty people as possible, in the form of a fixed payment for every criminal caught. However, they also face a per stop cost that increases whenever the proportion of people they stop from one racial group deviates from the proportion of this group in the whole population. Individual officers maximize their revenue, and only based on this do they make the decision about whom to stop.[4] Our goal is to study the long run equilibrium in this setup. Specifically, we wish to compare the ethnic distribution of the incarcerated population with the ethnic distribution of the entire criminal population.

In order for the police officers in our model to maximize revenue, they need information on the exact crime rates in the different ethnic populations. In reality, police officers do not have hard evidence of exact crime rates in each racial group. Instead, when deciding whom to stop and search, they rely on their experience, word of mouth among colleagues, and statistics on convicted criminals, to make their judgements about who is more likely to be involved in criminal activity. This lack of precise knowledge of actual crime rates in different ethnic populations is a feature which we attempt to capture in our model. In the first version of the model, we assume that the officers use the racial distribution of inmates in jail to estimate rates of criminality in the population, and that they make their decision about whom to stop according to these estimates. In the second version of the model we consider a situation where officers have access to information on past total number of stops by race, as well as conviction rates by race, of those stopped. The reason for making this assumption is that this type of data is currently being collected in many, but not all, states, and federal legislation under consideration would mandate the same type of data collection. This additional information enables officers to get a precise picture of the crime rates in each ethnic group, and this naturally affects their choices about whom to stop.

employs the definitions that both groups should be investigated with the same intensity, which in our setup would lead to a fair outcome. Gary S. Becker provides a somewhat different definition in his July, 2000 article in *Business Week* "... if stops of blacks and whites uncover evidence at about the same rate, that suggests the police are using reasonable criteria for deciding whom to stop and search." This will never occur in our model unless crime rates are indentical, and even then this may not be the best definition of fairness. See Chakravarty (2002) for a detailed discussion on this.

[3] The assumption that police officers are not racist is supported by findings in Knowles et al. (2001). They find that excess investigation of blacks is motivated purely on efficiency grounds rather than racial prejudice. Note, however, that we are modelling what happens even if the officers are not racist.

[4] A study by John Lamberth shows that 98.1% of all cars on a stretch of the New Jersey Turnpike are exceeding the speed limits. (See *New Jersey v. Soto*, 1996) This makes it reasonable to assume that police can stop any car passing by.

Through these two different information assumptions, we are able to examine the long-run effects of the belief formation on the incarceration rates of different ethnic groups.

In the framework described above, we show that providing incentives to arrest criminals always gives rise to excess incarceration of the ethnic group which is more criminal, compared to their share of the criminal population. In other words, if 11% of whites commit crimes, but only 10% of blacks do, in the long run, whites will comprise *more* than 11/21 of the incarcerated population. In that sense, jail rates will be "unfair", in that they do not represent the true rate of criminality among a race. The long-run result is in part due to a self-reinforcing flaw in the way officers estimate crime rates. The mechanism is as follows: The officers are given incentives to arrest criminals. Based on their estimate of the racial criminality, they decide how many people to arrest within a community. The power of incentives may then lead them to arrest more people within a given racial community. As a result, members of this community may be over-represented in jail, which in turn leads to inflated estimates of the crime rate among blacks in the next period.[5]

We find that improved information lessens the problem of over-incarceration, but does not completely eliminate it. This result lends credence to the suggestion made by many civil rights and anti-discrimination groups that police officers should record the race of the people who are stopped. It is worth noting that, for example, the ACLU does not see this data collection as a solution in itself, but rather as a way of documenting the existence of the problem. In contrast, we argue that simply dispersing better information on crime rates may diminish the magnitude of the problem.[6] In other words, if the police officers are provided with information about the stops they made and realize that for a given race the stops did not lead to as many convictions as expected, they will reduce the number of people stopped from that race, leading to less over-incarceration in the future. In spite of this reduction in over-incarceration, the incentives provided to the police officer ensure that he will continue to sample more people from the high crime group.

An interesting implication of the results we obtain is that over-incarceration of a specific ethnic group can occur, even if officers are not racist in the least. This, in turn, would imply that programs intended

[5] In Persico (2002) the fair steady state is reached even though there is no stigma attached to non-random sampling. There crime rates decline as a population is sampled, eventually equalizing the crime rates across populations. This implies that random sampling becomes optimal, and the fair steady state is achieved. Presumably, if a cost to non-random sampling was introduced, there might not be enough excess sampling of the high-crime group to equalize crime rates across different groups and it might be possible to obtain results similar to those presented in this paper.

[6] In their 1999 report on racial profiling, the ACLU discusses a bill that would make it mandatory to collect data on the race of people stopped on highways: "The idea behind the bill was that if the study confirmed what people of color have experienced for years, it would put to rest the idea that African Americans and other people of color are exaggerating isolated anecdotes into a social problem. Congress and other bodies might then begin to take concrete steps to channel police discretion more appropriately."

to educate police officers and make them 'color-blind' may have absolutely no effect whatsoever; not because the police officers refuse to reform, but because they are not racist in the first place! As long as they are provided with incentives to catch as many criminals as possible, the over-incarceration is a result of optimizing behavior on the part of police.

While our paper shares certain aspects of belief formation and self-fulfilling negative discrimination with many other papers, such as Farmer and Terrel (1996), Acemoglu (1995), Coate and Loury (1993), and more recently, Verdier and Zenou (2001), the similarities are somewhat superficial. The models in the papers mentioned above possess the common feature that there is a principal (often an employer) who acts upon a negative belief and takes a *cost-less* discriminatory action against the agent. The principal's action makes it optimal for the party who is discriminated against, to take such actions as to make the principal's "prophecy" self-fulfilling. In our setup, when the current generation of police use incarceration data to decide on which race to interdict more frequently, they may perpetuate their mistaken belief against the over-represented race, but at a cost: the cost of deviating from strictly random sampling. The aforementioned papers do not model this cost. Two additional points of contrast deserve mention here. First, we clearly demonstrate that long-run discrimination against a race can be obtained even without the assumption of a priori negative beliefs. Second, because we do not discuss issues relating to the supply of crime, we are, per force, silent on whether the race that is discriminated against truly ends up becoming more criminal, thereby "justifying" the discrimination.[7] Another important recent paper by Persico (2002) studies the question: is increased fairness on the part of the police necessarily deleterious for efficiency? He finds that if both races have the same population fraction of criminals (an equilibrium free-entry-into-crime condition in his static model) and if one group is policed with less intensity, then forcing the police to increase the intensity of interdiction on that group could raise both fairness and effectiveness of policing.

The rest of the paper is organized as follows. In Section 2 we set up and solve the model where police officers have limited information. In Section 3 we discuss the implication of the "End Racial Profiling Act of 2001" which, among other things, would mandate data collection such that officers have full information, and in Section 4 we conclude. Proofs are relegated to an appendix.

[7]Persico (2002) discusses the "general strain theory" and how unfairness in police practices may itself encourage crime among those being unfairly treated.

2 A Model of Updating with Limited Information

Consider a situation where there are N citizens in the district of a certain police officer. On each date $t = 1, 2, ..., \infty$, 50% of these are white and the rest are black.[8] Let c_i, $i = b, w$ be the unknown, *true* percentage of criminals among the black and white populations, respectively. Note that in practice these numbers are difficult to measure, because if we simply examine the arrests/convictions, those numbers are typically influenced by the behavior of the police and may be very inaccurate.

Unless otherwise explicitly stated, we will assume throughout that $c_b > c_w$. Clearly, we could just as easily assume that c_w is larger, without results being affected.[9] We will discuss separately the case where the rates are identical, as it is somewhat different.

In addition to the assumptions above, we will assume it is known to everyone that a fraction q of the entire population are criminals. That is, q is known, while the individual components, c_b and c_w are not.[10] Furthermore, we assume that c_b, c_w, and q are all time invariant.

Since the majority of data on racial profiling has been collected in the context of highway stops, this is the situation we have in mind when creating this model. As mentioned in the introduction, almost all cars traveling on the highway can be legally stopped because of minor or major traffic violations. Typically these violations are used as cause to stop the vehicle, but after the vehicle has been stopped it is searched for drugs and weapons or other signs of more serious offenses. We model the highway stop simply by saying that the police officer can legally stop any car, and after the car is stopped he can determine with 100% accuracy whether the person in the car is guilty of a serious offense. We assume that if the policeman stops a guilty person, that person is always subsequently convicted and spends the remainder of the period in jail. If a policeman stops an innocent person, that person is immediately released and will not be stopped for the rest of the period.[11] We assume that the officer will receive a compensation of $\$x$ for each guilty person he catches. This assumption is a proxy for all incentive rewards, such as pay increases, promotions, medals, etc. However, the policeman will receive \$0 when no one is arrested. Hence, the payments cannot be strictly negative. This assumption is a limited liability condition.

[8] This assumption does not qualitatively affect the results; it merely simplifies both the algebra and the intuition.

[9] Note that here we deviate from Persico (2002), who assumes that in equilibrium $c_b = c_w$. This is the natural outcome in his model, since he assumes that there is no cost associated with stopping only members of one race. This assumption implies that if $c_b > c_w$, only blacks will be stopped, which, in turn, will lead to the crime rate of whites to rise. With the assumption that it is very costly to the officer to stop only members of one race, this need not be the equilibrium outcome.

[10] The idea behind this assumption is that it is easier to determine the total number of crimes ($= q \cdot n = \frac{1}{2}(c_b + c_w)\, n$) and criminals than it is to determine the racial distribution of criminals.

[11] Note that since criminals are released at the end of the period, the percentage of criminals in the various populations remains constant period after period.

The policeman stops exactly $n << N$ people in each period. This number is pre-determined and beyond his control. If the police officer expends no special effort, he stops a racial mix which corresponds to the population proportion, that is $n/2$ blacks and $n/2$ whites. At a cost, however, he can change these proportions. Let S_b denote the proportion of the people he stops who are black. The associated cost function is $\Omega(S_b) = A\left(\frac{1}{4} - S_b(1-S_b)\right)$, $A >> 0$. This is the cost per stop of deviating from strictly random sampling. A is the exogenous parameter which determines the per-stop cost for a given racial mix. Note that $\Omega(.)$ obtains its minimum at $S_b = \frac{1}{2}$, is convex, and symmetric around $S_b = \frac{1}{2}$.[12] Our reasons for assuming that the cost function obtains its minimum at $\frac{1}{2}$ is that, in reality, random stopping is easier for the police officer, in the sense that there is no need to establish the race of the motorist before pulling him over, and that there is no need to wait for someone of the right ethnic group to pass by. Another reason for this feature is that we believe it is costlier for police officers to stop higher proportions of one race. This cost is associated with the stigma connected to accusations of bias. For example, Mr. Dunphy of the Los Angeles Police Department discusses the data on ethnicities collected on scantron sheets in a 2001 article in *National Review* and states, "If I show a pattern of filling in too many of the wrong circles there may be dire consequences awaiting me, no matter how benign my contacts have been."

We assume that the policeman is willing to accept the following contract if and only if his revenue exceeds his best outside option, which is normalized to 0 for simplicity. Therefore, "participation" in the offered contract will be obtained as long as the expected revenue is positive.

Initially, consider as a benchmark a case where c_b and c_w are known. The expected revenue, per stop, of the policeman is then:

$$R(S_b) = S_b \cdot c_b \cdot x + (1-S_b) \cdot c_w \cdot x - A\left(\frac{1}{4} - S_b(1-S_b)\right). \quad (1)$$

If $c_b = c_w$, the optimal choice is clearly $S_b^* = \frac{1}{2}$ such that the racial distribution of the incarcerated population will correspond to the racial distribution of the population in general.

If $c_b \neq c_w$, the optimal S_b is easily seen to be $S_b^* = \frac{1}{2} + \frac{(c_b - c_w)x}{2A}$. This suggests that if blacks in fact engage in more criminal activity than whites, more than 50% of the officer's stops will be of blacks.[13]

Now consider the more general case where the officer does not know c_b and c_w. In any period t, if

[12] Note that the symmetry of the cost function implies that it is equally costly to discriminate against either race. If the population consists of α whites and $1-\alpha$ blacks, the cost function should be adjusted to obtain its minimum at $1-\alpha$.

[13] In order to get an interior solution (i.e. $S_b^* \leq 1$), the set of allowable parameters values must be restricted. It is routine to check that if $A \in ((c_b - c_w)x, +\infty)$, S_b^* is strictly smaller than 1. Note that the equilibrium reward of the police officer is $R(S_b^*) = \frac{1}{2}x(c_b + c_w) + \frac{x^2}{4A}(c_b - c_w)^2 > 0$. Thus, participation is always insured.

there are J_b^t blacks and J_w^t whites in jail, the police officer estimates c_b and c_w by

$$\hat{c}_b^t = \frac{J_b^t}{J_b^t + J_w^t} \cdot q \cdot \frac{N}{N/2} = \frac{J_b^t}{J_b^t + J_w^t} \cdot 2q,$$

and

$$\hat{c}_w^t = \frac{J_w^t}{J_b^t + J_w^t} \cdot q \cdot \frac{N}{N/2} = \frac{J_w^t}{J_b^t + J_w^t} \cdot 2q.$$

Intuitively, these are the estimates obtained by assuming that the prison population is representative. To clarify the construction of these estimates, consider the estimate of c_w. The first term, $\frac{J_w^t}{J_b^t+J_w^t}$, is the proportion of the jail population that is white, and it thus estimates the percentage of the criminal population that is white. $(q \cdot N)$ is the total number of criminals (black and white) in the population. $\frac{J_w^t}{J_b^t+J_w^t} \cdot (q \cdot N)$ is therefore an estimate of the total number of white criminals. When the number is divided by the size of the white population, $N/2$, it provides an estimate of the proportion of criminals among the white population. These estimates would be unbiased estimates of c_b and c_w *if* the prison population were the result of random stops of the population.

The questions which we are interested in answering in this limited information framework are whether the fair steady state is ever obtained, and how the incentive payments to the officer and the information structure affect the long-run outcomes. Before moving on, it will therefore be useful to formally define the fair steady state. This is the steady state where the racial composition of the prison population is identical to the racial composition of the criminal population, in and out of jail. This means that we can denote the proportion of blacks in jail at the fair steady state by $F^* = \frac{c_b}{c_b+c_w}$. This provides a utopian benchmark to which we will compare the steady states of the various models in this paper.

Using aforementioned estimates of $\hat{c}_b^t$ and $\hat{c}_w^t$, the policeman stops a fraction $S_b^t = \frac{1}{2} + \frac{qx}{A}\left[\frac{J_w^t - J_b^t}{J_b^t + J_w^t}\right]$.

The expected number of black people jailed in the next period is then simply the number of black people stopped, nS_b, multiplied by the true crime rate, c_b, such that $J_b^{t+1} = c_b n S_b^t$.[14] Similarly $J_w^{t+1} = c_w n \left(1 - S_b^t\right)$. The racial distribution of the jail population for the next period is therefore given by

$$\begin{aligned} \frac{J_b^{t+1}}{J_b^{t+1} + J_w^{t+1}} &= \frac{c_b S_b^t}{c_w\left(1 - S_b^t\right) + c_b S_b^t} \\ \frac{J_w^{t+1}}{J_b^{t+1} + J_w^{t+1}} &= \frac{c_w\left(1 - S_b^t\right)}{c_w\left(1 - S_b^t\right) + c_b S_b^t}. \end{aligned} \tag{2}$$

Now, since S_b^t depends on $\frac{J_b^t}{J_b^t+J_w^t}$ and $\frac{J_w^t}{J_b^t+J_w^t}$, this is a system of difference equations. We denote steady states obtained as equilibria from (2) by J^*.

[14] Our model is a reduced form model. In a less parsemonious model, one would formulate an arrival process of criminals such that the expected value of J_b^{t+1} would be $c_b n S_b^t$. We feel that this would add complexity without adding insight.

At time 1, we will assume that there is an existing jail population, J_b^0 and J_w^0. The model is thus initiated by the officer calculating his first estimates, maximizing revenue and determining the optimal number of blacks (and whites) to stop. It turns out that the dynamical system is history independent, so the composition of the initial jail population does not affect the jail rates at the steady states.

Note that the dynamical system (2) can be simplified and written as the following unidimensional system:

$$z_{t+1} = \frac{\frac{1}{2}c_b + c_b \frac{qx}{A}\left[2z_t - 1\right]}{q + (c_b - c_w)\frac{qx}{A}\left[2z_t - 1\right]}, \ t = 1, 2, ...$$

We are now ready to state the main result of this paper:

Proposition 1 *There are two steady states in the model, J_L^* and J_H^*, with the following properties:*

a) $J_L^* < J_H^*$ and J_H^* is locally stable.

b) At J_H^*, when $x(A)$ increases (decreases), the proportion of blacks in the jail population increases and the proportion of whites decreases.

c) $J_H^* = F^*$ if and only if x is 0 or A is ∞.

d) When $x > 0$, $J_H^* > F^*$.

Proposition 1 states that there is a single stable steady state that can be approached from arbitrary initial conditions. Furthermore, we show that increasing incentives to officers will worsen the excess incarceration of blacks and, similarly, increasing the cost of non-random sampling will make the excess incarceration less severe. Finally, we show that only when there are no incentives will the fair steady state be obtained. For all other combinations of cost and incentives, blacks will be over-represented in the prison population. This result is driven by the fact that when blacks have higher crime rates than whites, more criminals will be caught if more blacks are stopped. In the absence of a cost of non-random stopping, police officers would stop only blacks. Since this is not cost-less, police officers will equalize marginal benefit to marginal cost and, therefore, stop more blacks than whites, but not solely blacks. If it were possible to offer no incentives ($x = 0$), police officers would not be willing to incur the costs of non-random sampling, hence the fair steady state would obtain. Note however, that for the officer to participate when there are no incentives, costs (A) must be zero as well.

To illustrate the dynamic evolution of the model, we present the results from a few simulations where we have started the model with some initial jail rates and let the jail population develop. We let these simulations run for 30 periods, and follow the evolution of the rate of whites and blacks in jail. In each graph we have plotted the fair rate for each race, to provide a benchmark. Graphs 1 and 2 have identical parameter values, except for the starting values of J_b and J_w. Comparing these two demonstrates the fact that the steady state is history independent (the fact can also be established by looking at the expression for J_H^* provided in the appendix). Graphs 2, 3 and 4 are created using identical parameter values, except for x, which varies, demonstrating the effect of the strength of the incentives on the steady state. Graphs 5, 6, and 7 illustrate the very special case where crime rates for blacks and whites are *exactly* identical. In this situation, the model behaves somewhat differently. Depending on the parameter values, one of two steady states is obtained: either a steady state where only one race is stopped or the fair steady state. In the steady state where only one race is stopped, which race gets stopped is entirely determined by initial beliefs. If incentives are strong enough, the police officer will never sample the race he perceives to be low-crime often enough to gain evidence against his initial beliefs. While it is interesting that the model is able to generate a situation where initial jail rates (possibly created by a racist police force of the past) cause persistent discrimination, we include this example mainly for the sake of completeness. In a situation where only one race is stopped, it becomes critically important to have the supply of crime included in the model (see Persico (2002)) and, in addition, the chances that crime rates are exactly the same seem very small given the vast socioeconomic differences between the populations.

3 Evaluating the effects of the "End Racial Profiling Act of 2001"

As a reaction to lawsuits and complaints by various civil rights groups, legislation is being introduced at both the state and federal level to monitor and/or curb racial profiling. The most recent, and by far the most comprehensive, federal legislation to be introduced in Congress is the "End Racial Profiling Act of 2001" (ERPA).[15] The ERPA starts off with a long list of findings. Among these are that the vast majority of law enforcement agents discharge their duties without bias, and that racial profiling is a practice that is well documented and undesirable. The ERPA states that "No law enforcement agent or law enforcement agency shall engage in racial profiling." In addition to the prohibition of the act of racial profiling, the ERPA mandates that detailed data must be collected on stops, searches, seizures, and arrest, such that it can be determined whether racial profiling is taking place.

Several previous laws suggesting that data on racial profiling be collected have been put before Congress, but ERPA is the first to clearly define exactly what is understood by racial profiling. In ERPA, it is written that: "The term 'racial profiling' means the practice of a law enforcement agent relying, to any degree, on race, ethnicity, or national origin in selecting which individuals to subject to

[15] Introduced in the Senate as S.989 and in the House as H.R.2074.

routine investigatory activities, or in deciding upon the scope and substance of law enforcement activity following the initial routine investigatory activity,[16] except that racial profiling does not include reliance on such criteria in combination with other identifying factors when the law enforcement agent is seeking to apprehend a specific suspect whose race, ethnicity, or national origin is part of the description of the suspect." This definition is substantially stronger than those typically seen in state legislation. At the state level, using race as an indicator for criminal activity is only considered racial profiling if an officer or agency relies *solely* on race, ethnicity, or national origin in selecting which individuals to stop.[17] This obviously is a significant difference, since it is extremely difficult to verify that there was no other cause for the stop than race, making it substantially more difficult to enforce the state legislation. The definition suggested in ERPA also rules out using knowledge that one ethnicity is much more likely to commit a given type of crime than another ethnicity. Using this type of knowledge is often called statistical discrimination and is deemed an important tool by law enforcement agencies. In the context of the type of traffic stops which we are modelling, this very strict definition has the implication that *only random sampling is allowed.* As a consequence, the legislation amounts to outlawing both racial and statistical (based on race) discrimination, and Congress is, in fact, attempting to put into force the fair steady state.

When attempting to determine the effects of this legislation, should it be enacted, there are two aspects we need to consider. The first is data collection and the additional information the data provides, and the second is the consequences of the act of racial profiling being outlawed. When we consider the effects of the data collection below, we will assume that this collection provides the officers themselves with additional information and examine the long-run consequences of this additional information. We model the prohibition of racial profiling as a sharp increase in the cost of deviating from random sampling, rather than the actual enforcement of random sampling. The reason for this is that actual random sampling is clearly impossible to enforce. But since racial profiling is now illegal, it seems reasonable to assume that the police officer will have to go to great lengths to attempt to hide any non-randomness, and that the potential cost if he is caught has suddenly become substantially higher than before the legislation was enacted.

To examine the effects of the additional information in the context of the model, we assume that the

[16] In the ERPA, routine investigatory activities are defined in the following manner: "The term 'routine investigatory activities' includes the following activities by law enforcement agents: traffic stops; pedestrian stops; frisks and other types of body searches; consensual or nonconsensual searches of the persons or possessions (including vehicles) of motorists or pedestrians; ..."

[17] Note that since race is the only heterogeneity among motorists in our model, the model cannot distinguish between the two definitions in its current form.

officers now have precise information on the race of the people they have stopped as well as the conviction rates for each race.[18] This would enable the officers to estimate crime rates in the two populations, using the conviction rates from these populations. This would, at time $t+1$, provide estimators of the form: $\hat{c}_b^{t+1} = \frac{1}{t}\sum_{j=1}^{t} \frac{J_b^j}{S_b^j n}$ and $\hat{c}_w^{t+1} = \frac{1}{t}\sum_{j=1}^{t} \frac{J_w^j}{S_w^j n}$, where the denominators are the number of people of the given race who are stopped, and the numerator represents the number of convictions among this group. Note that under broad sets of assumptions on the arrival process of vehicles, these are unbiased and consistent estimators, since on average $J_b^t = c_b S_b^t n$.

Because we are interested in the long-run outcomes of the legislation, we will approximate the consistent estimators described above by the true values, c_b and c_w (it also significantly simplifies the analysis of the model). Thus, the steady state we are interested in will be the one where c_b and c_w are known. From (1), we know that the per-stop revenue is defined as

$$R(S_b) = S_b \cdot c_b \cdot x + (1 - S_b) \cdot c_w \cdot x - A(1 - S_b(1 - S_b)),$$

which results in the following first order condition of the officer's revenue maximization problem:

$$c_b \cdot x - c_w \cdot x + A - 2AS_b = 0.$$

The optimal fraction is then seen to be $S_b^* = \frac{1}{2} + \frac{(c_b - c_w)x}{2A}$. This, in turn, provides the steady state black fraction of the jail population as

$$J_{ERPA}^* = \frac{J_b^*}{J_b^* + J_w^*} = \frac{c_b + c_b(c_b - c_w)\frac{x}{A}}{c_b + c_w + (c_b - c_w)^2 \frac{x}{A}}. \tag{3}$$

Note that since the model is no longer dynamic, it is seen that this is a unique steady state. This is the steady state which results *simply from the collection of data*. We have not yet considered the impact of the actual ban of racial profiling. It is interesting to compare this steady state to that obtained in the previous section because data collection is an essential part of ERPA, but it is also interesting because some states are currently collecting this data, while others are not. Since the data collection itself is not cost-less, it is of interest to determine whether it provides other benefits than the enforcement of the legislation.[19] To examine the effects of the ban on racial profiling, we need to determine how J_{ERPA}^* changes when costs of deviating from random sampling increase. Finally, it is of interest to compare the ERPA steady state to the fair steady state which the legislation is attempting to obtain.

[18] In reality they would have similar data from other officers working the same highway, and, thus, have access to even more information.

[19] In a 2001 article in The Times-Picayune, Representative Cedric Richmond, D-New Orleans is quoted as saying that such data collection would cost Louisianna about $1 million per year.

Proposition 2 *a) The ERPA steady state always has fewer blacks in jail than the limited information steady state, but more than the fair steady state.*
b) The level of the ERPA steady state can be decreased either by increasing the cost of non-random stopping A, or by decreasing the level of x.
c) ERPA equals the fair steady state when there is no incentive scheme ($x = 0$) or when costs of non-random sampling are infinity ($A = \infty$).

The implication of this theorem is that the excess jailing of blacks can be reduced by ensuring that police officers have access to enough data to obtain accurate estimates of crime rates in the different populations. Thus, the data collection, which the legislation views solely as a means of enforcing the ban on racial profiling, in itself decreases the burden on the population that is being investigated with higher intensity. While the steady state becomes more fair with the additional information, even the ban on racial profiling does not force the police to sample completely randomly unless any and all incentives to catch guilty criminals can be eliminated, or alternatively the cost of non-random sampling is infinity. We believe that this policy point is important because implementing the ban on racial profiling is deemed to be very costly. Our model shows that, even though a state is not willing or able to enforce the ban on racial profiling, it is likely to derive benefits from the data collection as long as this information is dispersed among its police officers.

Comparing our results to those of other papers in the literature, for example Persico (2002), it is worth noting that those other papers obtain the fair steady state even though they have no cost of non-random sampling. This is driven by the fact that crime rates decline as a population is sampled. In those models the crime rates are eventually equalized through this mechanism, random sampling becomes optimal, and the fair steady state is achieved. It is conceivable that if a cost of non-random sampling was to be introduced in these models, there might not be enough excess sampling of the high-crime group to equalize crime rates across different groups and it might be possible to obtain results similar to those presented in this paper.

In short, we find that when the crime rates cannot equalize, as clearly they have not yet done in the US, it is not possible to obtain the fair steady state, as long as police officers are provided with incentives to arrest criminals. Interestingly, the ERPA actually addresses the issue of incentives. Under the header "FINDINGS", the following statement is made: "Current local law enforcement practices, such as ticket and arrest quotas, and similar management practices, may have the unintended effect of encouraging law enforcement agents to engage in racial profiling." In section 201, under the header

"POLICIES TO ELIMINATE RACIAL PROFILING," the ERPA states that "IN GENERAL - Federal law enforcement agencies shall cease existing practices that encourage racial profiling." While the bill does not go so far as to eliminate incentives, this is purely due to the fact that Congress does not find it to be well established that incentive pay leads directly to racial profiling.

4 Conclusion

In this paper, we show that statistical discrimination is typically optimal in the context of highway stops. This is in line with empirical findings in the literature. Specifically, Knowles, Todd, and Persico (2001) develop a test to determine whether the decision by a policeman to stop a vehicle is triggered by efficiency reasons or by racist motives. They apply the test to data gathered in a recent legal case in which state troopers have been accused of racial discrimination. The authors are unable to show that racial discrimination is motivated by reasons other than efficiency. This sample data exhibits an overwhelming number of searches performed on African-American motorists, but the proportion of motorists found guilty within the African-American and the white group are very close.[20] This data is also consistent with our results that even marginally higher crime rates in one group can lead to a higher percentage of stops from this group, and, in turn, imply that the high crime group will be over-represented in jails.

An important issue that we have refrained from discussing so far is the welfare implications of the results. When considering societal welfare, there are three aspects that we consider important: Efficiency, that is, getting as many criminals as possible off the streets at the lowest possible cost; fairness; and the cost of data collection. Any welfare function would have to contain these three parts. If a welfare function contained these three elements, it would be possible to make inferences about the amount of data to collect and the level of incentives to provide to police officers. Unfortunately, it is unclear how one would weigh issues of fairness against efficiency, or whether there is even a trade-off between the two (see Persico (2002)). As such, we leave these issues for future research.

[20] They find that among a total of 1530 searches, 63.4% were performed on African-Americans, whereas only 29.3% were performed on white motorists. However, within the set of people arrested, 34% of the blacks were found guilty of drug possession, as were 32% of the whites.

References

[1] Acemoglu, D. (1995), "Public policy in a model of long-term unemployed", *Economica* 62, pp 161-178.

[2] American Civil Liberties Union Freedom Network (1999), "Driving While Black. Racial Profiling on Our Nation's Highways." *Special Report*, June.

[3] Becker, G. S. (1971), "The Economics of Discrimination," *University of Chicago Press,* Chicago.

[4] — (2000), "Tough Justice Is Saving Our Inner Cities," *Business Week*, July 17.

[5] Borooah, V.K. (2001), "Racial bias in police stops and searches: an economic analysis" *European Journal of Political Economy* 17, pp 17-37.

[6] — (2002), "Racial bias in police stops and searches: a reply," *European Journal of Political Economy* 18, pp 607-608.

[7] Chakravarty, S.P. (2002), "Economic analysis of police stops and searches: a critique,"'*European Journal of Political Economy* 18, pp 597-605.

[8] Coate, S. and G.C. Loury (1993), "Will affirmative-action policies eliminate negative stereotypes?" *American Economic Review* 83, pp 1220-1240.

[9] Doster, S. (2001), "Racial Profiling Law Lacks Teeth; Loophole Lets Police Off Hook, Some Say," *The Times-Picayune* (New Orleans), December 16.

[10] Dunphy, J. (2001), "Counting Noses," *National Review*, December 21.

[11] Farmer, A. and D. Terrel (1996), "Discrimination, Bayesian updating of employers belief, and human capital accumulation", *Economic Inquiry* 34, pp 204-219.

[12] Kersten, Katherine (2001), "Race to Conclusions: What the activists don't tell you about racial profiling," *The Weekly Standard*, August 20.

[13] Knowles J., P. Todd , and N. Persico. (2001), "Racial Bias in Motor-Vehicle Searches: Theory and Evidence," *Journal of Political Economy* 109(1), pp. 203-29.

[14] Persico N. (2002), "Racial Profiling, Fairness, and Effectiveness of Policing," Forthcoming, *The American Economic Review*.

[15] Verdier T., Y. Zenou. (2001), "Racial Beliefs, Location and the Causes of Crime," mimeo.

Appendix

A Proof of Proposition 1

Let $Q = \frac{qx}{A}$, $z_t = \frac{J_b^t}{J_b^t + J_w^t}$ and $y_t = \frac{J_w^t}{J_b^t + J_w^t}$. Then we can write equations (2) as

$$\begin{aligned} z_{t+1} &= \frac{\frac{1}{2}c_b + c_b Q\left[z_t - y_t\right]}{q + (c_b - c_w)\, Q\left[z_t - y_t\right]} \\ y_{t+1} &= \frac{\frac{1}{2}c_w - c_w Q\left[z_t - y_t\right]}{q + (c_b - c_w)\, Q\left[z_t - y_t\right]} \end{aligned}$$

Now, since $z_t + y_t = 1$, this system reduces to a one-dimensional dynamical system. Specifically, we will examine the system

$$z_{t+1} = \frac{\frac{1}{2}c_b + c_b Q\left[2z_t - 1\right]}{q + (c_b - c_w)\, Q\left[2z_t - 1\right]} \tag{A1}$$

The steady states can be found by equating z across periods, such that

$$\begin{aligned} z^* &= \frac{1}{4Q\left(c_b - c_w\right)} \times \\ &\left(3Qc_b - \frac{1}{2}\left(c_b + c_w\right) - Qc_w \pm \sqrt{\left(c_b + c_w\right)^2 Q^2 + \left(c_b - c_w\right)^2 Q - 4c_b c_w Q + \frac{1}{4}\left(c_b + c_w\right)^2}\right) \end{aligned}$$

Now, clearly two steady states exist as long as the term under the square root is positive. It is easy to verify that this is always the case. Note that the condition states that a second degree polynomial in Q must be positive:

$$\left(c_b + c_w\right)^2 Q^2 + \left(c_b - c_w\right)^2 Q - 4c_b c_w Q + \frac{1}{4}\left(c_b + c_w\right)^2 > 0 \tag{A2}$$

Now the roots of the polynomial in (A2) are

$$Q^{\maltese} = \frac{1}{2\left(c_b + c_w\right)^2}\left(-c_w^2 + 6c_b c_w - c_b^2 \pm 16\sqrt{c_b c_w}\sqrt{-\left(c_b - c_w\right)^2}\right).$$

Clearly these are both complex, and, therefore, the inequality in (A2) is always satisfied. As a result, there are always two steady states.

A.1 Proof of a)

Now name the steady states such that

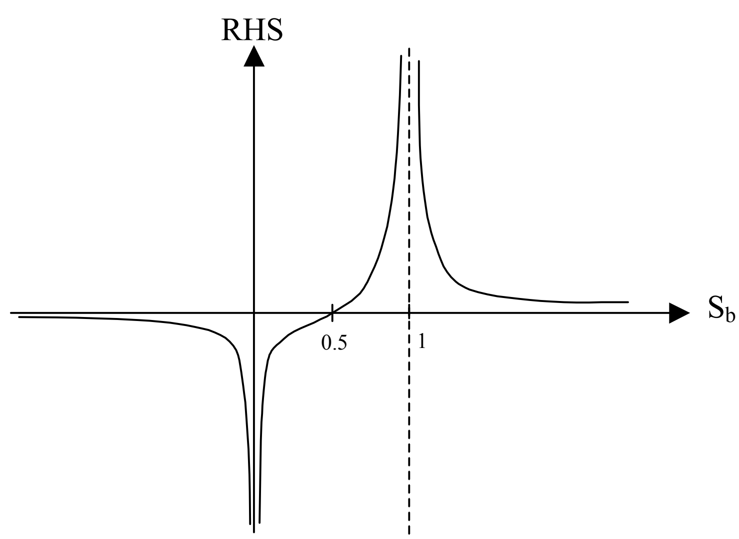
$$\begin{aligned} J_L^* &\equiv \frac{1}{4Q\left(c_b-c_w\right)}\left(3Qc_b-\frac{1}{2}\left(c_b+c_w\right)-Qc_w\right. \\ &\left.-\sqrt{\left(c_b+c_w\right)^2Q^2+\left(c_b-c_w\right)^2Q-4c_bc_wQ+\frac{1}{4}\left(c_b+c_w\right)^2}\right), \text{ and} \\ J_H^* &\equiv \frac{1}{4Q\left(c_b-c_w\right)}\left(3Qc_b-\frac{1}{2}\left(c_b+c_w\right)-Qc_w\right. \\ &\left.+\sqrt{\left(c_b+c_w\right)^2Q^2+\left(c_b-c_w\right)^2Q-4c_bc_wQ+\frac{1}{4}\left(c_b+c_w\right)^2}\right). \end{aligned}$$

Clearly, by this definition, $J_L^* < J_H^*$ when $c_b > c_w$. It then remains to be proven that J_H^* is stable. To simplify notation, we will define the function $f : \mathbb{R} \to \mathbb{R}$ by

$$f\left(z\right)=\frac{\frac{1}{2}c_b+c_bQ\left(2z-1\right)}{q+\left(c_b-c_w\right)Q\left(2z-1\right)},$$

such that $z_{t+1}=f\left(z_t\right)$. Stability is then implied for a steady state J^* if $f'\left(J^*\right)<1$. Therefore, it is necessary to examine the properties of f in some detail. First note that $f'\left(z\right)=\frac{8bQw}{\left(\frac{1}{2}(b+w)+(b-w)Q(2z-1)\right)^2}>0$, and that f has an asymptote at $z_a\equiv\frac{1}{2}-\frac{1}{4Q}\frac{(b+w)}{(b-w)}$. In addition, since $f''\left(z\right)=\frac{-8Q(b-w)}{\left(\frac{1}{2}(b+w)+(b-w)Q(2z-1)\right)}f'\left(z\right)$, it is easy to see that $f''\left(z\right)<0$ for $z>z_a$ and $f''>0$ for $z<z_a$.

To show that the steady state J_H^* is stable, we need to show that $f'\left(J_H^*\right)<1$. To prove this, we will show that J_L^* lies to the right of the asymptote. Since $J_H^*>J_l^*$, this would also imply that J_H^* lies to the right of the asymptote, and we will have established that f can be depicted by Figure A1 below.



Clearly, if this is the correct depiction, it must be the case that $f'(J_H^*) < 1$, since f cuts the 45-degree line from above at J_H^*. For J_L^* to lie to the right of the asymptote, it must be the case that

$$\frac{1}{4Q(c_b-c_w)}\Bigg(3Qc_b - \frac{1}{2}(c_b+c_w) - Qc_w - \sqrt{(c_b+c_w)^2 Q^2 + (c_b-c_w)^2 Q - 4c_bc_wQ + \frac{1}{4}(c_b+c_w)^2}\Bigg) > \frac{1}{2} - \frac{1}{4Q}\frac{(c_b+c_w)}{(c_b-c_w)}$$

This can easily be reformulated as

$$\left(\frac{1}{2}+Q\right)(c_b+c_w) > \sqrt{(c_b+c_w)^2 Q^2 + (c_b-c_w)^2 Q - 4c_bc_wQ + \frac{1}{4}(c_b+c_w)^2}.$$

Since both sides of this equation are positive, we can square both sides and rearrange to get the following very simple expression:

$$0 > -4c_bc_wQ$$

Clearly, this always holds, and we have established that J_H^* is a stable steady state.

A.2 Proof of b)

To establish the signs of these derivatives, we will first show that $\partial J_H^*/\partial Q > 0$:

$$\begin{aligned}\frac{\partial J_H^*}{\partial Q} &= \frac{1}{8}\frac{(c_b+c_w)}{(c_b-c_w)} \\ &\quad -\frac{1}{8}\frac{2Qc_b^2 - 12c_bc_wQ + 2Qc_w^2 + c_b^2 + 2c_bc_w + c_w^2}{Q^2(c_b-c_w)\sqrt{(c_b+c_w)^2 Q^2 + (c_b-c_w)^2 Q - 4c_bc_wQ + \frac{1}{4}(c_b+c_w)^2}}\end{aligned}$$

Simple algebra verifies that this expression is positive if

$$\begin{aligned}&(c_b+c_w)\sqrt{\left(4Q^2(c_b+c_w)^2 + 4Qc_b^2 - 24c_bc_wQ + 4Qc_w^2 + (c_b+c_w)^2\right)} \\ > \;& 2Qc_b^2 - 12c_bc_wQ + 2Qc_w^2 + (c_b+c_w)^2\end{aligned}$$

The expression to the left of the inequality is definitely positive, but the expression on the right might be negative. If it is negative, the inequality clearly holds, and if it is positive, we can square both sides while maintaining the direction of the inequality. Squaring both sides and simplifying the expression, we are left with

$$c_w^2 + c_b^2 - 2c_bc_w = (c_b-c_w)^2 > 0,$$

which trivially holds for all parameter values. Now, since $Q = \frac{qx}{A}$, $\partial J_H^*/\partial x = \partial J_H^*/\partial Q \cdot \partial Q/\partial x > 0$ and $\partial J_H^*/\partial A = \partial J_H^*/\partial Q \cdot \partial Q/\partial A < 0$, and the proof is complete.

A.3 Proof of c) and d)

Here we need to investigate when $J_H^* = F^*$. Recall that $F^* = c_b/(c_b + c_w)$ and

$$
\begin{aligned}
J_H^* &= \frac{1}{4Q(c_b - c_w)} \times \\
&\left(3Qc_b - \frac{1}{2}(c_b + c_w) - Qc_w + \sqrt{(c_b + c_w)^2 Q^2 + (c_b - c_w)^2 Q - 4c_b c_w Q + \frac{1}{4}(c_b + c_w)^2}\right).
\end{aligned}
$$

Note that when $Q \to 0$,

$$
\begin{aligned}
\lim_{Q\to 0} J_H^* &= \lim_{Q\to 0} \frac{1}{4Q(c_b - c_w)} \times \\
&\lim_{Q\to 0}\left(3Qc_b - \frac{1}{2}(c_b + c_w) - Qc_w + \sqrt{(c_b + c_w)^2 Q^2 + (c_b - c_w)^2 Q - 4c_b c_w Q + \frac{1}{4}(c_b + c_w)^2}\right) \\
&= \lim_{Q\to 0} \frac{1}{4Q(c_b - c_w)} \times \left(-\frac{1}{2}(c_b + c_w) + \sqrt{\frac{1}{4}(c_b + c_w)^2}\right)
\end{aligned}
$$

Using L'Hospital's rule,

$$
\begin{aligned}
\lim_{Q\to 0} J_H^* &= \lim_{Q\to 0} \frac{1}{4(c_b - c_w)} \times \\
&\lim_{Q\to 0}\left(3c_b - c_w + \frac{1}{2}\left((c_b + c_w)^2 Q^2 + (c_b - c_w)^2 Q - 4c_b c_w Q + \frac{1}{4}(c_b + c_w)^2\right)^{-\frac{1}{2}} \times \right. \\
&\left.\left(2(c_b + c_w)^2 Q + (c_b - c_w)^2 - 4c_b c_w\right)\right) \\
&= \frac{3c_b - c_w + \frac{1}{2}\left(\frac{1}{4}(c_b + c_w)^2\right)^{-\frac{1}{2}}\left((c_b - c_w)^2 - 4c_b c_w\right)}{4(c_b - c_w)} \\
&= \frac{3c_b - c_w + \frac{1}{c_b + c_w}\left((c_b - c_w)^2 - 4c_b c_w\right)}{4(c_b - c_w)} = \frac{4c_b^2 - 4c_b c_w}{4(c_b - c_w)(c_b + c_w)} \\
&= \frac{c_b}{(c_b + c_w)}
\end{aligned}
$$

Now, since $Q = 0$ is equivalent to $x = 0$ or $A = \infty$, we have shown that the fair steady state is implemented for this value. Furthermore, since $\partial J_H^* / \partial Q > 0$, the jail rate for blacks will be higher than the fair steady state for all other values of Q. This completes the proof of c) and d).

B Proof of Proposition 2

First, we will prove that $\frac{J_b^*}{J_b^* + J_w^*} > F^*$. By (3)and the definition of F^*, this can be written as

$$
\frac{c_b + c_b(c_b - c_w)\frac{x}{A}}{c_b + c_w + (c_b - c_w)^2 \frac{x}{A}} > \frac{c_b}{c_b + c_w}.
$$

This expression easily simplifies to

$$2c_w > 0,$$

which holds since we have assumed that $1 > c_i > 0, i = w, b$.

Now, from $(A1)$, we know that

$$z^* = \frac{\frac{1}{2}c_b + c_b Q\left[2z^* - 1\right]}{q + (c_b - c_w)\, Q\left[2z^* - 1\right]}. \tag{A3}$$

We want to prove that z^* is greater than $J^*_{ERPA} = \frac{J^*_b}{J^*_b + J^*_w}$. By $(A3)$and (3)this can be written as

$$\frac{\frac{1}{2}c_b + c_b Q\left[2z^* - 1\right]}{q + (c_b - c_w)\, Q\left[2z^* - 1\right]} > \frac{c_b + c_b\left(c_b - c_w\right)\frac{x}{A}}{c_b + c_w + \left(c_b - c_w\right)^2 \frac{x}{A}}.$$

Re-arranging provides the expression

$$\left(c_b + c_w\right)\left[2z^* - 1\right] > \left(c_b - c_w\right),$$

which in turn simplifies to the condition that

$$z^* > \frac{c_b}{c_b + c_w}.$$

But this condition is exactly the condition that $z^* > F^*$, which we have just verified, and the proof of part a) is, therefore, complete.

To prove part b) we need to show that $\partial J^*_{ERPA}/\partial A < 0$ and that $lim_{x \to 0} J^*_{ERPA} = F^*$. Simple algebra provides us with the following expression for the derivative:

$$\frac{\partial J^*_{ERPA}}{\partial A} = \frac{-2c_w c_b\left(c_b - c_w\right)\frac{x}{A^2}}{\left(c_b + c_w + \left(c_b - c_w\right)^2 \frac{x}{A}\right)^2},$$

which is clearly negative. The final part of the proof will consist of determining $\lim_{x \to 0} J^*_{ERPA}$ and $\lim_{A \to \infty} J^*_{ERPA}$

$$\begin{aligned} \lim_{x \to 0} J^*_{ERPA} &= \lim_{x \to 0} \frac{c_b + c_b\left(c_b - c_w\right)\frac{x}{A}}{c_b + c_w + \left(c_b - c_w\right)^2 \frac{x}{A}} = \frac{c_b}{c_b + c_w}, \text{ and} \\ \lim_{A \to \infty} J^*_{ERPA} &= \lim_{x \to 0} \frac{c_b + c_b\left(c_b - c_w\right)\frac{x}{A}}{c_b + c_w + \left(c_b - c_w\right)^2 \frac{x}{A}} = \frac{c_b}{c_b + c_w}, \end{aligned}$$

which completes the proof.

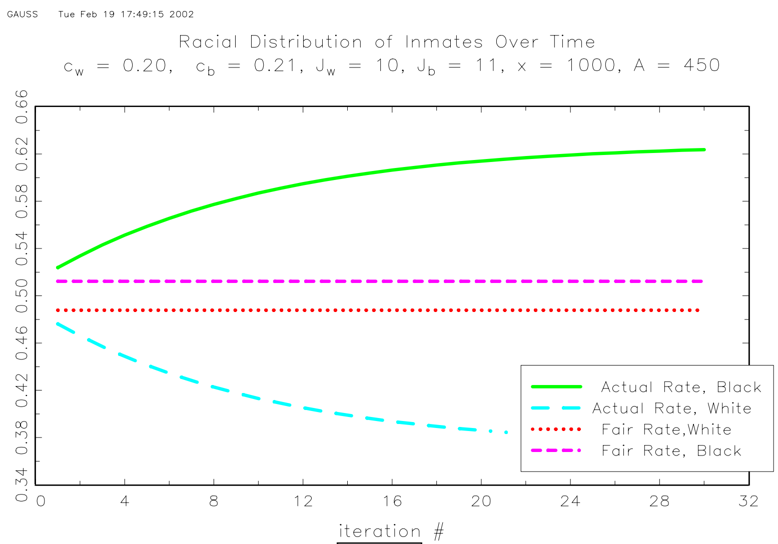


Figure 1

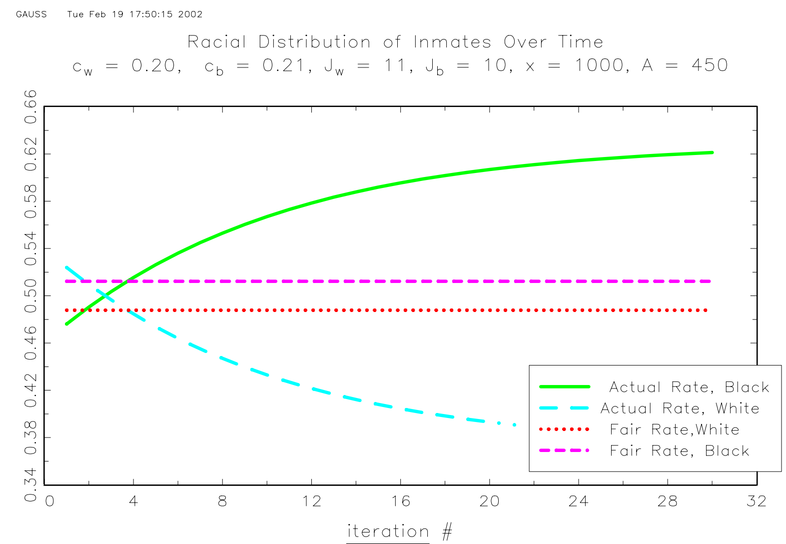


Figure 2

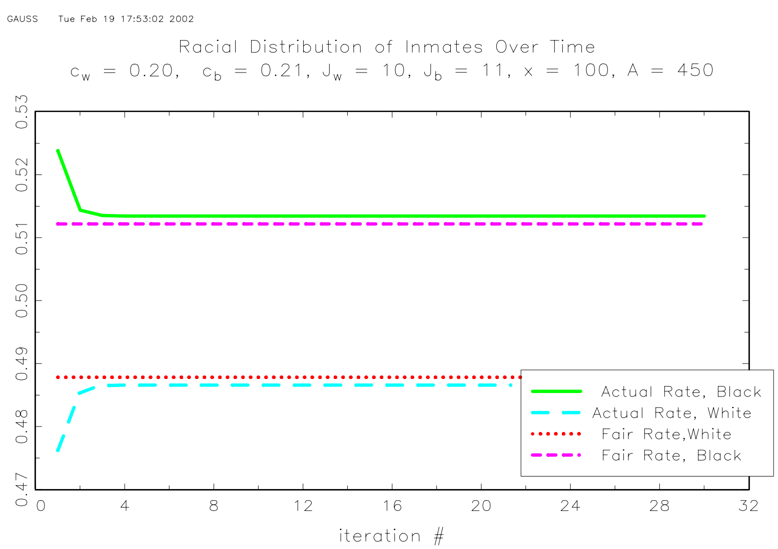


Figure 3

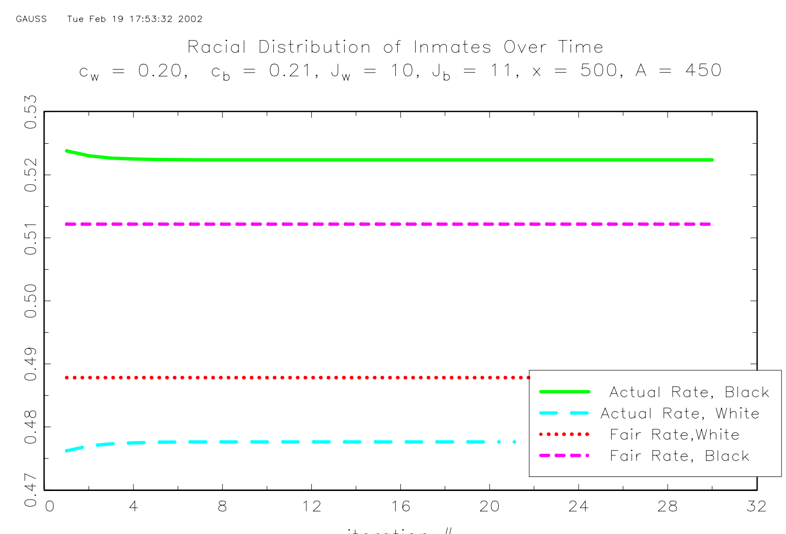


Figure 4

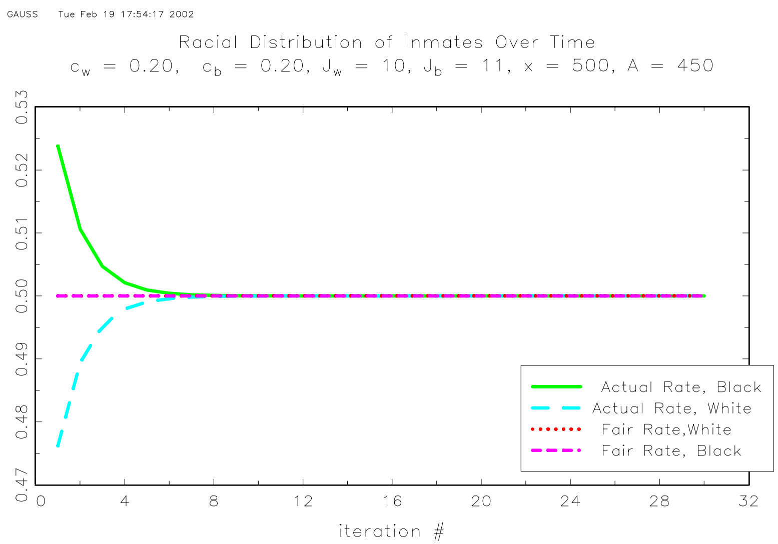


Figure 5

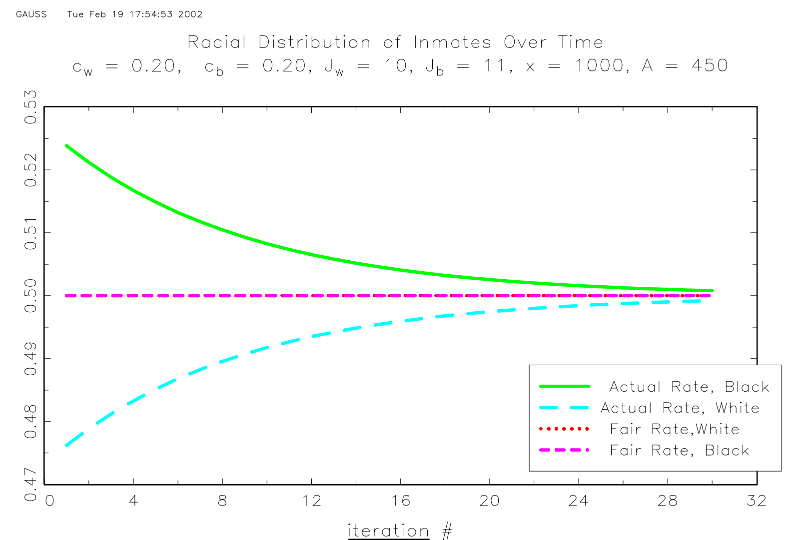


Figure 6

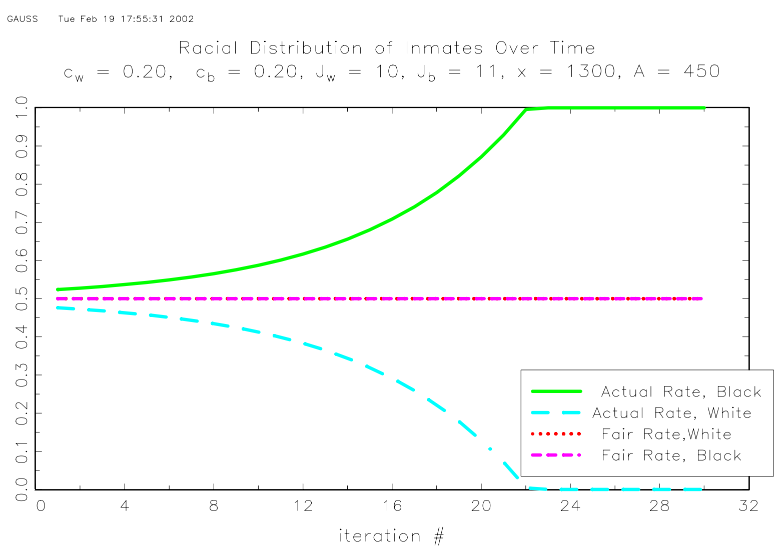


Figure 7