Statistical (and Racial) Discrimination, ‘Banning the Box’, and Crime Rates

Murat C. Mungan*

April 24, 2017

Abstract

This article presents law enforcement models where employers engage in statistical discrimination, and the visibility of criminal records can be adjusted through policies (such as ban the box campaigns). I show that statistical discrimination leads to an increase in crime under plausible conditions. This suggests that societies in which group membership is salient (e.g. through greater racial or religious heterogeneity) are, ceteris paribus, likely to have higher crime rates. Attempting to fix the negative impacts of statistical discrimination through policies that reduce the visibility of criminal records increases crime further. Moreover, such policies cause a greater negative effect for law abiding members of the disadvantaged group than members of the statistically favored group. In addition to these findings, the article demonstrates that ‘self-fulfilling expectations’ cannot exist when there are diminishing returns to deterrence from increasing expected sanctions. This suggests that heterogeneities across groups’ criminal tendencies remains the only explanation for rational statistical discrimination.

Keywords: Statistical discrimination, ban the box, crime, deterrence, racial heterogeneity.

JEL classification: D82, K00, K14, K42, J71, J78.

1. Introduction

Statistical discrimination refers to using individuals’ observable characteristics, such as their race or gender, as inputs in estimating their quality in some relevant dimension, for instance, their expected productivity. This type of discrimination does not explicitly aim to exclude or harm particular individuals, but, instead, is a result of people’s desire to increase their own well-being by incorporating as much information as they can in estimating the value from transacting with another person. Based on this definition, some may claim that statistical discrimination is only as bad as acting with self-interest in market transactions generally. However, numerous studies discuss the potential harms
that may flow from this type of discriminatory behavior, especially harms to those who are statistically discriminated against.\footnote{See, e.g., Fang and Moro (2011) reviewing the economics literature on statistical discrimination.} This article adds to this list of potential harms by providing an explanation of how statistical discrimination may lead to an increase in the overall crime rate.

The dynamics that lead to more crime can be explained as follows. Members of disadvantaged groups often have less access to acquiring the skills necessary to increase their productivity in the labor market, and they may simply have inferior education opportunities. Therefore, employers who engage in statistical discrimination, i.e. consider membership to these groups as a relevant factor in estimating the productivity of individuals, may rationally offer members of disadvantaged groups inferior wages and/or working conditions. Once members of disadvantaged groups begin to expect this type of unequal treatment in the hiring process, they are less likely to invest in increasing their work-productivity and more likely to engage in criminal activity, because they have less to lose (in the form of lower wages) by becoming ex-convicts and having even worse job prospects. Therefore, statistical discrimination is likely to cause an increase in crimes committed by people in disadvantaged groups.

This observation, on its own, does not suffice to conclude that statistical discrimination increases the overall crime rate. This is because the above observations are silent on how statistical discrimination affects the incentives of people who are not members of disadvantaged groups. In a world where statistical discrimination does not exist, an employer pools individuals from all groups together for purposes of estimating their productivity. Therefore, the predicted average wage of individuals in non-disadvantaged groups is lowered. Thus, the reputational cost of being convicted of a crime is lowered for these people, and, they are more likely to commit crime. In sum, statistical discrimination increases the crime rate among people in disadvantaged groups, and has the opposite effect for people who benefit from statistical discrimination. Hence, the effect of statistical discrimination on the aggregate crime rate is, a priori, ambiguous.

This ambiguity vanishes under the following two intuitive conditions. The first condition requires "diminishing deterrent returns" to increasing expected punishment to both groups (Nagin 2013 p. 230). In addition to being intuitive, this assumption is consistent with the observations made in the literature that increases in the expected sanction tend to have a large deterrent effect only if the sanction is not very large to begin with.\footnote{See, e.g., Nagin (2013) pp. 230-232 making these observations in contexts where the punishment is monetary as well as in contexts where the punishment is imprisonment.} The second condition requires the disadvantaged group to be more responsive to policies that increase the expected cost to committing crime, which is likely, since members of this group are the ones who commit crimes more often. These two conditions constitute sufficient (but not necessary) conditions for statistical discrimination to increase the overall crime rate.

It is worth briefly pointing out that the dynamics described above rely on
statistical discrimination which emerges as a result of differences in the opportunities available to the two groups, e.g. inferior education opportunities for the members of the disadvantaged group. However, statistical discrimination may occur even when the two groups have identical characteristics, if it arises as a result of "self-fulfilling expectations" (Rasmusen 1996). This would require employers to have different expectations regarding the behavior of the members of the two groups, which would provide asymmetric incentives to members of these groups. In turn, these asymmetric incentives would need to result in equilibrium behavior that validate employers’ beliefs. The presence of this type of self-fulfilling expectations requires multiple equilibria, since different beliefs regarding identical groups’ behavior must be supportable in equilibrium. Interestingly, the analysis reveals that this type of multiple equilibria cannot exist when there are diminishing returns to deterrence from increased expected sanctions. This is essentially because, when there are diminishing returns to deterrence, an increase in stigma beyond that which supports the first equilibrium is insufficient to reduce the crime rate to the point where the increase in stigma is supportable. Thus, the conditions that imply an increase in crime due to discrimination also imply that statistical discrimination, if it exists, must be due to differences in the characteristics of the two groups. This point also qualifies results regarding the possibility of multiple equilibria in the law enforcement literature by illustrating intuitive conditions under which such equilibria can be ruled out.

Even if statistical discrimination can be shown to increase crime, and even if the reasons for such discrimination can be identified, one may suggest that this type of discrimination is impossible or extremely costly to fully eliminate. Even so, there are at least two reasons why the relationship between statistical discrimination and criminal activity deserves academic interest. First, the problem of statistical discrimination can only exist if membership to the disadvantaged group is relatively salient. Therefore, ceteris paribus, societies with more salient race (or ethnic) heterogeneity are likely to have higher crime rates. This observation is of direct interest to research focusing on the determinants of crime rates across different societies (e.g. Spamann 2016) and is also related to social disorganization theory (Shaw and McKay 1929) which posits that racial heterogeneity is likely to increase crime rates by weakening local informal control mechanisms. Second, there are policies, such as ‘ban the box’ campaigns, aimed at indirectly reducing unfavorable labor market outcomes for members of disadvantaged groups, and recent research suggests that such policies may encourage statistical discrimination (Doleac and Hansen (2017) and Agan and Starr (2016)). Extending the theory of statistical discrimination presented in this article allows an assessment of such policies, and reveals that they are not only likely to cause a further increase in the crime rate, but as suggested by recent empirical research, harm disadvantaged groups more than any other group.

The next section presents a brief literature review. Section 3 presents a law enforcement model where convictions generate stigma (as in Rasmusen (1996)) for members of two distinct groups whose membership may or may not be ob-
servable by third parties. The same section derives sufficient conditions under which statistical discrimination increases crime. Section 4 incorporates governmental policies that may reduce the visibility of criminal records, and shows that such policies are likely to increase crime, and cause more harm to individuals in the disadvantaged group. Section 5 concludes.

2. Literature Review

Discrimination is a topic which has drawn the attention of economists, starting at least as early as Becker (1957). There is an important distinction between what economists call ‘taste-based’ discrimination and statistical discrimination. The former occurs when some people have a preference, unrelated to commonly studied goals such as profit maximization, for transacting with people from a particular group rather than another. Statistical discrimination, on the other hand, occurs when rational utility maximizers use people’s salient characteristics (e.g. their race or gender) to increase the accuracy of their estimates regarding a relevant quality, such as their productivity or likelihood of committing crime. Becker (1957) noted in his analysis of taste-based discrimination that firms which engage in such discrimination will be out-competed by non-discriminating firms in equilibrium. Perhaps because of this powerful argument, most later articles have focused on statistical discrimination. The instant article, too, focuses on statistical, and not taste-based, discrimination.

The seminal articles on statistical discrimination are Phelps (1972) and Arrow (1973), and the large literature that builds on these two articles is reviewed in Fang and Moro (2011). The existing literature focuses on two sources of statistical discrimination. Fang and Moro (2011) describes these sources as follows:

In Phelps (1972), and the literature that originated from it, the source of inequality is some unexplained exogenous difference between groups of workers, coupled with employers’ imperfect information about workers’ productivity. (...) [Whereas in] the literature that originated from Arrow (1973), average group differences in the aggregate are endogenously derived in equilibrium, without assuming any ex-ante exogenous differences between groups. Even in this strand of literature decision makers hold asymmetric beliefs about some relevant characteristic of members from different groups, but the asymmetry of beliefs is derived in equilibrium. This is why these beliefs are sometimes referred to as “self-fulfilling stereotypes”.

In the law enforcement literature, Rasmusen (1996) demonstrates that self-fulfilling expectations may lead to multiple equilibria in a setting where criminal records signal lower productivity. Thus, criminal records generate stigma in the form of a wage penalty.\(^3\) Although Rasmusen (1996) considers a single group, the analysis extends through trivial modifications to a context where there are multiple groups, and, therefore, can provide the basis for statistical discrimination. The analysis presented in this article uses Rasmusen’s (1996) model

\(^3\)A number of articles have analyzed criminal stigmatization since Rasmusen (1996) including Iacobucci (2014), Zasu (2007), Mungan (2016a), and Mungan (2016b).
of stigmatization, and demonstrates that the existence of self-fulfilling expectations requires non-diminishing returns to deterrence from increased expected sanctions. This not only clarifies the conditions under which multiple stigmatization equilibria can exist, but, also suggests that, in this setting, if statistical discrimination exists, it must be due to differences in groups’ characteristics as in Phelps (1972) and the subsequent literature.\textsuperscript{4}

Fang and Moro (2011) also explains that although most of the literature on statistical discrimination analyzes incentive effects, only "a small literature has been devoted to analyzing the different sources of inefficiency arising from statistical discrimination" (Fang and Moro (2011) p. 191).\textsuperscript{5} The instant article adds to the literature also in this dimension by identifying conditions under which statistical discrimination is likely to generate social costs, in the form of increased crime, and thereby cause inefficiencies. Moreover, because statistical discrimination requires salient differences between the two groups, the link between statistical discrimination and increased crime suggests that important (e.g. racial, ethnic or religious) heterogeneities in a countries’ population may contribute to its crime rate through statistical discrimination. Thus, such heterogeneities may partially explain the cross-country variation in crime rates, which is the focus of a sizeable empirical literature (see, e.g., Spamann (2016), Dills et. al (2010) and McCrary and Sanga (2012)). Moreover, this observation also complements the implication of social disorganization theory (Shaw and McKay 1929) that racial heterogeneity is likely to be associated with higher crime rates. Social disorganization theory suggests that this is likely to happen not due to statistical discrimination, but because such heterogeneity is likely to weaken informal control mechanisms within a neighborhood.

Statistical racial discrimination is, of course, a practice that has been criticized extensively, because it causes undue burdens for people simply as a result of their skin color. This type of discrimination is, however, also quite difficult to reduce or eliminate. Thus, policy makers have looked for additional ways to reduce the unequal burdens faced by minorities, most notably African Americans in the United States. As part of this endeavor, many states have launched ‘ban the box’ campaigns, which prohibit certain employers from asking job applicants whether they have a criminal record in job application forms.\textsuperscript{6} Some policy makers have explicitly stated that one of the objectives of these campaigns is to reduce the unemployment rate among African Americans (see Agan and Starr 2016 p. 2). The current article shows that policies which reduce the visibility of criminal records may have the unintended consequence of lowering the expected earnings of law-abiding members of groups that are targeted by statistical discrimination, in addition to increasing the over-all crime rate fur-

\textsuperscript{4}It is worth briefly noting that the model considered here is also similar to Moro and Norman (2004) in that it is a type of general equilibrium model where people make ‘investment decisions’ in the form of whether or not to commit crime. Therefore, it also highlights that when the investment decision takes the particular form studied in this article, multiple equilibria, also identified in Moro and Norman (2004), can be ruled out in some circumstances.\textsuperscript{5}See, section 7 of Fang and Moro (2011) for a review of efficiency related points made in the literature.\textsuperscript{6}See, Agan and Starr (2016) for a more detailed description of these campaigns.
ther. This finding is consistent with recent empirical work investigating the likely unintended consequences of ‘ban the box’ campaigns (Doleac and Hansen (2017) and Agan and Starr (2016)).

3. A Model of Crime with and without Observable Group Membership

Consider a continuum of individuals who have varying benefits \( b \) from committing crime. Suppose further that there are two groups \( H \) and \( M \) (for majority), and that the cumulative distributions among individuals in these two groups are also denoted \( H \) and \( M \). This corresponds to assuming that an individual draws his criminal benefit \( b \) from his group’s cumulative distribution function (CDF) prior to all other events described below. Although \( H \) and \( M \) are common knowledge, a person’s criminal benefit \( b \), is private information.

A person who refrains from criminal activity devotes more time to acquiring skills necessary for legal occupations, and increases his productivity, as a worker, from \( \omega \) to \( \omega + \alpha \).\(^7\) To incorporate the idea that members of group \( H \) have less access to acquiring these skills and/or more access to criminal opportunities, I assume that \( H \) first order stochastic dominates \( M \), i.e. \( H(b) < M(b) \) for all \( b \in [0, \infty) \), and I assume that \( H' = h \) and \( M' = m \) are positive for all \( b \in [0, \infty) \).

The relative size of group \( M \) [and \( H \)] is denoted by \( \phi \) [and \( 1 - \phi \)].

Members of both groups decide whether or not to participate in criminal activity. If a person commits crime, he receives benefit \( b \), but, is caught by law enforcers with probability \( p \), and receives a formal sanction of \( s \). In addition to the formal sanction, when caught, he receives a criminal record, which reduces the wages he receives in the labor market.

Individuals’ wages in the labor market are determined by rational employers who offer each individual his expected productivity given all information he has about the individual. There are two types of information that the employer may have: (i) a person’s criminal background and (ii) his group membership. In this section I assume that employers always, and accurately, see whether a person has a criminal record.\(^8\) However, whether they can observe the person’s group membership depends on whether or not group characteristics are salient. To formalize this idea, I assume that when group membership is observable, an employer knows that an individual from group \( H \) [or \( M \)] has drawn his benefit from \( H \) [or \( M \)]. However, when group membership is not observable, the employer only knows that each individual draws his criminal benefit from the CDF associated with the entire population, namely,

\[
N \equiv \phi M + (1 - \phi) H
\]

since the relative size of group \( M \) is \( \phi \). An alternative way to think about the case where group membership is unobservable is to assume that all individuals belong to group \( N \) which is associated with the CDF defined in (1).

Given the notation introduced, we can denote a person’s criminal record as \( i \in \{\psi, \nu\} \) (\( \psi \) for criminal record and \( \nu \) for no record) and his group membership

\(^7\)This assumption corresponds to that in Rasmusen’s (1996) moral hazard model.

\(^8\)This assumption is relaxed in the next section where I consider the effects of policies that reduce the visibility of criminal records.
as \( J \in \{H, M, N\} \) where \( H \) and \( M \) are group memberships when membership is observable, and \( N \) denotes the entire population when membership is not observable. Thus, a person’s wage, depending on his criminal record and his group membership is given by: \( w_{J \in \{H, M, N\}} \).

A point that is worth highlighting is that, employers engage in statistical discrimination, only when group membership is observable, which enables them to offer different wages to people with similar criminal records from different groups. Thus, in cases where group membership is observable, if employers could somehow be forced to not engage in statistical discrimination, they would be constrained to act as if group membership is unobservable. Therefore, the impact of statistical discrimination, when it exists, can be studied by comparing various outcomes when group membership is observable and when it is not observable. This is one of the reasons for why both cases are introduced. Next, I describe how individuals’ criminal decision making processes and employers’ wage offers interact. These dynamics are very similar to those introduced in Rasmusen’s (1996) moral hazard model.

### 3.1. Individuals’ Decisions

Each individual takes wages as given, because there are a continuum of individuals, which implies that no single individual’s behavior has an impact on employers’ expectations. Thus, a person with group membership \( J \in \{H, M, N\} \) expects a payoff of \( b + p(w_J^\psi - s) + (1 - p)w_J^\nu \) from committing crime and a payoff of \( w_J^\nu \) from refraining from committing crime. Therefore, he commits crime if:

\[
b_J = p(s + w_J^\psi - w_J^\nu) = p(s + \sigma_J) < b
\]

where \( \sigma_J \) is the group-dependent reduction in wages from having a criminal record, which may alternatively be called the stigma associated with a criminal record. Given this group-dependent threshold, the crime rate in both groups can be calculated as:

\[
\theta_J = 1 - J(b_J) \text{ for } J \in \{H, M, N\}
\]

These crime rates depend on people’s expectations regarding the wages to be offered in the market. However, wages are also endogenously determined by rational employers, and, their decision making process is outlined next.

### 3.2. Wages

Wages are offered in a competitive labor market, and thus, they equal an individual’s expected productivity given his group membership and criminal record. Specifically, among individuals in group \( J \), only \( p\theta_J \) proportion of people have criminal records, because they are the only ones who have committed crimes and were caught. Thus, these individuals’ wages are \( w_J^\psi = \omega \), because all criminals have productivity \( \omega \). On the other hand, in group \( j \), among individuals who do not have records, \( (1 - \theta_J) \) have no records only because they were not convicted subsequent to committing crime. Thus, these individuals have an average productivity of

\[
w_J^\nu = \frac{\omega(1 - p)\theta_J + (1 - \theta_J)(\omega + \alpha)}{1 - p\theta_J}
\]
which equals the wages offered to individuals in group $j$ who do not have criminal records. Therefore, the stigma associated with employers’ wage offers is given by:

$$
\sigma_j^o \equiv w_j^o - w_j^\psi = \frac{\omega(1 - p)\theta_J + (1 - \theta_J)(\omega + \alpha)}{(1 - p\theta_J)} - \omega = \frac{1 - \theta_J}{1 - p\theta_J}
$$

(5)

Here, $\sigma_j^o$ denotes the stigma caused by employers’ offers, given the crime rate in each group. As explained in section 3.1., these crime rates depend on the wages people expect to earn with and without criminal records. Thus, an equilibrium emerges when people’s expectations match the actual wages offered by employers. This equilibrium condition is identified next.

### 3.3. (Unique) Equilibrium Characterization

In a Perfect Bayesian Equilibrium, wages offered by employers must equal the wages individuals expect to earn when they make decisions on whether or not to commit crime. Thus, an equilibrium is obtained when

$$
C_J(b^*_J) = b^*_J - p(\sigma_J^o(b^*_J) + s) = 0 \text{ for } J \in \{H, M, N\}
$$

(6)

where $b^*_J$ denotes the equilibrium threshold benefit for individuals in group $J$, and $C_J(b) \equiv b_J - p(\sigma_J^o(b_J) + s)$ is used to characterize equilibria. By focusing on this condition lemma 1, below, notes that an equilibrium exists and identifies conditions under which it is unique.

**Lemma 1:** There exists an equilibrium, and it is unique if $H$ and $M$ are concave.

**Proof:** Note that $C_J(0) = -p(\sigma_J^o(0) + s) = -ps < 0$ and that $\lim_{b \to \infty} C_J(b) = \infty$. Thus, there exists a positive and finite $b^*_J$ such that $b^*_J = p(\sigma_J^o(b^*_J) + s)$ for both $J \in \{H, M, N\}$. Next, let $B^*_J = \{b > 0 | C_J(b) = 0\}$ and note that

$$
\frac{dC_J}{db_J} = 1 - p\frac{d\theta_J}{db_J} \frac{\partial \sigma_J^o}{\partial \theta_J} = 1 - p\theta_J \alpha \frac{1 - p}{(1 - p\theta_J)^2}
$$

(7)

where $j = J'$ is the density function associated with $J \in \{H, M, N\}$, and the second equality follows from (5). Thus,

$$
\frac{d^2C_J}{db_J^2} = -p(1 - p)\alpha \frac{j''(b_J)(1 - p\theta_J) - 2p(j(b_J))}{(1 - p\theta_J)^3}
$$

(8)

which implies that

$$
\frac{d^2C_J}{db_J^2} \geq 0 \text{ iff } j'(b_J) \leq \frac{2p(j(b_J))}{1 - p\theta_J}
$$

(9)

---

\(^9\)This expression is derived in exactly this form (where there is a single group) in Rasmusen (1996). Although other formulations are possible (see Mungan (2016a)), this specification most tractably captures the idea that refraining from committing crime can be seen as an investment.
which holds whenever \( j' < 0 \), i.e. whenever \( J \) is concave. This observation, combined with the fact that \( \lim_{b \to -\infty} C_J(b) = \infty \), implies that \( \frac{dC_J(b)}{db} > 0 \) for all \( b \), and, thus, \( |B_J^*| = 1 \) whenever \( J \) is concave for both \( J \in \{ H, M \} \).

Lemma 1 identifies a sufficient, but not necessary, condition that rules out multiple equilibria. Thus, the result shows that diminishing deterrent effects from expected punishment are enough to rule out the possibility of multiple equilibria considered in Rasmusen (1996), and justify focusing on a single equilibrium. Therefore, the next sections focus on the case where there is a unique equilibrium.

3.4. Crime Rates with and without Observable Membership

When group membership is observable, a natural question to ask is whether there is an unambiguous relationship between the magnitude of stigma faced by members of the two groups. Intuitively, members of group \( H \) are likely to suffer less from stigmatization as a result of receiving a criminal record, because they are likely to be discriminated against to begin with. This implies a higher crime rate in group \( H \) than in group \( M \) due to two reasons. First, people in group \( H \) face lower informal sanctions associated with committing crime, and, thus, they are less deterred than individuals in group \( M \), i.e., in symbols, \( b_H < b_M \) as can be inferred from (2). Second, even if people in groups \( H \) and \( M \) were to face the same amount of stigma, the crime rate in group \( H \) would still be higher, simply because individuals in this group typically have higher criminal benefits, i.e., in symbols, \( \theta_H(b) > \theta_M(b) \) for all \( b \) because \( H \) first order stochastically dominates \( M \). The next proposition formalizes these results. In what follows a * superscript denotes equilibrium values and subscripts always refer to the relevant group.

**Proposition 1:** In equilibrium, (i) the magnitude of stigma associated with conviction is higher for people in group \( M \) than for people in group \( H \) (i.e. \( \sigma_H^* < \sigma_M^* \)), (ii) people in group \( M \) require higher criminal benefits to commit crime than people in group \( H \) (i.e. \( b_H^* < b_M^* \)), (iii) the crime rate within group \( H \) is higher than the crime rate within group \( M \) (i.e. \( \theta_H > \theta_M^* \)). These observations also imply that \( w_A^\psi = w_M^\psi < w_M^\nu < w_A^\nu \).

**Proof:** First, note that \( H(b) = 1 - \theta_H(b) < 1 - \theta_M(b) = M(b) \) for all \( b > 0 \), since \( H \) first order stochastically dominates \( M \). This, in turn, implies that \( \sigma_H^* < \sigma_M^* \), since \( d((1-\theta)/(1-\rho))^2)/d\theta = (1-p)(1-\theta)/(1-\rho)^2 = \frac{b_H^*}{1-p} < 1 \), which implies that \( \sigma^\rho \) is decreasing in \( \theta \).

(ii) Thus, per (6), it follows that \( b_H^* < b_M^* \).

(iii) \( \theta_M^* = 1 - M(b_M^*) < 1 - H(b_M^*) = \theta_H^* \), since \( M(b_M^*) > M(b_H^*) > H(b_H^*) \) where the first inequality follows from part (ii), and the second inequality follows from first order stochastic domination.

(i) Since \( \theta_H^* > \theta_M^* \), it follows that \( \sigma_H^* = \sigma^\rho(\theta_H^*) < \sigma_M^*(\theta_M^*) = \sigma_M^* \), because, as demonstrated in the beginning of this proof \( \sigma^\rho \) is decreasing in \( \theta \).

Proposition 1 validates intuitive conjectures when group membership is observable. A natural question to ask is what happens to the magnitude of stigma and crime rates in the absence of statistical discrimination, or equivalently, when group membership is unobservable. Intuitively, when employers cannot observe
group membership, they form milder expectations regarding the single group that they are confronted with, and, thus, individuals face an amount of stigma that lies in between the amount faced by members of the two groups under statistical discrimination. The result is an intermediate crime rate. Proposition 2 formalizes this result.

**Proposition 2:** $\sigma_H^* < \sigma_N^* < \sigma_M^*$, and, therefore $b_M^* > b_N^* > b_H^*$ and $\theta_M^* < \theta_N^* < \theta_H^*$.

**Proof:** $N(b) = \phi M(b) + (1-\phi)H(b)$ implies that $\theta_H(b) > \theta_N(b) > \theta_M(b)$ for all $b > 0$, since $H$ first order stochastically dominates $M$. This, in turn, implies that $\sigma_H^*(\theta_H(b)) < \sigma_N^*(\theta_N(b)) < \sigma_M^*(\theta_M(b))$ for all $b$, since $\sigma^\circ$ is decreasing in $\theta$ as shown in proposition 1. This, in turn, implies per (6) that $b_H^* < b_N^* < b_M^*$, and, therefore, $\sigma_H^* < \sigma_N^* < \sigma_M^*$ and $\theta_M^* < \theta_N^* < \theta_H^*$. 

Although proposition 2 demonstrates that the crime rate among group $H$ goes down and the crime rate among group $M$ goes up when statistical discrimination is eliminated, the proposition is silent on what happens to the total crime rate. The next section investigates the total crime rate.

### 3.5. The Impact of Statistical Discrimination on Total Crime

When group membership is observable and statistical discrimination is present, the total crime rate is given by:

$$\theta_S^* \equiv \phi \theta_M^* + (1-\phi)\theta_H^* = \phi[1-M(b_M^*)] + (1-\phi)[1-H(b_H^*)]$$

(10)

since $\phi$ denotes the relative size of group $M$. However, in the absence of statistical discrimination (or observable groups), the crime rate is given by:

$$\theta_N^* = \phi[1-M(b_N^*)] + (1-\phi)[1-H(b_N^*)]$$

(11)

because having a criminal record results in the same amount of stigma for members of both groups, and, therefore, the threshold criminal benefit is common and equal to $b_N^*$ in both groups. Therefore, statistical discrimination leads to more crime, if

$$\phi M(b_N^*) + (1-\phi)H(b_N^*) > \phi M(b_M^*) + (1-\phi)H(b_M^*)$$

(12)

which can equivalently be expressed as:

$$(1-\phi)[H(b_N^*) - H(b_M^*)] > \phi[M(b_M^*) - M(b_N^*)]$$

(13)

The inequality in (13) simply corresponds to the case where statistical discrimination increases the number of crimes committed by members of group $H$ by more than it reduces the number of crimes committed by members of group $M$. There is, a priori, no reason to suggest that this condition holds. However, because people in group $M$ have, on average, lower criminal benefits from crime than people in group $H$, relatively small sanctions (formal + informal) are likely to deter most of these individuals. Therefore, in equilibrium, this group is likely to be the less responsive of the two groups to changes in expected punishment. An assumption that formalizes this idea is that $h > m$ for all relevant policies (i.e. for all $b \in [b_H^*, b_M^*]$). As stated in the next proposition, this condition
coupled with an assumption of weak concavity of $H$ and $M$ over the relevant range of criminal benefits provides sufficient conditions under which statistical discrimination leads to more crime.

**Proposition 3:** Statistical discrimination increases crime, if $h(b) > m(b)$, and $H$ and $M$ are (weakly) concave for all $b \in [b_H^*, b_M^*]$.

**Proof:** See Appendix.

**Figure 1:** $\theta_h^* > \theta_m^*$ when $M$ and $H$ are concave and $h > m$. 
The proof of the proposition is relegated to an appendix. However, figure 1, above, graphically depicts some of the steps in the proof. \( S \), the dashed line connecting the points \((b_H^*, H(b_H^*))\) and \((b_M^*, M(b_M^*))\), represents convex combinations of these points. Therefore, the height of this line represents the level of deterrence obtained when group membership is salient, for any chosen level of \( \phi \) (which determines the weight assigned to the point \((b_M^*, M(b_M^*))\)). On the other hand, the level of deterrence when group membership is unobservable is depicted by the line labeled \( N \). This line is concave and lies above \( S \) under the conditions specified in proposition 3. Moreover, the same conditions imply that \( b_N^* > \phi b_M^* + (1 - \phi) b_H^* \). Thus, deterrence is greater when group membership is unobservable due to two reasons. First, \( N(\phi b_M^* + (1 - \phi) b_H^*) > \phi M(b_M^*) + (1 - \phi) H(b_H^*) \), i.e. if group membership were unobservable and people committed crime only if their benefits exceeded the average threshold criminal benefits obtained under observable membership, there would be less crime than among the corresponding population where group membership is salient. Second, \( b_N^* > \phi b_M^* + (1 - \phi) b_H^* \), i.e. a switch from observable group membership to non-observable group membership increases the average threshold benefit people require to commit crimes, which implies that deterrence is enhanced further.

The analysis in this section focuses on the potential impact of statistical discrimination on the incentives of individuals in different social groups to commit crimes, and, it also derives intuitive sufficient conditions under which such discrimination may lead to more crime. The absence of statistical discrimination is viewed as a situation in which either group membership is not observable, which makes statistical discrimination impossible, or, a situation in which employers either do not, or are forced to not, engage in such discrimination. Thus, the entire analysis revolves around whether information in one dimension, namely group membership, is visible or usable. The next section focuses on the visibility of information regarding the second type of information considered in the model, namely criminal records.

4. A Model of Crime with Different Degrees of Criminal Record Visibility

In this section, I focus exclusively on a population which consists of two groups for which membership is observable, and suppose that the government can enact policies that affect the likelihood with which a person’s record affects his wage (e.g. ban the box). In the previous section, this probability was implicitly assumed to be 1. Instead, suppose now that it is \( v \in (0,1] \). In this case, a person’s benefit from committing crime is, \( b + p(v w_J^v + (1 - v) w_J^v) \) and his benefit from not committing crime is \( w_J^v \). Thus, a person commits crime if:

\[
 b > p[v \sigma J + s] = b_J^v 
\]  

(14)

Re-calculating wages as in the previous section reveals that stigma is given by:

\[
 \sigma_J^v = \frac{1 - \theta J}{1 - pv \theta J} 
\]  

(15)

12
Thus, the equilibrium cut-off benefits are:

\[ b_J^* = p(v\alpha \frac{(1 - \theta_J)}{1 - pv\theta_J} + s) \text{ for } J \in \{H, M\} \]  

(16)

As (16) demonstrates, holding the crime rate constant, expected stigma is increasing in the visibility of criminal records. This is partly because visibility reduces the proportion of criminals without observable records (and thereby increases the average productivity of individuals without visible records), but, also because visibility increases the odds with which a person is stigmatized for having a record. This simple observation has a number of important implications which are summarized by proposition 4.

**Proposition 4:** (i) Reducing the visibility of criminal records increases the crime rate in both groups, i.e. \( \frac{\partial b_J^*}{\partial v} < 0 \). (ii) This leads to a reduction in the wages of individuals without visible criminal records.

**Proof:** (i) As can be inferred from (16) \( \frac{\partial b_J^*}{\partial v} > 0 \), this implies that \( \frac{\partial \sigma_J^*}{\partial v} < 0 \) for \( J \in \{H, M\} \). (ii) \( \frac{\partial \sigma_J^*}{\partial v} = \frac{\partial \sigma_J^*}{\partial b_J^*} + \frac{\partial \sigma_J^*}{\partial \theta_J} \frac{\partial b_J^*}{\partial v} > 0 \), since \( \frac{\partial \sigma_J^*}{\partial \theta_J} > 0 \) as can be inferred from (15), and, thus, \( \frac{\partial \sigma_J^*}{\partial v} = \frac{\alpha(1-pv)j(b_J^*)}{(1-pv\theta_J)^2} \frac{\partial b_J^*}{\partial \theta_J} > 0 \), which implies that \( \frac{\partial \sigma_J^*}{\partial \theta_J} > 0 \). Therefore, \( w_J^* = \sigma_J^* + \omega \) is increasing in \( v \).

Given recent research on the effects of policies such as the ‘ban the box’ (BTB) campaign, an important question is whether policies that tend to reduce the visibility of criminal records are likely to have differential effects on the members of groups \( H \) and \( M \). In particular, Agan and Starr (2016) find that BTB increases the gap between the callbacks that similarly situated white and black job applicants receive. They suggest that this may be because employers engage in more statistical discrimination in response to BTB policies. The next proposition shows that reducing the visibility of criminal records is likely to harm members of group \( H \) more than members of group \( M \), even when employers always statistically discriminate between members of these groups.

**Proposition 5:** A reduction in the visibility of criminal records leads to a greater percentage reduction in the wages of people without (visible) criminal records in group \( H \) than the percentage reduction in the wages of people without (visible) criminal records in group \( M \), as long as wages for people with visible criminal records (i.e. \( \omega \)) is sufficiently low.

**Proof:** Note that

\[
\frac{\partial w_J^*}{\partial v} = \frac{\partial \sigma_J^*}{\partial v} + \frac{\partial \sigma_J^*}{\partial b_J^*} \frac{\partial b_J^*}{\partial v} = \frac{p\theta_J(1 - \theta_J)\alpha}{(1 - pv\theta_J)^2} + \frac{\alpha(1 - pv)j(b_J^*)}{(1 - pv\theta_J)^2} \frac{po(1 - \theta_J)}{1 - pv\theta_J^2 - (1 - pv)pvoj(b_J^*)} \]  

(17)

Thus,

\[
\left. \frac{\partial w_J^*}{\partial v} \right|_{\omega=0} = \frac{p\theta_J(1 - \theta_J)\alpha}{(1 - pv\theta_J)^2} + \left( \frac{\alpha(1 - pv)j(b_J^*)}{(1 - pv\theta_J)^2} \frac{po(1 - \theta_J)}{1 - pv\theta_J^2 - (1 - pv)pvoj(b_J^*)} \right) \]  

(18)
which simplifies to:

\[
\frac{\partial w_j^v}{\partial v} \frac{1}{w_j^v} |_{\omega=0} = p \frac{\theta_j}{1 - pv \theta_j} + \frac{\alpha(1 - pv)j(b_j^v)}{(1 - pv \theta_j)((1 - pv \theta_j)^2 - (1 - pv)pv \alpha_j(b_j^v))} = 0
\]

From (19) it can easily be observed that

\[\frac{\partial w_j^v}{\partial v} \frac{1}{w_j^v} |_{\omega=0} > 0\] and

\[\frac{\partial w_j^v}{\partial v} \frac{1}{w_j^v} |_{\omega=0} > 0.\]

This implies that \(\frac{\partial w_j^v}{\partial v} \frac{1}{w_j^v} |_{\omega=0} < \frac{\partial w_j^v}{\partial v} \frac{1}{w_j^v} |_{\omega=0}\) since \(h(b_h^*) > m(b_m^*)\) and \(\theta_h^* > \theta_M^*.\) This implies that the same relationships hold for sufficiently small but positive \(\omega.\)

Proposition 5 demonstrates that policies like BTB can have greater negative effects for law abiding people in disadvantaged groups, even if the frequency of statistical discrimination remains unchanged before and after policy changes. Thus, the proposition highlights that the magnitude of the negative consequences associated with statistical discrimination may matter just as much as the frequency with which statistical discrimination takes place.

5. Conclusion

Statistical discrimination is likely to cause great harms to people in disadvantaged groups, and, moreover, it reduces the incentives of these individuals to invest time and effort into becoming more productive members of society. Similar dynamics that remove or reduce the incentives of these individuals to invest in productive human capital may also cause them to commit crimes more frequently. A corollary of this observation is that racial heterogeneity may be one of the determinants of higher crime rates in societies, due to its crime inducing effect through statistical discrimination. Although the harms from statistical discrimination may be large, the analysis in this article suggests that one should exercise caution in designing policies to mitigate such harms. In particular, attempting to indirectly reduce the gap between unemployment rates across different groups by reducing the visibility of criminal records may have unintended consequences in the form of higher crime rates and a widening of the gap between the labor market prospects of individuals in different groups.

Appendix

The proof of proposition 3 relies on a lemma which makes use of the following observation. Let \(b^c(x)\) be defined as

\[
b^c(x) \equiv p \left(\alpha \frac{L^x(b^c)}{(1 - p) + pL^x(b^c)} + s\right) \text{ for all } x \in [0,1]
\]

Here, \(L^x = xM(b^c) + (1 - x)H(b^c)\) is a fictitious CDF synthesized by combining members of the two groups with different weights reflected by \(x.\) Thus, it follows, from (5) and (6), that \(b^c(x)\) is characterized by the condition:

\[
C_{L_x}(b^c) = b^c - p(\sigma_{L_x}^c(b^c) + s) = 0
\]

and, therefore,

\[
b_N^* = b^c(\phi), b_M^* = b^c(1) \text{ and } b_H^* = b^c(0)
\]
The properties of $b^c$ play a key role in proving proposition 3, which are formalized by the following lemma.

**Lemma 2**: $b^c$ is increasing and concave in $x$ whenever $M$ and $H$ are concave and $m < h(b)$.

**Proof**: Since $\frac{dC_i}{db^c} > 0$, as explained in lemma 1, it follows that $\frac{db^c}{dx} = -\frac{\partial C_{L^*}}{\partial x} > 0$ if

$$\frac{\partial C_{L^*}}{\partial x} = -\frac{(1 - p)p\alpha(M - H)}{(1 - p) + pL^*(b^c)} < 0$$

(A.4.)

which is true since $M > H$. Thus, $\frac{db^c}{dx} > 0$. To obtain the second derivative of $b^c$ with respect to $x$, the expression for $\frac{\partial C_{L^*}}{\partial x}$ in (7) can be used to express $\frac{d^2 b^c}{dx^2}$ more specifically as:

$$\frac{db^c}{dx} = \frac{(1 - p)p\alpha(M - H)}{1 - pl^x\alpha\frac{1 - p}{K^2}} = \frac{(1 - p)p\alpha(M - H)}{K^2 - (1 - p)p\alpha L^x}$$

(A.5.)

where

$$K \equiv (1 - p) + pL^x(b^c)$$

(A.6.)

is used to abbreviate expressions. Differentiating again, we have that

$$\frac{1}{p(1 - p)\alpha} \frac{d^2 b^c}{dx^2} = \frac{db^c}{dx}$$

(A.7.)

$$\frac{d^2 b^c}{dx^2} = \frac{db^c}{dx} \left( (1 - p)p\alpha L^x - (M - H) \left( 2K \frac{dK}{dx} - (1 - p)p\alpha (m - h) + \frac{db^c}{dx} \frac{d^2 x}{dx^2} \right) \right)$$

Therefore, $\frac{d^2 b^c}{dx^2} < 0$ if:

$$\frac{db^c}{dx} \frac{M - h}{M - H} \left( (1 - p)p\alpha L^x + (1 - p)p\alpha (m - h) + \frac{db^c}{dx} \frac{d^2 x}{dx^2} \right) < 2K \frac{dK}{dx}$$

(A.8.)

The left hand side of (A.8.) is negative whenever $m < h$ and $m', h' < 0$, since then $\frac{db^c}{dx} > 0$, $\frac{d^2 b^c}{dx^2} = x_m' + (1 - x)h' < 0$, and $K^2 > (1 - p)p\alpha L^x$ as explained in (7). Therefore, $\frac{d^2 b^c}{dx^2} < 0$ since $\frac{dK}{dx} = \frac{db^c}{dx} p\alpha L^x + (M - H) > 0$.

**Proof of Proposition 3**:

Note that the concavity of $b^c$ implies that $b_N^* > \phi b_M^* + (1 - \phi) b_H^*$. Therefore,

$$N(b_N^*) > N(\phi b_M^* + (1 - \phi) b_H^*)$$

(A.9.)

$$= \phi M (\phi b_M^* + (1 - \phi) b_H^*) + (1 - \phi) H (\phi b_M^* + (1 - \phi) b_H^*)$$

$$> \phi (\phi M (b_M^* + (1 - \phi) M (b_H^*))) + (1 - \phi) (\phi H (b_M^*) + (1 - \phi) H (b_H^*))$$

where the second inequality follows from the concavity of $M$ and $H$. Next, note that

$$\phi (\phi M (b_M^*) + (1 - \phi) M (b_H^*)) + (1 - \phi) (\phi H (b_M^*) + (1 - \phi) H (b_H^*))$$
which holds whenever $h > m$ for all $b \in [b_H^*, b_M^*]$. Thus, (A.9.) and (A.10.) imply that $\phi M(b_M^*) + (1 - \phi) H(b_H^*) > \phi M(b_M^*) + (1 - \phi) H(b_H^*)$, which is the condition in (12).

References


