Empirical Implications of Statistical Discrimination on the Returns to Measures of Skill*

Andrea Moro[§] and Peter Norman[¶]

October 2003

Abstract

We investigate how lack of information may bias the investigator's assessment of the presence of statistical discrimination. We show that the nature of the bias is such that statistical discrimination may be rejected in a Mincerian regression even when the data is generated from an equilibrium with statistical discrimination. This may occur even when the investigator has a more informative signal of productivity the employers have.

1 Introduction

Few empirical regularities are as well documented as the black-white wage gap and the gender wage gap. Some researchers interpret the observed wage gaps as evidence of prejudice. Others think the main source of inequality are some "intrinsic" differences. Still others believe that *statistical discrimination* is an important factor.

The term statistical discrimination refers to models where differential treatment of groups is driven by informational imperfections that create incentives for rational firms to use race or gender as a proxy for productivity. In this article, we will only discuss the more modern version of the theory of statistical discrimination, where there are feedback effects between the beliefs held

*We thank Ananth Seshadri, Dan Silverman, Erzo G.J. Luttmer, two anonymous referees and the editor, Yves Zenou, for helpful discussions and suggestions. Support from NSF Grants #SES-0003520 and #SES 0001717 is gratefully acknowledged.

[§]Department of Economics, University of Minnesota. Email: amoro@atlas.socsci.umn.edu

[¶]Department of Economics, University of Wisconsin-Madison. Email: Pnorman@facstaff.wisc.edu

by firms and workers' incentives to accumulate human capital. Such models usually assume that groups are identical in everything but an "irrelevant" label (typically interpreted as race or gender). The group characteristic is a useful proxy only if groups behave differently in equilibrium. Hence, income inequality between groups is generated purely as an equilibrium phenomenon. Members of a discriminated group are paid lower wages and are more often assigned to low-skilled jobs because employers assume that the discriminated group have less human capital on average. Moreover, the returns to acquire skills for the discriminated group are lower than for other workers, which makes the "stereotyping" by the employers consistent with a rational expectations equilibrium. Crudely put, workers from the discriminated group don't invest in skills because they will be assumed to have low skills no matter what (Coate and Loury [9], Moro and Norman [15],[16]).[1]

If statistical discrimination of this form contributes to the black-white wage gap it must be that:

1. blacks have less human capital than whites;
2. the return on investments in human capital are lower for blacks than for whites.

There is arguably some consensus that racial differences in human capital in the U.S. are as predicted by the model, and it may appear straightforward to estimate the returns to skills for the different groups by, say, returns to schooling. However, confronting the implications of the model with returns to observable measures of skill misses the point with the theory. The fundamental reason that race or sex is useful to employers (in an equilibrium with statistical discrimination) is that productivity cannot be perfectly observed. Years of schooling can be observed, so all workers should have returns for an additional year of schooling pinned down by how much this improves expected productivity.[2] The theory is thus about returns to *imperfectly observable* human capital investments, such as how much effort an individual puts into his or her education. If human capital

[1]Assuming that groups are identical is probably unrealistic. Differences in family backgrounds, the quality of local public schools and many other factors are plausible sources for asymmetries between groups. Our view is that a crucial question is why the racial wage gap is so persistent (the black-white wage gap has increased slightly during the last twenty years). An explanation that relies on socioeconomic characteristics measured during childhood cannot address this question in the absence of a theory for why black "background characteristics" don't improve over time. While admittedly stylized, a model of statistical discrimination that ignores plausible real world asymmetries provides a simple theory for the lack of convergence.

[2]If schooling has "signalling value" there could be statistical discrimination on the basis of schooling. This does however not change the point: whether returns to schooling are higher or lower for blacks than for whites is a separate issue from whether returns to unobservable human capital investments are higher or lower for blacks than for whites.

investments are made by altruistic parents, we may also think about investments in basic skills like punctuality, work ethic, self-control and other social skills.[3]

In this article, we show that even if the theory is about *unobservable* human capital investments, it has implications on returns to *observables.* We develop two versions of the model where the relation between perceived returns to human capital is a direct scaling of the observed returns to the proxy for productivity used by the firms. The scaling factor is group-independent, so the incentive to invest is higher for whites than for blacks if the return to the proxy for skill is higher for whites than for blacks and vice versa. Differences in returns to investments in human capital will thus be reflected by differences in the coefficient on the proxy for skill. That is, if blacks have worse incentives to accumulate skills, as should be the case if statistical discrimination is part of the explanation for the wage gap, then the "return to the skill proxy" should also be lower.

It is crucial for the argument above that *the proxy for skill used by the empirical researcher is the same as the one used by the firms.* We don't think this is particularly realistic. Our analysis will therefore focus on what we view as a more realistic case, where the researcher has access to a signal which is different from the signal observed by the firms. Our belief is that this setup is a good description of the situation of a researcher working with the NLSY dataset, or any other dataset with wage data and measures of cognitive ability such as the Armed Forces Qualification Test (AFQT).

To fix ideas, suppose that productivity is a one dimensional variable and that the AFQT test score is a noisy signal of this true underlying productivity. Furthermore, assume that firms use some other noisy indicator of skill as a proxy for productivity. Also suppose that the econometrician runs a wage regression where the AFQT test score enters as an explanatory variable. It should not be surprising that the coefficient on the AFQT test is an inconsistent estimator of the true returns to human capital if the data is generated by the model of statistical discrimination. Due to measurement error, estimates of the return to AFQT are downwards biased estimates of the returns to the proxy used by firms (which in itself must be scaled up to get a measure of the actual returns). This is neither unexpected or a problem in itself.

What is problematic is that the size of the bias may be different for different groups. We

[3]Much recent empirical literature has emphasized that such basic skills are crucial determinants for earnings. See for example Cawley et al [8].

demonstrate that this may lead the investigator to reject statistical discrimination in an arbitrary large sample even with data generated from an equilibrium with statistical discrimination. The relative precision of the signals is irrelevant, so this may occur even if the signal the investigator is basing the regressions on is a more accurate signal of ability than the signal(s) used by firms.

We show that a model of statistical discrimination can match the main *qualitative* features of the NLSY dataset. Whether the model holds up quantitatively is an open question that we do not resolve in this article. The main obstacle to a serious calibration of the model is that using moments from the observed distribution of wages alone leaves too many free parameters.

This article is related in spirit to a growing empirical literature that is investigating whether there is empirical support for statistical discrimination in various contexts.[4] However, we are only aware of two previous articles discussing differences in rates of returns to skills. Both of them argue that incentives cannot explain the differences between groups. In their well-known study of the black-white wage gap, Neal and Johnson [17] use NLSY data to show (see Table 2 in their article) that the interaction of AFQT with race produces insignificant coefficients. They conclude that there is no support for statistical discrimination, as the returns to AFQT are no different between groups. While this conclusion may be right, our analysis shows that with a standard model of statistical discrimination and using what we view as natural informational assumptions, the step between returns to AFQT and returns to human capital involves a leap of faith. Persico, Postlewaite and Silverman [21] investigate wage inequality between men with different height. They find that current height does not significantly affect wages if height at adolescent age is included as a regressor. Hence, factors that affect youth development, such as self-esteem, participation in extracurricular activities, (which are correlated with height at young age) affect future wages. However, they argue against the hypothesis that this is due to statistical discrimination, since the returns to different measures of self-esteem are no significantly different between short and tall young men.

The rest of the article is organized as follows. In Section 2 a relatively general model of statistical

[4]In the context of labor economics, a few examples are Altonji and Pierret [3], Altonji and Blank [2], Antonovics [4], Foster and Rosenzweig [11],[12], Moro [14], Neumark [18], Oettinger [19]. There is also a significant literature on statistical discrimination in mortgage lending. Bowlus and Eckstein [6] instead disentangle intrinsic skill differences from “taste-based” discrimination.

discrimination is discussed, and Sections 3 and 4 discuss two special cases where the consequences of measurement error in the proxy for skill observed by the researcher can be discussed. Section 5 discusses the patterns in NLSY against the background of the model and Section 6 concludes.

2 A Model of Imperfectly Observable Human Capital Investments

In order to relate the models presented in sections 3 and 4 to the literature we now briefly discuss a somewhat general model of statistical discrimination. Special cases of the setup include models considered in Coate and Loury [9], Moro and Norman [15], [16], and Lundberg and Startz [13].

Consider an economy with a unit mass of workers indexed by $i \in I \subset R$. Each worker belongs to one of two identifiable groups, B or W, and we denote with λ^J the proportion of workers that belongs to group J. Prior to entering the labor market each worker makes a costly human capital investment. We denote the set of possible levels of human capital by H, a generic element by h, and assume that preferences over income (private consumption) and human capital investments for an agent with index i are represented by $u(w) - C(h; i)$. It is easy to allow the distribution of i to be group specific, and for empirical analysis this may be desirable. The cleanest formulation, however, is to assume that there exists some distribution G over I such that the distribution of the cost parameter is equal to G in both groups.

With respect to human capital accumulation by workers, the only departure from standard models in the spirit of Becker [7] is that firms cannot *perfectly* observe productivity. Workers are therefore rewarded on the basis of their *expected* rather than actual human capital. This may seem like a trivial modification, but it changes the analysis in important ways. The reason is that firms must make inferences about individual workers, so any prior knowledge about the distribution of human capital in each group will be used. In particular if firms know that blacks and whites follow different investment strategies, this knowledge will affect the probability estimates of worker productivity.

This signal extraction problem is modeled by assuming that there is a one-dimensional variable $\theta \in \Theta$ summarizing all of the observable information about a worker. When θ is a continuous variable (θ is a binary variable in Section 3) we let $f(\theta|h)$ denote the density of signals conditional on h. Since the choice h is unobservable (and mechanisms inducing self-selection are ruled out by assumption) wages, hiring decisions, and job assignments can be conditioned only on θ and J.

Changes in the distribution of human capital affect posterior beliefs about worker productivity, so equilibrium wages will depend on the distribution of human capital. Exactly *how* equilibrium wages are related to the distributions depends on assumptions about the production technology and the form of competition in the labor and product markets. For now, we treat the labor market as a black box. In Sections 3 and 4 we replace this black box with the simplest possible model of a competitive labor market.

Regardless of the exact equilibrium conditions we observe that $\mathcal{W} = \{w | w : \Theta \to R\}$ is the set of all possible wage schemes. Writing $\Delta(H)$ for the set of distributions over H we may thus think of the labor market equilibrium conditions as a mapping $E = \left(E^B, E^W\right)$, where $E^J : \Delta(H) \times \Delta(H) \to \mathcal{W}$ for $J = B, W$. For any wage profile $w^J \in \mathcal{W}$ the problem for the individual worker (i, J) is to choose the optimal level of human capital solving:

$$\max_{h \in H} \int_{\theta \in \Theta} u\left(w^J(\theta)\right) f(\theta | h)\, d\theta - C(h; i). \tag{1}$$

At this general level there is of course nothing that ensures that there is a unique optimal level of human capital for every worker (i, J). However, suppose for simplicity that this is the case, and for each $w^J \in \mathcal{W}$ let $\eta\left(i; w^J\right)$ be the unique solution to (1), The unique distribution of human capital in group J that is consistent with rational choice by workers when they face wage scheme w^J is then given by $R\left(\cdot; w^J\right)$, where $R\left(h; w^J\right) = \int_{i:\eta(i;w^J) \leq h} dG(i)$, for every $h \in H$. If (1) has a unique maximizer for all i we thus know there is a unique distribution of human capital consistent with w^J. If there are multiple solutions to (1), then each solution generates a distribution of human capital, so in general optimal investment behavior is a correspondence $V = \left(V^B, V^W\right)$, where $V^J : \mathcal{W} \to \Delta(H)$.

Combining optimal investments and labor market equilibrium conditions we obtain a composite map $V \circ E : \Delta(H) \times \Delta(H) \to \Delta(H) \times \Delta(H)$. An equilibrium of the model can therefore be thought of as a distribution of human capital that is a fixed point under the mapping $V \circ E$. The interpretation is that the equilibrium is a distribution of human capital in each group. Such distribution is generated by optimal investment decisions given a wage scheme which is consistent with firms best responding to the equilibrium distribution of human capital.

3 A Variant with a Binary Investment Choice

To fix ideas we will first consider a version of the model where the investment is a binary choice. This model is too stylized to be used as a serious tool in measurement, but it illustrates the problems arising when the econometrician and the firms use different sources of information.

We let $H = \{0, 1\}$ where $h = 0$ is interpreted as "no investment" (or "low investment") and $h = 1$ as "investment" (or "high investment"). By normalization we may then take the cost of the no investment to be zero, that is $C(0; i) = 0$ for all $i \in I$. Moreover, there is no further loss of generality to assume that $C(1, i) = i$ for all $i \in I$, so that the index of an individual worker refers to her (extra) cost to acquire the high level of human capital.

We refer to workers who invest in human capital as "qualified" workers. Qualified workers productivity is equal to 1, and workers who are not qualified do not contribute at all to output.

We also assume that the noisy signal θ that firms observe is a binary variable. That is, $\Theta = \{\theta_0, \theta_1\}$,where

$$\Pr[\theta = \theta_1 | h = 1] = p_q > p_u = \Pr[\theta = \theta_1 | h = 0]. \tag{2}$$

Hence, θ_1 is "good news" about the productivity of the worker since qualified workers are more likely to draw this signal. The labor market is assumed to be competitive. Each worker is thus paid her expected marginal productivity conditional on the signal and group identity.

3.1 Equilibrium

Since human capital takes on only two values, we may describe the distribution of human capital in group J as a single number π^J, which denotes the fraction of qualified workers. Employers know this proportion in equilibrium, so π^J will be the relevant prior probability of investment for a worker from group J. Qualified (unqualified) workers have productivity 1 (0), so the expected productivity coincides with the posterior probability of being qualified. Competitive wages can thus be calculated applying Bayes' rule:

$$\begin{aligned} w^J(\theta_1; \pi^J) &= \frac{\pi^J p_q}{\pi^J p_q + (1 - \pi^J) p_u} \\ w^J(\theta_0; \pi^J) &= \frac{\pi^J (1 - p_q)}{\pi^J (1 - p_q) + (1 - \pi^J)(1 - p_u)}. \end{aligned} \tag{3}$$

The expressions above summarize the labor market equilibrium conditions for the model.[5] What remains is to determine which proportions π^J are compatible with optimal investment behavior given that wages are given by (3). Assuming for simplicity that $u(w) = w$ the decision problem for an individual worker may be written as

$$\max_{h \in \{0,1\}} \left\{ E\left[w^J(\theta; \pi^J) | h = 1\right] - i, E\left[w^J(\theta; \pi^J) | h = 0\right] \right\},$$

where the first term in the parenthesis is the expected utility if the worker becomes qualified. To collapse down the equilibrium conditions it is convenient to let $I\left(\pi^J\right)$ be the *incentive to invest,* that is, the difference in expected earnings for a worker who invested and one who did not,

$$\begin{aligned} I\left(\pi^J\right) &= E\left[w^J(\theta; \pi^J) | h = 1\right] - E\left[w^J(\theta; \pi^J) | h = 0\right] \\ &= p_q w^J(\theta_1; \pi^J) + (1 - p_q) w^J(\theta_0; \pi^J) - \left[p_u w^J(\theta_1; \pi^J) + (1 - p_u) w^J(\theta_0; \pi^J)\right] \\ &= (p_q - p_u)\left[w^J(\theta_1; \pi^J) - w^J(\theta_0; \pi^J)\right] \end{aligned} \quad (4)$$

One can check that if $0 = p_u < p_q < 1$, then $I(\cdot)$ is strictly decreasing, if $0 < p_u < p_q = 1$, then $I(\cdot)$ is strictly increasing, and if $p_u = 0$ and $p_q = 1$, then $I(\cdot)$ is constant in π^J. These cases (with revealing signals) are somewhat special. In all other cases $I(\cdot)$ is a hump-shaped function.

Assuming $0 < p_u < p_q < 1$ it is easy to verify that $I(0) = I(1) = 0$. That is, when everyone behaves in the same way and there is some error in the signal, employers simply choose to ignore the signal and only rely on the prior information. For intermediate values of π^J, the signal is informative, i.e. higher signals are more likely to have invested in human capital, hence will receive higher wages, which implies positive incentives to acquire human capital. The proportion of group J workers who will rational invest given that firms correctly believe that a proportion π^J invests

[5] The reader may notice that *only investments in the own group is relevant in (3).* This feature comes from the assumption that the productivity of individual workers depend only on their skills and not on the relative abundance of qualified and unqualified workers. When there is separability between groups, statistical discrimination becomes a "coordination failure theory". In Moro and Norman [16] we argue that this separability is unappealing and show that by introducing a neoclassical production function of "skilled" and "unskilled" labor, the separability disappears. The model with curvature in production creates incentives for groups to specialize, which results in a more interesting and plausible model of discrimination. In this paper, our primary focus in this paper is to discuss how to confront the model with data on returns to ability. For simplicity we have therefore chosen to consider specifications where groups are separable despite the shortcomings of such models.

is thus $G\left(I\left(\pi^J\right)\right)$. Hence, a proportion π^J of investors in group J is consistent with equilibrium if and only if it solves $\pi^J = G(I(\pi^J))$.

3.2 Statistical Discrimination

It is not difficult to construct examples such that there are multiple equilibria, which differ by the proportion of investors and are Pareto rankable with the equilibrium with the highest level of investments being the most desirable equilibrium. The crucial feature that opens up the possibility for multiple equilibria is an informational externality. A useful way to think of this externality is to observe that employers' perception of a group is like a public good. Individual workers, however, ignore the feedback from investments to the perception about the group, which creates an obvious "under investment problem" as well as a possibility of multiple equilibria.

If the fixed point equation has multiple solutions it is possible for the two groups to coordinate on different solutions, in which case we say that there is income inequality between groups in equilibrium "generated" by statistical discrimination.

3.3 Implications for Observable Returns to Skills

The most fundamental assumption in a model of statistical discrimination is that firms cannot observe productivity perfectly. It seems unlikely that the researcher has perfect information about workers productivity, so in order to take the theory seriously it seems to us that the empirical strategy must deal with imperfect measures of productivity in the data.[6]

A very basic observation is that the model has empirical implications also if actual productivity is unobservable. Suppose that the researcher has access to the information available to the firms. Also, suppose that the imperfect measure of productivity is racially unbiased, so that the probability of drawing θ_1 depends only on the level of human capital and not on group identity. We observe that if group B is statistically discriminated against so that $\pi^B < \pi^W$, it must be that $I(\pi^B) < I(\pi^W)$. The only way this can happen is if the "return to a high signal" is higher in group W than in group

[6] A notable exception is Foster and Rosenzweig [11],[12], who use detailed micro data from rural economies. Several observations of piece-rate and time-rate earnings are available. Using piece-rate earnings as an independent measure of productivity can allow for the identification and testing of statistical discrimination.

B, which is equivalent to:

$$w^B(\theta_1;\pi^J) - w^B(\theta_0;\pi^J) < w^W(\theta_1;\pi^J) - w^W(\theta_0;\pi^J). \tag{5}$$

If the researcher in this simple world observed that the average wage in group B is below the average wage in group W, but that (5) fails, the conclusion would be that there is something wrong with the model. Hence the hypothesis of statistical discrimination could be falsified without directly observing individual productivities.

3.4 The Consequences of Using a Measure of Skill not Observed by Firms

Suppose that the researcher does not observe θ but observes instead a different signal $\Theta \in \{\Theta_1, \Theta_0\}$ which is not observed by the firms. We assume that also the signal Θ is racially unbiased, so that the probability of obtaining Θ_1 depends only on whether the worker is qualified or not and not on group identity. Hence, we let

$$\Pr\left[\Theta = \Theta_1 | h = 0\right] = r_u < r_q = \Pr\left[\Theta = \Theta_1 | h = 1\right],$$

where it is irrelevant for the discussion below whether θ or Θ is the more precise measure of skill.

Since θ and Θ are both proxies of unobservable skills it seems reasonable to simply replace θ with Θ in (5) and check whether the returns to the *observable* measure of skill is lower for the group with less human capital. That is, we take

$$R^J(\pi^J) = E\left[w^J|\Theta_1\right] - E\left[w^J|\Theta_0\right] \tag{6}$$

as the operational measure of return to skill in group J. To assess this strategy we will assume that our simple model of statistical discrimination is true and that the equilibrium that generates the data is an equilibrium where $\pi^B \neq \pi^W$. It is then straightforward to derive what the investigators measure of observable skill should be in terms of π^J and parameters of the model, as

$$\begin{aligned}
R(\pi^J) &= E(w^J|\Theta_1) - E(w^J|\Theta_0) \\
&= \Pr(\theta_1|\Theta_1,\pi^J)w^J(\theta_1;\pi^J) + \left[1 - \Pr(\theta_1|\Theta_1,\pi^J)\right] w^J(\theta_0;\pi^J) \\
&\quad - \Pr(\theta_1|\Theta_0,\pi^J)w^J(\theta_1;\pi^J) + \left[1 - \Pr(\theta_1|\Theta_0,\pi^J)\right] w^J(\theta_0;\pi^J) \\
&= [\Pr(\theta_1|\Theta_1,\pi^J) - \Pr(\theta_1|\Theta_0,\pi^J)][w^J(\theta_1;\pi^J) - w^J(\theta_0;\pi^J)],
\end{aligned} \tag{7}$$

where, using Bayes' rule,

$$
\begin{aligned}
\Pr(\theta_1|\Theta_1,\pi^J) &= \frac{\pi^J p_q r_q + \left(1-\pi^J\right) p_u r_u}{\pi^J p_q r_q + (1-\pi^J)\, p_u r_u + \pi^J\,(1-p_q)\, r_q + (1-\pi^J)\,(1-p_u) r_u} \\
&= \frac{\pi^J p_q r_q + \left(1-\pi^J\right) p_u r_u}{\pi^J r_q + (1-\pi^J) r_u} \\
\Pr(\theta_1|\Theta_0,\pi^J) &= \frac{\pi^J p_q(1-r_q) + \left(1-\pi^J\right) p_u(1-r_u)}{\pi^J p_q(1-r_q)+\left(1-\pi^J\right)p_u(1-r_u)+\pi^J(1-p_q)(1-r_q)+(1-\pi^J)(1-p_u)(1-r_u)} \\
&= \frac{\pi^J p_q(1-r_q) + \left(1-\pi^J\right) p_u(1-r_u)}{\pi^J\,(1-r_q) + (1-\pi^J)\,(1-r_u)}.
\end{aligned} \tag{8}
$$

Comparing (4) with (7) we observe that the only difference between the expressions is that $p_q - p_u$ in $I\left(\pi^J\right)$, is replaced with $\Pr(\theta_1|\Theta_1,\pi^J) - \Pr(\theta_1|\Theta_0,\pi^J)$ in $R\left(\pi^J\right)$. Hence, the true incentives is the wage premium for a good "employer signal", multiplied with the difference in obtaining a good signal between a qualified worker and an unqualified worker. In the expression for the observed measure of returns to skill the wage premium is instead multiplied with the difference in the probability of obtaining a good employer signal between a worker with a good "investigator signal" and one with a bad "investigator signal".

Since there are both qualified and unqualified workers that draw Θ_1 it follows that $\Pr(\theta_1|\Theta_1,\pi^J) < p_q$, and symmetrically $\Pr(\theta_1|\Theta_0,\pi^J) > p_u$. Hence, the measured incentives will be a biased measure of the true incentives. This should be intuitive, and, since we are mainly concerned with the relative magnitude of the incentives rather than a dollar value, it is not a problem in itself.

The more serious issue is that $\Pr(\theta_1|\Theta_1,\pi^J) - \Pr(\theta_1|\Theta_0,\pi^J)$ changes with π^J. Hence, different compositions of investors and non-investors in the two groups lead to different magnitudes of the bias. Moreover, there is nothing in the theory that suggests that the bias should either be larger or smaller in the statistically discriminated group: it all depends on parameters of the model and we can see no way to a priori judge which case is the more likely.

The "composition effects" that lead to group differences in $\Pr(\theta_1|\Theta_1,\pi^J) - \Pr(\theta_1|\Theta_0,\pi^J)$ may be strong enough to make "observed returns to skills" the same or even higher for the group with lower returns to skill investments. As an example, Figure 1 illustrates the possibility that the returns to skill measured by the signal available to the researcher are the same for both groups in spite of the wage differences being generated by statistical discrimination. The two functions $I(\pi)$ and $R(\pi)$ are constructed using parameters $p_q = r_q = .8$, $p_u = 0.5$, and $r_u = 0.1$. We have indicated two possible proportions of investors, π^B and π^W with $\pi^B < \pi^W$. The associated incentives to

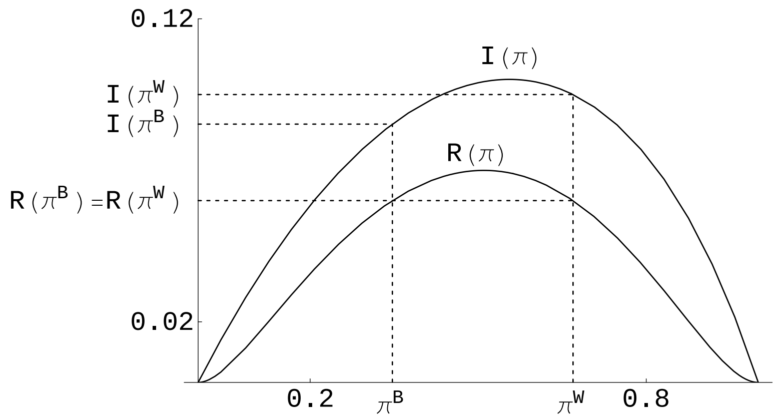

Figure 1: Incentives and observed incentives

acquire human capital $I(\pi^B)$ and $I(\pi^W)$ satisfy $I\left(\pi^B\right) < I\left(\pi^W\right)$ hence there exist (an abundance of) distributions G such that the equilibrium conditions hold for π^B and π^W. Even though "true incentives" are different between the two groups, they appear to be the same to the investigator observing a different signal. Notice that this is true even though the researcher's signal is more informative than the signal used by firms, having higher correlation with productivity for any π.

So far we have only discussed bad news. However, for any $\pi^J \in (0,1)$

$$\begin{aligned} \lim_{(r_q,r_u)\to(1,0)} \Pr(\theta_1|\Theta_1,\pi^J) &= p_q \\ \lim_{(r_q,r_u)\to(1,0)} \Pr(\theta_0|\Theta_1,\pi^J) &= p_u. \end{aligned} \tag{9}$$

There is some continuity in the model implying that if the researcher's signal is a "good enough" indicator of human capital, then differences in observed returns to skill cannot be explained away with group specific differences in the bias of the estimated returns to human capital accumulation.

4 Continuous Human Capital

In this section we will consider a version of the setup which, while still being very special, is somewhat more flexible than the discrete example in the previous section. Our objective is to generate a model of statistical discrimination where log wages are linear in the proxy for skill. This allows us to analyze the bias that is generated when the firms and the empirical investigator observe different proxies of skills in a setup that should look familiar to most readers. Returns to skill in this variant can be measured as the coefficient on the imperfect measure of skill from a Mincerian regression.

The distributional assumptions necessary for this log-linear version of the model imply that there is only one non-trivial equilibrium that we can compute. Discrimination can thus *not* be a pure equilibrium phenomenon in this version. However, we may still use this model to assess whether there are differences in incentives by allowing certain (in principle estimable) parameters to differ.

Again the two groups can be analyzed in isolation, so we will for simplicity drop the superscript J indicating group identity. We will generate log wages linear in the measure of skill by setting the model up so that the distribution of human capital is lognormal in equilibrium. Recall that variable x is lognormal with parameters μ, σ if $\ln(x)$ is normal with mean μ and standard deviation σ. We assume that the individual i has a cost function for human capital accumulation given by

$$C(h;i) = \frac{h}{i} \tag{10}$$

and that $\ln(i) \sim N(\mu_i, \sigma_i)$. We parametrize preferences by assuming that $u(w) = \ln(w)$. As in the previous example we assume that a worker with human capital h produces output worth h no matter which firm the worker is employed in or what the composition of the labor force is. For now, we simply assume that the distribution of h in the population is lognormal with parameters μ_h and σ_h, i.e. $\ln(h) \sim N(\mu_h, \sigma_h)$. Once we have derived the associated equilibrium wages given this distributional assumption we will go back and verify that this is true given the parametrization of the preferences and cost functions assumed above.

In order for wages to be log-linear in θ we need to assume that $\theta = \ln(h) + \varepsilon$, where $\varepsilon \sim N(0, \sigma_\varepsilon)$.[7] Standard properties of the Normal distribution imply that the conditional distribution of $\ln(h)$ given θ is also Normal:

$$f(\ln(h)|\theta) \sim N\left(\mu_h \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_h^2} + \theta \frac{\sigma_h^2}{\sigma_\varepsilon^2 + \sigma_h^2}, \sqrt{\frac{\sigma_h^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_h^2}}\right). \tag{11}$$

Risk neutral firms pay workers $E(h|\theta)$ in equilibrium. This expectation is easily calculated: if $\ln x \sim N(\mu, \sigma)$, then $E(x) = \exp\left(\mu + \sigma^2/2\right)$. We thus conclude that the equilibrium wage scheme

[7] Or, equivalently that $\theta = h\varepsilon$, where $\ln \varepsilon \sim N(0, \sigma_\varepsilon)$. Since θ is an ordinal variable there is no difference between this setup and the one we use.

must be

$$w(\theta) = E(h|\theta) = \exp\left(\mu_h \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_h^2} + \theta \frac{\sigma_h^2}{\sigma_\varepsilon^2 + \sigma_h^2} + \frac{1}{2} \frac{\sigma_h^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_h^2}\right).$$

To simplify notation we define constants

$$\begin{aligned} \alpha &\equiv \mu_h \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_h^2} + \frac{1}{2} \frac{\sigma_h^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_h^2} \\ \beta &\equiv \frac{\sigma_h^2}{\sigma_\varepsilon^2 + \sigma_h^2} \end{aligned} \tag{12}$$

so that $w(\theta) = \exp(\alpha + \beta\theta)$. The individual decision problem is to solve,

$$\max_{h \in H} E_\theta \left[u\left(w(\theta)\right)|h\right] - C(h; i) = \max_{h \geq 0} E_\theta \left[\ln\left(w(\theta)\right)|h\right] - \frac{h}{i}. \tag{13}$$

But given that the worker chooses human capital h she knows that her signal is normal with mean $\ln h$, i.e. $\theta = \ln(h) + \varepsilon$, where $\varepsilon \sim N(0, \sigma_\varepsilon)$, so

$$E_\theta \left[\ln\left(w(\theta)\right)|h\right] = E_\theta \left[\ln\left(\exp\left(\alpha + \beta\theta\right)\right)|h\right] = E_\theta \left[\alpha + \beta\theta|h\right] = \alpha + \beta E_\theta\left(\theta|h\right) = \alpha + \beta \ln h \tag{14}$$

Hence, the maximization problem reduces to

$$\max_{h \geq 0} \alpha + \beta \ln h - \frac{h}{i}. \tag{15}$$

This is a strictly concave maximization problem and since $\lim_{h \to 0} \beta/h - 1/i \to \infty$ for every $i > 0$ the solution must be interior. The first order condition to the maximization problem is $\beta/h - 1/i = 0$ and from this we conclude that $h(i) = \beta i$, implying that human capital is distributed log-normally as we conjectured when we derived the wage schemes. More precisely,

$$\ln(h) \sim N\left(\mu_i + \ln\beta, \sigma_i\right). \tag{16}$$

The wage scheme the worker expects in order for (16) to be the equilibrium human capital distribution was derived from the assumption that $\ln(h) \sim N(\mu_h, \sigma_h)$. Hence, for the distribution of h to be a rational expectations equilibrium it must be that

$$\begin{aligned} \sigma_h &= \sigma_i \\ \mu_h &= \mu_i + \ln\beta = \mu_i + \ln\beta = \mu_i + \ln\left(\frac{\sigma_h^2}{\sigma_\varepsilon^2 + \sigma_h^2}\right) \\ &= \mu_i + \ln\left(\frac{\sigma_i^2}{\sigma_\varepsilon^2 + \sigma_i^2}\right). \end{aligned} \tag{17}$$

Since μ_i, σ_i and σ_ε all are exogenous parameters this calculation shows that there is only one equilibrium in the model where the distribution of human capital follows a non-degenerate lognormal distribution. There is also a trivial equilibrium where no worker invests at all, and there may be equilibria where the distribution is non-degenerate, but not lognormal.

The skill distribution in the lognormal equilibrium has the following mean and variance:

$$\begin{aligned} E(h) &= \exp\left(\mu_i + \ln\left(\frac{\sigma_i^2}{\sigma_\varepsilon^2+\sigma_i^2}\right) + \frac{\sigma_i^2}{2}\right) \\ Var(h) &= \exp\left(2\mu_i + 2\ln\left(\frac{\sigma_i^2}{\sigma_\varepsilon^2+\sigma_i^2}\right) + \sigma_i^2\right)\exp\left(\sigma_i^2 - 1\right) \end{aligned}$$

This shows that the average human capital investment is inversely related to the variance of the signal noise but increases with the variance of the cost distribution.

To summarize, we found that there are potentially multiple equilibria, but there is only one interesting equilibrium that we are able to compute. Hence we have to resort to exogenous differences between groups in order to rationalize any differences in behavior. In particular, if we make the more or less standard assumption that the signal is noisier for minorities than it is for whites, then that will create lower investments in human capital in the minority.[8]

4.1 The Consequences of Relying on Tests that are not Observed by Firms

Assume that the investigator instead observes a variable Θ, correlated with h satisfying $\Theta = \ln(h) + \delta$, where $\delta \sim N(0, \sigma_\delta)$ is independent from the noise in the firm signal ε. Moreover, imagine that the investigator tries to determine whether differences in incentives to invest can explain the differences in average wages between group B and group W by regressing log wages on a constant and Θ. Dropping group superscript J, and assuming that the investigator observes the sample $[\Theta_i, w_i]_{i=1}^N$ we can define the following least squares estimator:

$$\widehat{b} = \frac{Cov_N(\Theta_i, \ln(w_i))}{Var_N(\Theta_i)}$$

where Var_N and Cov_N denote the sample variance and covariance.

We now measure the extent in which $\widehat{b}$ is a biased measure of the returns to skill β. We have assumed $\Theta_i = \ln h_i + \delta_i$ and from the model we derived $\ln(w_i) = \alpha + \beta\theta_i = \alpha + \beta(\ln h_i + \varepsilon_i)$. Variables

[8] The most common way to generate statistical discrimination in the literature is to assume that signals of the performance of the discriminated group are less accurate than for the dominant group. See for example Phelps [20], Aigner and Cain [1], Lundberg and Startz [13], and Cornell and Welch [10].

$\ln h_i$, ε_i, δ_i are drawn from i.i.d. normal distributions with mean and variance respectively equal to $\left(\mu_i + \ln\left(\frac{\sigma_i^2}{\sigma_\varepsilon^2+\sigma_i^2}\right), \sigma_i^2\right)$, $\left(0, \sigma_\varepsilon^2\right)$, $\left(0, \sigma_\delta^2\right)$. Hence,

$$\begin{aligned} Cov_N(\Theta_i, \ln(w_i)) &= Cov_N\left(\ln h_i + \delta_i, \alpha + \beta(\ln h_i + \varepsilon_i)\right) \\ &= \beta Cov_N(\ln h_i, \ln h_i) + \beta Cov_N(\ln h_i, \varepsilon_i) \\ &\quad + \beta Cov_N(\delta_i, \ln h_i) + \beta Cov_N(\delta_i, \varepsilon_i) \end{aligned} \tag{18}$$

But the realizations of $(\ln h, \varepsilon)$ are i.i.d. and so are (δ, ε) and $(\delta, \ln h)$. Hence we can apply the law of large number and observe that the first term in (18) converges to $\beta\sigma_i^2$, while the others converge to 0. Similarly, $Var_N(\Theta_i)$ converges to the population variance of Θ which is equal to $\sigma_\delta^2 + \sigma_i^2$.

We can therefore derive the following asymptotic result:

$$\mathrm{p}\lim(\widehat{b}) = \beta\left(\frac{\sigma_i^2}{\sigma_\delta^2 + \sigma_i^2}\right) \tag{19}$$

so that the asymptotic bias is $-\beta\sigma_\delta^2/\left(\sigma_\delta^2 + \sigma_i^2\right)$.

The only difference between our setup and the canonical setup with measurement error is that there are two independent error terms generating the difference between θ and Θ rather than a single error term. However, for large samples it is only the noise in the signal Θ that matters directly (indirectly our theory implies that the other error term affects the distribution of human capital). In any other way the intuition from the canonical model with measurement error goes through: estimates are downwards biased and the larger is the measurement error the larger is the downward bias.

Since the group with higher human capital has higher β, it is possible that the estimated difference in returns to skill is smaller than the true difference, in which case the investigator may erroneously believe there is no statistical discrimination. This however depends on the details on how statistical discrimination is generated. The next section investigates this issue and discusses how we may be able to reconcile NLSY data with the presence of statistical discrimination.

5 Discussion: Estimates of Returns to Skill Using NLSY Data

In their study of black-white wage differences, Neal and Johnson [17] found that skill differences accounts for a quite large share of the wage gap. This fact alone is obviously consistent with wage differentials generated by statistical discrimination, but they argue explicitly against such an interpretation. Instead they argue that the differences are created during childhood and are driven by differences in family backgrounds and other obstacles that make it hard for black children to acquire human capital. Their reason for dismissing the hypothesis that statistical discrimination has anything to do with the differences comes from the fact that the OLS estimates of return to AFQT are roughly the same for blacks and whites.

We know from the discussion above that OLS estimates of the AFQT coefficient are biased estimates of the return to acquire skills *if data is generated from the model discussed in this article.* Also notice from (19) that the magnitude of the bias depends on the variance of the skill distribution, which in turn (according to the theory) is related to the observable variance of wages. In addition there are restrictions from the theory, which together with the observation that we do observe the wage variance means that it is quite possible that the finding that the OLS estimates are the same in both groups cannot be consistent with the model.

In (20) below we summarize the summary statistics from the NLSY sample (restricted to male workers) that relates directly to the theory.

	Black	White
Observations	466	825
$\overline{w}^j \equiv$Average($\ln(w)$)	6.64	6.89
$\sigma_w^J \equiv$Stdev($\ln(w)$)	.46	.43
$\sigma_\Theta^J \equiv$Stdev($\ln(\Theta)$)	.82	.93
$\widehat{b}^J$	.19 (.02)	.18 (.02)

(20)

Notice that we interpret the AFQT test score as the variable Θ in our model and that the estimated coefficient resulting from regressing log wages on AFQT scores (standard errors in parenthesis) is not significantly different between groups.[9]

We can ask whether it is possible to reconcile the facts in (20) with the model and whether there are any reasons to believe that blacks returns to skill are lower than whites', so that incentives determine to the lower skill level among blacks.

The model implies (using (12) and $\sigma_i = \sigma_h$) that wage means and the variances should be

[9] Neal and Johnson [17] also find that the coefficient is not significantly different when controlling for age and the square of AFQT.

related to the fundamental parameters of the model as follows (omitting group superscripts):

$$\begin{aligned}
\overline{w} &= E(\log(w)) = \alpha + \beta E(\theta) \\
&= \underbrace{\mu_h \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_i^2} + \frac{1}{2}\frac{\sigma_i^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_i^2}}_{\alpha} + \underbrace{\left(\frac{\sigma_i^2}{\sigma_\varepsilon^2 + \sigma_i^2}\right)}_{\beta}\mu_h \\
&= \mu_h + \frac{1}{2}\frac{\sigma_i^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_i^2} = \mu_i + \ln\left(\frac{\sigma_i^2}{\sigma_\varepsilon^2 + \sigma_i^2}\right) + \frac{1}{2}\frac{\sigma_i^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_i^2} \qquad (21) \\
\sigma_w^2 &= Var(\log(w)) = \beta^2 Var(\theta) = \frac{\left(\sigma_i^2\right)^2}{\sigma_\varepsilon^2 + \sigma_i^2} \qquad (22) \\
\sigma_\Theta^2 &= Var(\Theta) = Var(\log(h)) + \sigma_\delta^2 = \sigma_i^2 + \sigma_\delta^2 \qquad (23) \\
p\lim \widehat{b}^J &= \frac{\sigma_i^2}{\sigma_\varepsilon^2 + \sigma_i^2}\frac{\sigma_i^2}{\sigma_\delta^2 + \sigma_i^2} \qquad (24)
\end{aligned}$$

Because the scale of the AFQT variable is arbitrary, we have chosen to initially ignore σ_Θ^2. Assume that AFQT precision doesn't vary between races (i.e. $\sigma_\delta^B = \sigma_\delta^W = \sigma_\delta$), and ignore small sample issues and treat the estimate $\widehat{b}^J$ as the asymptotic limit of the least square coefficient. Then, (21), (22), and (24) provide six equations in seven unknowns: $\sigma_i^J, \sigma_\varepsilon^J, \mu_h^J$, $J = B, W$ and σ_δ. There is one free parameter, so there may be several solutions. For example, two sets of parameters that satisfy all six equations are:

Example 1

	Black	White
σ_δ^2	1.90	
σ_i^{2J}	.52	.49
σ_ε^{2J}	.069	.067
μ_h^J	6.73	6.99

Example 2

	Black	White
σ_δ^2	1.94	
σ_i^{2J}	.481	.449
σ_ε^{2J}	.022	.020
μ_h^J	6.67	6.92

We computed in each example the implied value of the returns to investment $\beta^J = \sigma_i^{J2} / \left(\sigma_i^{J2} + \sigma_\varepsilon^{J2}\right)$. Parameters in Example 1 imply $\beta^B = 0.882 > \beta^W = 0.879$, while those in Example 2 imply $\beta^B = 0.956 < \beta^W = 0.958$. Hence, depending on the assumed value of σ_δ whites may have higher or lower returns to investment than blacks. However, we cannot rationalize *large positive* differences in returns to investment between whites and blacks. We also observe that the parameters computed in each example imply lower AFQT variance in whites, which is counterfactual.

If one is willing to relax the assumption that AFQT precision is the same and allow $\sigma_\delta^B \neq \sigma_\delta^W$, then it is possible to reconcile *all* of the eight restrictions from (21)-(24) together with large, positive differences in returns to investment between whites and blacks. However, to achieve this we need $\sigma_\delta^B < \sigma_\delta^W$, which some may find counterintuitive. This perhaps suggests that the data from the NLSY sample is hard to reconcile with black males being significantly statistically discriminated.

We believe that a fruitful approach to investigate whether statistical discrimination is a plausible factor for the explanation of the wage gap is to force the model to match moments of the AFQT distribution as well. The way to do this, however, is not obvious because the scale of the AFQT variable is arbitrary and it is doubtful whether results that rely on the cardinal properties of this variable should be taken seriously. We have chosen to investigate this issue in future research.

6 Conclusion

The basic message with this article is that if the theory of statistical discrimination is to be taken seriously, then the empirical work should reflect the most fundamental assumption in the theory: true productivity is observed only with error. We have argued that the most crucial implication of the theory is that the returns to accumulate imperfectly observable skills should be lower for groups with lower average wages.

If the researcher can observe the same variables that firms use in choosing their wage offers, these returns can be quantified by measuring the returns to the observable proxies of productivity used by employers. However, using measures of the returns skills that are not observed by the firms generates biased estimates of the returns to skill due to measurement error. The nature of the bias is such that the researcher may conclude that the returns to skill are the same for both group even if the data is generated by an equilibrium with statistical discrimination.

The reader should notice that we are not saying that estimates of return to ability using AFQT or other measures of ability are necessarily contaminated by measurement error. In the context of the very model we are using, the procedure is fine if employers could observe the AFQT test scores in the dataset. If human capital can be perfectly observed by firms or if the informational asymmetry can be "contracted away" the procedure would also be legitimate, but then one could not interpret this as evidence either against or in favor of statistical discrimination since it would be impossible to support statistical discrimination in such a model.

References

[1] Aigner, Dennis J., and Glen Cain, "Statistical Theories of Discrimination in the Labor Market," *Industrial and Labor Relations Review* **30,** 1977, 175-187.

[2] Altonji, Joseph, and Rebecca Blank: "Race and gender in the labor market" , in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics, vol. 3c* Elsevier Science B.V. (1999), 3144-3259.

[3] Altonji, Joseph, and Charles Pierret, "Employer Learning and the Signaling Value of Education" ,*Forthcoming in* I. Ohashi and T. Tachibanaki Ed*., Employment Adjustment, Incentives and Internal Labour Market*, McMillan Publishing Co.

[4] Antonovics, K., "Persistent Racial Wage Inequality", University of California, San Diego Discussion Paper n. 2002-05, February 2002.

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

[6] Bowlus, Audra J. and Zvi Eckstein, "Discrimination and Skill Differences in an Equilibrium Search Model", *International Economic Review,* 2002, 1309-46.

[7] Becker, Gary S., *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education,* Third Edition, University of Chicago Press, Chicago, 1993.

[8] Cawley, John, James Heckman, and Edward Vytlacil, "Three Observations on Wages and Measured Cognitive Ability," *Labour Economics* **8,** 2001, 419-442.

[9] Coate, Stephen and Glenn C. Loury, "Will Affirmative Action Policies Eliminate Negative Stereotypes?," *American Economic Review* **83**, December 1993, 1220-40.

[10] Cornell, Brad, and Ivo Welch. "Culture, Information and Screening Discrimination." *Journal of Political Economy* 104-3, June 1996, 542-571.

[11] Foster, Andrew, and Mark Rosenzweig, "Information, Learning, and Wage Rates in Rural Labor Markets", *Journal of Human Resources*, **28**(4), 1993, 759-790.

[12] Foster, Andrew, and Mark Rosenzweig, "Comparative Advantage, Information and the Allocation of Workers to Tasks: Evidence from an Agricultural Labor Market", *Review of Economic Studies* **63**(3), July 1996, 347-374.

[13] Lundberg, Shelly J. and Richard Startz, "Private Discrimination and Social Intervention in Competitive Labor Markets", *American Economic Review* **73** (1983), 340-47.

[14] Moro Andrea, " The Effect of Statistical Discrimination on Black-White Wage Differentials: Estimating a Model with Multiple Equilibria", *forthcoming, International Economic Review,* 2003.

[15] Moro, Andrea and Peter Norman, "Affirmative Action in a Competitive Economy", *forthcoming, Journal of Public Economics*, 2003.

[16] Moro, Andrea and Peter Norman, "A General Equilibrium Model of Statistical Discrimination", *forthcoming, Journal of Economic Theory,* 2003.

[17] Neal, Derek A. and William R. Johnson (1996), "The Role of Premarket Factors in Black-White Wage Differentials", *Journal of Political Economy* **104**, 869-895.

[18] Neumark, David, "Labor Market Information and Wage Differentials by Race and Sex, " *NBER Working Paper* No. *6573,* May 1998.

[19] Oettinger, Gerald, "Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap", *Journal of Labor Economics* **14**, January 1996, 52-78.

[20] Phelps, Edmund S., "The Statistical theory of Racism and Sexism," *American Economic Review* **62** (September 1972), 659-61.

[21] Persico, Nicola, Andrew Postlewaite and Dan Silverman, "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height", mimeo, University of Pennsylvania, 2002.