Education and Labor-Market Discrimination

Kevin Lang **Michael Manove**

Boston University

February, 2008

Short Abstract: We combine statistical discrimination with educational sorting. Our model predicts that blacks get more education than whites of similar ability and that blacks have lower wages than similarly educated whites, both confirmed in the data. Our model suggests, that when measuring the black-white earnings differential, one should control for both AFQT and education. In that case, a substantial wage differential emerges. We explore and reject the hypothesis that school quality differences explain the wage and education differentials, suggesting that some of the wage differential reflects the operation of the labor market.

Abstract: We propose a model combining statistical discrimination and educational sorting that explains why blacks get more education than do whites of similar cognitive ability. Our model explores the relation between education and AFQT and that between wages and education, and it explains why these relations differ for blacks and whites. Although the model cannot explain why, conditional only on AFQT, blacks earn no more than do whites, it does suggest, that when comparing the earnings of blacks and whites, one should control for both AFQT and education. In that case, a substantial black-white wage differential emerges. We explore and reject the hypothesis that differences in school quality between blacks and whites explain the wage and education differentials. Our findings support the view that some of the black-white wage differential reflects the operation of the labor market.

JEL classification: .

Authors' Addresses:

Kevin Lang
lang@bu.edu

Michael Manove
manove@bu.edu

Dept. of Economics
Boston University
270 Bay State Road
Boston MA 02215

Acknowledgment: This research was supported in part by NSF grant SEC-0339149. We thank Derek Neal and Eli Berman for extremely helpful discussions and for helping us to replicate his results. We also thank participants in seminars and workshops at Boston College, Boston University, Carnegie-Mellon University, Carlos III, Duke, George Washington University, the LSE, Michigan, the NBER, NYU, Pompeu Fabra, Rochester, Syracuse, the University of Houston/Rice, the University of Western Ontario, UCSD and Yale for helpful comments and criticisms. The usual caveat applies.

1 Introduction

Models of statistical discrimination often imply that African Americans—or others suffering from discrimination—invest less in themselves than otherwise comparable whites do, because African Americans receive a lower return from investment in human capital than whites (see, for example, Lundberg and Startz, 1983 and Coate and Loury, 1993). The critical assumption underlying these models is that the market receives imperfect information about productivity and no direct information about investments in human capital. Although many human-capital investments are undoubtedly unobservable, others, such as educational attainment, are readily observed. Observable investments can signal productivity as in Spence (1973), and the value of the signal will be greater the less reliable is the employers' information about worker productivity. Moreover, evidence suggests that employers find it more difficult to evaluate black job candidates than to evaluate white candidates. Therefore, when observable investments in human capital are available, it is plausible that statistical discrimination will induce blacks to invest in themselves more than whites do, not less.

To formalize this intuition, we construct a model of statistical discrimination in which a worker's race and educational attainment are observable by prospective employers, but a worker's ability is not observable except by the worker himself. We analyze the model from a theoretical perspective, and then we apply it to the data.

In our empirical analysis, we proxy ability by the performance of high school students on the Armed Forces Qualification Test (AFQT), often conditioning on educational as of the time of testing, and we measure educational attainment by the highest grade completed. We find that educational attainment, conditional on AFQT, is higher among blacks than it is among whites, a result that is consistent with our signaling hypothesis. We also predict that whites and blacks will have similar earnings at very low and very high levels of education (not controlling for ability) but that blacks will have lower earnings at intermediate levels of education. We do not test this for women because of selection issues (Neal, 2004), but we do confirm the hypothesis for men. Our particular formalization implies that blacks would have levels of education similar to those of whites at low and high levels of ability and levels exceeding those of whites at intermediate levels of ability. This is confirmed only for men (black women have more education than whites at all but the lowest ability levels).

Of course, there are plausible alternative explanations for these findings. In particular, if African Americans attend lower quality schools than whites do, and if AFQT is mostly determined by school inputs, then blacks get less of an AFQT benefit than whites from schooling, so that for a given amount of schooling, blacks will have lower AFQT scores. This means that for a given AFQT score, blacks will have more education than whites. We test the alternative hypothesis directly by controlling for measurable differences in school quality, and we find that school quality cannot explain the education differential. In addition, we show that our results are robust to controlling for a number of other factors that influence educational attainment.

For the purposes of our analysis, AFQT must be exogenous to eventual educational attainment, but it need not be a measure of purely innate ability. In fact, we view AFQT as a measure of both innate ability and ability created by acquired personal traits. We do not attempt to explain the behavior of children and adolescents prior to the administration of the AFQT, nor do we assume such behavior takes account of the value of investment in education or human capital. But we do assume that students act as rational agents after the administration of the AFQT, who, aside from wanting to invest in their human capital, are motivated by a desire to signal.

Our model and findings have important implications for the debate over the role of market and premarket factors in explaining black-white wage differentials. In a highly influential article, Derek Neal and William Johnson (1996; see also O'Neill, 1990) argue that wage differentials between blacks and whites can be explained by productivity-related personal characteristics (premarket factors) rather than by labor-market discrimination. To this end, Neal and Johnson (hereafter NJ) show that the black-white wage differential is dramatically reduced, and in some cases eliminated, by controlling for performance on the AFQT. Therefore, if discrimination were responsible for black-white wage differentials, it must be true that discrimination (or the anticipation of future discrimination) adversely affects black performance on the AFQT. Neal and Johnson show, however, that the effect of AFQT on the earnings of blacks is at least as large as on the earnings of whites,[1] so that labor-market discrimination should not have caused blacks to reduce their investments in the cognitive skills that the AFQT measures. Thus they conclude that labor-market discrimination does not account for black-white earnings differentials.

In what follows, we attempt to reopen this question. We argue that when examining black and white wage differentials, it is not appropriate to control only for AFQT. Our argument is based on the fact that with AFQT performance given, blacks get more education than whites do. In the absence of discrimination, blacks should be rewarded for their greater education. The similar earnings of blacks and whites when we control only for AFQT suggest that blacks are not so rewarded. We show that when we control for both education and AFQT, wage differentials between blacks and whites are substantially larger than when we control for AFQT alone.

In the general case, our model suggests that if we control for AFQT but not for education, blacks should earn more than whites do at all but the highest and lowest levels of ability.[2] We test this hypothesis for men and reject it. The failure of the model in this regard could reflect either missing variables or labor market discrimination. This is an old debate and not one we will pretend to resolve. We do explore whether the wage differential can be explained by differences in the

---

[1] Note, however, that if AFQT is influenced by investments in human capital, then we should expect the observed returns to AFQT to be equated to the agent's discount rate in equilibrium. If black and white workers have similar discount rates, they would also have similar equilibrium rates of return to AFQT. But it is perfectly possible that at a given AFQT score, the returns would differ between blacks and whites.

[2] The only exception occurs when education at the margin functions as a pure signal with no effect on productivity, a special case described in Section 2.1.3 below.

quality of schools attended by blacks and whites and find no evidence to support this hypothesis.

2 Why Blacks Get More Education than Whites: A Signaling Model

In this section, we argue that statistical discrimination against blacks creates incentives for them to signal ability through education. We believe that ethnographic evidence supports the view that blacks see education as a means of getting ahead. Newman (1999) finds that blacks in low-skill jobs in Harlem view education as crucial to getting a good job and that blacks with low levels of education have difficulty obtaining even jobs that we would not normally think of as requiring a high school diploma. Kirschenman and Neckerman (1991) also find that employers are particularly circumspect in their assessment of low-skill blacks, a finding consistent with our approach.

Our theoretical model merges a standard model of statistical discrimination (e.g., Aigner and Cain, 1977) with a conventional sorting model. In a sense, it stands Lundberg and Startz on its head, by dealing with observable investment in contrast with the unobservable investment in that paper. As is standard in the statistical discrimination literature, we assume that the productivity of blacks is less easily observed than the productivity of whites. However, consistent with our reading of the ethnographic literature, we add a nonstandard feature to our model: as education levels increase, the ability of firms to assess the productivity of both black and white workers improves, until, for sufficiently high levels of education, firms assess productivity for the two races equally well. This assumption is not required for our principal result, that blacks generally get more education than equivalent whites do, but it is needed for some of the other predictions of the model. In addition, our model specifies that firms observe the productivity of the highly educated perfectly, although our results require only that there be no asymmetry of information between highly-educated workers and firms.

Our strategy for analyzing statistical discrimination is to develop a game-theoretic signaling model of educational attainment and apply it separately to blacks and to whites, setting parameters for each group consistent with the above suppositions. Then we compare the equilibrium outcomes of the two groups. The principal result is that because employers have greater difficulty directly observing the productivity of blacks than of whites, they put more weight on education (an observable trait related to productivity) when making offers to blacks than they do when making offers to whites. In response, blacks choose to get more education than do whites of similar ability. However, because observability changes, this effect disappears at the highest levels of education.

2.1 The Signaling Game

In this section, we construct and analyze a signaling game between a racially homogeneous group of job candidates. In the next section we will compare the outcome of the game between whites with the outcome of the game between blacks.

Workers of a given race differ in ability and educational attainment, both of which have a positive effect on productivity. Ability can be observed only by the worker himself, but educational attainment can be observed precisely by potential employers. In addition, potential employers can observe a candidate's productivity directly but with error. The worker then has an incentive to use education as a signal of productivity. The greater the error in the employer's direct observation, the more weight the employer will place on the education signal.

In the standard job-market signaling model, worker signaling is costly, and the cost of the signal falls as ability (and productivity increases). Therefore, there is a separating equilibrium in which high-ability agents buy more of the signal than low-ability agents. In our model, however, the opportunity cost of education is higher for more able workers. Therefore, if education fully revealed productivity, the candidate separating equilibrium would unravel. To surmount this problem, we make the very realistic assumption that workers cannot perfectly predict their own productivity from knowledge of their ability and education alone. In our model, productivity depends not only on characteristics known to the worker, but on a random element that the worker does not observe. Consequently, the employer would always have an incentive to place a positive weight on his direct observation of the worker's productivity. A low ability person who deviates from a separating equilibrium and chooses more schooling in order to pool with a high ability person, would nevertheless have an expected return smaller than that of high ability person. This makes the existence of a separating equilibrium possible.

Consider now a game between a continuum of workers of different ability levels a, where a is continuously distributed over some fixed interval. Each worker must choose a level of education s. Because we assume that education and ability are complementary inputs in the creation of productivity (in a sense defined below), we shall search for a separating equilibrium in which the workers' strategy profile is described by a continuous and differentiable function $S(a)$, strictly increasing in a, where $s = S(a)$ is the education obtained by a worker of ability a. Firms in our model simply follow the rules of a competitive labor market—they play no strategic role in the game. (But employers beliefs about $S(a)$ in equilibrium are required to be correct.)

Suppose that a worker's productivity p^*, conditional on his education level s and ability a, has the log-normal distribution given by

$$p^* = Q(s,a)\,\hat{\varepsilon}, \tag{1}$$

where $Q(s,a)$ is a deterministic function of education and ability and where $\varepsilon \equiv \ln \hat{\varepsilon}$ is a normal random variable with mean 0 and variance σ_ε^2. Letting $q(s,a) \equiv \ln Q(s,a)$ denote the mean of $\ln p^*$, we can write log productivity as

$$\ln p^* = q(s,a) + \varepsilon. \tag{2}$$

We assume that the effect of education on log-productivity is characterized by diminishing returns ($q_{ss} < 0$) but that ability complements the productivity-increasing effects of education ($q_{sa} \geq 0$).

A potential employer can observe a worker's education level s but not his true productivity p^*. However, the employer does observe a productivity signal p given by

$$\ln p = \ln p^* + u, \tag{3}$$

where u is a random error of observation. The error term u has variance $\sigma_u^2(s)$, which is common to all firms, continuous and decreasing in s. We assume that ε and u are independently distributed.

Let $\lambda(s) \in [0,1]$ be given by

$$\lambda(s) \equiv \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_u^2(s)}.$$

If $\lambda(s)$ is near 0, then $\sigma_u^2(s)$ must be large, in which case the employer's ability to observe worker productivity directly is poor. Conversely, if $\lambda(s) = 1$, then $\sigma_u^2(s) = 0$, and the employer can observe worker productivity perfectly. In the latter case, workers would have no incentive to signal their productivity to employers, and they would obtain the efficient level of education.

2.1.1 The Equilibrium Competitive Wage

In the candidate separating equilibrium described by the workers' strategy profile S, an employer can infer a worker's ability a from his knowledge of the worker's education s. If $\hat{q}(s)$ denotes the employer's equilibrium inference about the value of $q(s,a)$ conditional on s, it follows that $\hat{q}(s) \equiv q(s, A(s))$, where $A \equiv S^{-1}$.

Proposition 1 *From the point of view of an employer who has observed a worker's productivity signal p and education level s, the conditional mean and variance of the unobservable random element ε is given by*

$$E[\varepsilon \,|\, p, s] = \lambda(s)(\ln p - \hat{q}(s)) \tag{4}$$

and

$$\sigma^2[\varepsilon \,|\, p, s] = (1 - \lambda(s))\, \sigma_\varepsilon^2. \tag{5}$$

Proof. Because the values of $\ln p - \hat{q}(s)$ and s uniquely determine p, we know that any expectation conditioned on p and s will remain unchanged if conditioned on $\ln p - \hat{q}(s)$ and s instead. Therefore we can write

$$E[\varepsilon \,|\, p, s] \equiv E[\varepsilon \,|\, \ln p - \hat{q}(s)\,, s]\,. \tag{6}$$

Moreover, (2) and (3) imply that

$$\ln p - \hat{q}(s) = u + \varepsilon \tag{7}$$

in equilibrium. The proposition now follows from (6) and from standard results for the sum of independent normal random variables. ∎

In a competitive labor market, an employer will offer the wage $\hat{w}(p,s) \equiv E[p^* \,|\, p, s]$ to a worker with observed characteristics p and s. We show:

Proposition 2 *The log of the equilibrium competitive wage is given by*

$$\ln \hat{w}(p,s) = \lambda(s) \ln p + (1 - \lambda(s))\left(\hat{q}(s) + .5\sigma_e^2\right). \tag{8}$$

Proof. We calculate the expected values of the terms of equation (2) conditional on the observed p and s. This yields

$$E[\ln p^* \,|\, p, s] = \hat{q}(s) + E[\varepsilon \,|\, p, s]\,. \tag{9}$$

Applying Proposition 1 give us

$$E[\ln p^* \mid p, s] = \lambda(s) \ln p + (1 - \lambda(s))\, \hat{q}(s)$$

and

$$\sigma^2[\ln p^* \mid p, s] = (1 - \lambda(s))\, \sigma_\varepsilon^2.$$

A lognormally-distributed random variable x satisfies

$$\ln E[x] = E[\ln x] + \frac{1}{2}\sigma^2[\ln x],$$

which, applied to $\hat{w}(p, s) \equiv E[p^* \mid p, s]$, yields the proposition. ■

2.1.2 Workers' Equilibrium Strategies

Each worker knows his own ability a. But a worker must choose his level of education s before ε and u are realized. When other workers have the strategy profile $S(a)$, a designated worker's expectation of his wage, conditional on his own s and a, is given by $E_{\varepsilon,u}[\hat{w}(p, s)]$, where $E_{\varepsilon,u}$ integrates over ε and u.

As a first step in deriving the best response of a worker with characteristics (s, a) to the profile $S(a)$, we compute the value of $\ln E_{\varepsilon,u}[\hat{w}(p, s)]$. From (8), (2) and (3), we see that

$$\ln \hat{w}(p, s) = \lambda(s)(q(s, a) + u + \varepsilon) + (1 - \lambda(s))\big(\hat{q}(s) + .5\sigma_e^2\big),$$

which is a normally distributed random variable with mean

$$E_{\varepsilon,u}[\ln \hat{w}(p, s)] = \lambda(s)\, q(s, a) + (1 - \lambda(s))\big(\hat{q}(s) + .5\sigma_e^2\big)$$

and variance

$$\sigma^2[\ln \hat{w}(p, s)] = \lambda(s)^2\big(\sigma_e^2 + \sigma_u^2(s)\big) = \lambda(s)\, \sigma_e^2.$$

Again, from the standard properties of log-normal random variables, we have

$$\ln E_{\varepsilon,u}[\hat{w}(p, s)] = E_{\varepsilon,u}[\ln \hat{w}(p, s)] + \frac{1}{2}\sigma^2[\ln \hat{w}(p, s)],$$

so that

$$\ln E_{\varepsilon,u}[\hat{w}(p, s)] = \lambda(s)\, q(s, a) + (1 - \lambda(s))\, \hat{q}(s) + .5\sigma_e^2. \tag{10}$$

This confirms the intuition that a designated worker's expected wage depends both on his actual ability and the ability level inferred by the employer, which in turn depends on $S(a)$.

Workers maximize expected discounted net income. Assume that the only cost of education is its opportunity cost in terms of lost income while in school. If r is the worker's discount rate, the expected present value at time $t = 0$ of the future income of a worker with characteristics (s, a) is given by

$$v(s, a) \equiv \int_s^\infty e^{-rt} E_{\varepsilon,u}[\hat{w}(p, s)]\, dt \equiv \frac{1}{r} e^{-rs} E_{\varepsilon,u}[\hat{w}(p, s)]$$

or

$$\ln v(s,a) \equiv -r - rs + \ln E_{\varepsilon,u}[\hat{w}(p,s)] \,.$$

The first-order condition for maximizing $v(s,a)$ with respect to s is

$$\frac{\partial}{\partial s} \ln E_{\varepsilon,u}[\hat{w}(p,s)] = r. \tag{11}$$

This restates the well-known proposition that when the only cost of schooling is the student's opportunity cost, the worker will continue to obtain information so long as rate of return to additional education exceeds the discount rate r. We restrict the class of equilibria we consider to those for which (11) has a unique solution.

We are now in a position to describe a separating equilibrium of the wage/education game among the class of strategy profiles that are "well behaved" (continuous, differentiable, monotonically increasing and specify a unique best response for every worker type). In all of the material below, we use the term "equilibrium" to refer to a perfect-Bayesian equilibrium of the signaling game.

Proposition 3 *If the support of worker abilities is the interval* $[a_0,\ a_1]$*, then any well-behaved separating equilibrium* S *has the property that the education level* $S(a_0)$ *of the lowest-type worker must be efficient and not influenced by signaling.*

Proof. In an equilibrium with $S(a)$ strictly increasing in a, the employer would infer that a worker with education $S(a_0)$ has ability a_0, the lowest level in the support. If $S(a_0)$were inefficiently high, the worker of ability a_0 could safely deviate to the lower efficient level of education without lowering the employer's inference of his ability, and so raise his payoff. If $S(a_0)$ were inefficiently low, the worker of ability a_0 would deviate to $s > S(a_0)$ even without consideration of the positive payoff from signaling. ■

We can now provide a complete description of any well-behaved equilibrium.

Proposition 4 *Suppose* $[a_0,\ a_1]$ *is the support of worker abilities. If a workers' equilibrium strategy profile* $S(a)$ *is well behaved, then its inverse* $A(s)$ *must satisfy the differential equation*

$$q_s + (1-\lambda)\, q_a A' = r. \tag{12}$$

For $0 \le \lambda < 1$ *, this equation is equivalent to*

$$S' = \frac{(1-\lambda)\, q_a}{r - q_s}. \tag{13}$$

For $\lambda = 1$*, the equilibrium condition is given by the solution for* s *of the equation*

$$q_s(s,a) = r. \tag{14}$$

This solution of (14) defines the efficient level of education, which we denote by $S^*(a)$*. Furthermore, we have* $S(a_0) = S^*(a_0)$ *for any function* $\lambda(s)$*. Therefore each* $\lambda(s)$ *corresponds to exactly one well-behaved equilibrium.*

Proof. Substituting (10) into (11) yields the differential equation

$$\frac{\partial}{\partial s}\left(\lambda(s)\, q(s,a) + (1-\lambda(s))\, \hat{q}(s) + .5\sigma_e^2\right) = r,$$

or

$$\lambda'(s)\, q(s,a) + \lambda(s)\, q_s(s,a) - \lambda'(s)\, \hat{q}(s) + (1-\lambda(s))\big(q_s(s,A(s)) + q_a(s,A(s))\, A'(s)\big) = r. \tag{15}$$

This equation implicitly defines the best response s of a worker with ability a to the strategy profile S. Consequently, $a = A(s)$ and $q(s,a) = \hat{q}(s)$ in equilibrium, and (15) reduces to (12). Equation (13) follows from the fact that the derivative of S is the reciprocal of the derivative of A. Proposition 3 implies that $S(a_0) = S^*(a_0)$. ■

The left-hand side of equation (12) represents the worker's rate of return to a marginal unit of education, which, given appropriate concavity conditions and the strategy-profile-inverse A, must be equated to r. This rate of return arises from the direct and indirect effects of education.

First, consider the direct effect of additional education on the employer's inference of productivity when inferred ability is held constant. The direct effect works through two channels. For any given productivity signal p, additional education leads the employer to infer higher productivity, which increases the return to education by $(1-\lambda)\, q_s$. But additional education also increases the expected value of the p signal, and the increase in p causes the expected return to education to increase by λq_s. These effects sum to q_s, the first term of (12).

Second, in equilibrium, an increase in education causes the employer to increase inferred ability. The rate of increase of inferred ability with respect to education is A', the effect of increased ability on expected log productivity is q_a and the weight that the employer puts on this inference (as opposed to his signal) is $1-\lambda$. The second term, $(1-\lambda)\, q_a A'$, is the product of these effects.

In (12), the term $(1-\lambda)\, q_a A'$ is always nonnegative so that for any equilibrium $S(a)$, we have $q_s(S(a)\,, a) \leq r$. Because we have assumed that q_{ss} is negative, and because the efficient level of education $S^*(a)$ is defined by $q_s(S^*(a)\,, a) = r$, the following proposition holds:

Proposition 5 *Let $S(a)$ describe any separating equilibrium of the workers' signaling game. Then for all $a \in [a_0,\ a_1]$, $S(a) \geq S^*(a)$*

Because an equilibrium strategy profile satisfies $q_s(s,a) = r$ whenever $\lambda(s) = 1$ (see Proposition 4), and because that equation characterizes full-information level of education, we have:

Proposition 6 *Let s^* be the lowest value of s such that $\lambda(s) = 1$ for all $s \geq s^*$, and let $a^* = A(s^*)$. Then for $a \geq a^*$, $S(a)$ is the same as in the case where information about productivity is perfect at all levels of education.*

2.1.3 Example: Ability as the capacity to be educated

We now analyze a special case of this model in which ability is viewed as the capacity to be educated. Let productivity p^* be given by

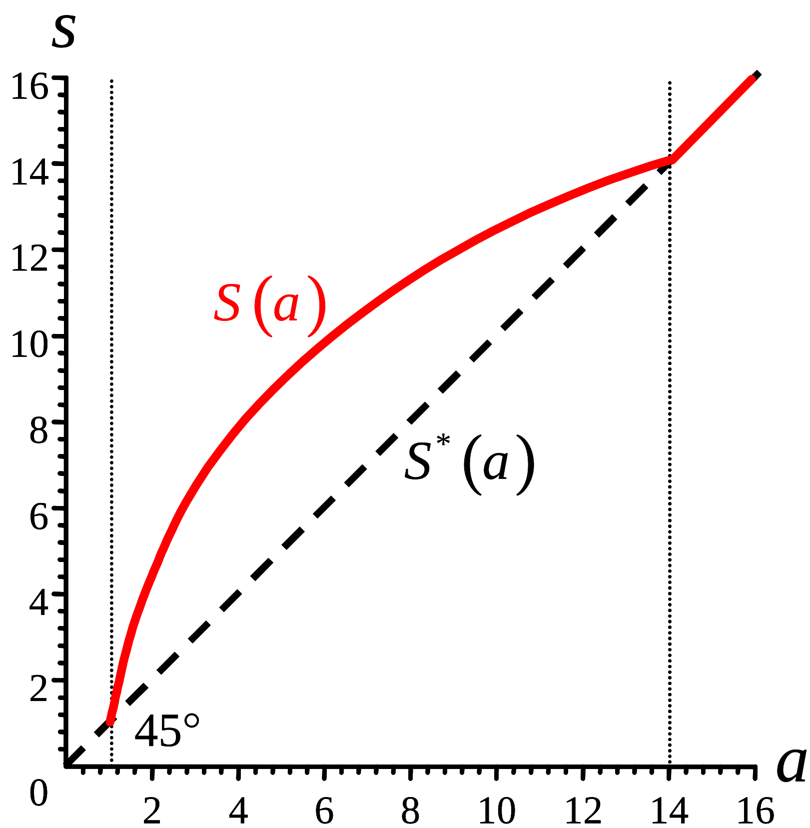
$$p^* = \min\{s,a\}\, \hat{\varepsilon},$$

where $\hat{\varepsilon} = \exp(\varepsilon)$ is a lognormal random variable. This yields a special case of (2) in which $q(s,a)$ is defined by

$$q(s,a) = \min\{\ln s, \ln a\}\,. \tag{16}$$

In this example, additional education is productive only when $s < a$. But when $s < a$, additional ability is not productive, so that the worker has no incentive to use additional education to signal ability. When $s > a$, additional ability is productive but additional education is not, so if the worker obtains additional education, signaling can be his only purpose. Therefore, in this example, we have decoupled the productivity and signaling effects of added education.

Figure 1.



We now find the efficient level of education in this framework. From (16) we have

$$q_s(s,a) = \begin{cases} 1/s & \text{for} \quad s < a \\ 0 & \text{for} \quad s > a \end{cases},$$

which defines the social rate of return to education. Additional education is efficient so long as $q_s(s,a) > r$. This means that the efficient level is given by

$$S^*(a) = \min\left\{\frac{1}{r}, a\right\}. \tag{17}$$

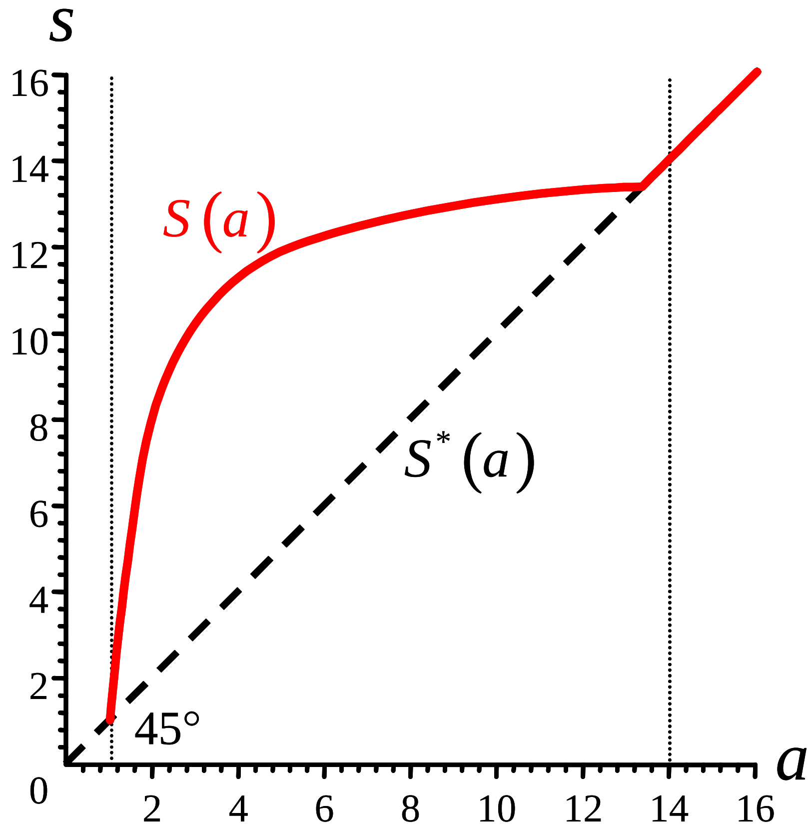
This is the equilibrium level of education when information is perfect ($\lambda = 1$).

For $\lambda(s) < 1$ and $s > a$, equation (13) yields the equilibrium condition

$$S'(a) = \frac{1-\lambda(s)}{r}\frac{1}{a}. \tag{18}$$

Let $\tilde{S}(a)$ be a solution of (18) For Figure 1 we specify $r = .0625$ ($1/r = 16$), and we normalize a so that the lowest level of ability is given by $a_0 = 1$. From (17) we see that the efficient level of

Figure 2.



education $S^*(a)$ increases along the 45-degree line until $s = 16$ and is constant at 16 thereafter. From Proposition 5, it follows that for $a \leq 16$, $S(a) = \tilde{S}(a)$ whenever $\tilde{S}(a) > a$ and $S(a) = a$ otherwise.

Proposition 4 tells us that $S(1) = S^*(1) = 1$, the efficient level of education for $a = 1$. For constant $\lambda(s) = \lambda_0$, the solution of (18)is

$$\tilde{S}(a) = \frac{1 - \lambda_0}{r} \ln a + 1, \tag{19}$$

which describes the equilibrium in the region $s > a$ (above the 45-degree line). The function $S(a)$ is graphed in Figure 1 with λ constant at $\lambda_0 = .692$, a value calibrated to cross the diagonal at $s = 14$.

In Figure 2, we illustrate the situation in which $\lambda(s) = s/b$ (λ increases linearly in s and reaches 1 at $s = b$). In that case, the differential equation for an equilibrium in the region $s > a$ becomes

$$S'(a) = \frac{b - s}{br} \frac{1}{a},$$

and if we require $S(1) = 1$, its unique solution is

$$\tilde{S}(a) = b + (1 - b)\, a^{-\frac{1}{br}}$$

Again, for $a \leq 16$, $S(a) = \tilde{S}(a)$ whenever $\tilde{S}(a) > a$ and $S(a) = a$ otherwise. This is graphed in Figure 2 for $b = 14$. Note that $S(a)$ becomes equal to $S^*(a)$ before $s = 14$ when perfect information is obtained.

2.2 Statistical Discrimination

We are now in a position to model statistical discrimination. The literature on statistical discrimination suggests that firms observe the productivity of blacks less accurately than that of whites. This is almost a convention in the literature, but it can be justified on the grounds that blacks have poorer networks than do otherwise comparable whites or on the basis of language differences. Considerable research shows that blacks and whites use different nonverbal listening and speaking cues and that this can lead to miscommunication (Lang, 1986).

Given that firms can observe the race of applicants, differences in the accuracy of productivity observations induce firms to put a relatively higher weight on education and a lower weight on observed productivity for black workers as compared with white workers. Therefore, education is a more valuable signal of ability for blacks than it is for whites, which leads us to expect blacks to obtain more education than do whites of equal ability. This means that at any level of education, blacks will be of lower ability and have lower wages. However, at any level of ability, since blacks get more education, they should have higher wages if we do not hold education constant. We derive these results formally below.

Let the subscript b denote black workers and w white workers. If black productivity is observed less accurately than white productivity for $s < s^*$, then $\lambda_b(s) < \lambda_w(s)$ there. The following proposition shows that under these circumstances, blacks will get more education than whites of equal ability for all intermediate ability levels.

Proposition 7 *Given $\lambda_b(s) < \lambda_w(s)$ for all $s < s^*$, we have $S_b(a) > S_w(a)$ for all $a \in (a_0,\ a^*)$ in equilibrium.*

Proof. >From (13) we know that for $\lambda_i(s) < 1$, the equilibrium S_b and S_w are characterized by

$$S_i'(a) = \frac{(1 - \lambda_i(s))\, q_a(s,a)}{r - q_s(s,a)}, \tag{20}$$

where i is either b or w. If for $s < s^*$ blacks and whites have the same values of a and s, then from $\lambda_b(s) < \lambda_w(s)$ we know that $S_b'(a) > S_w'(a)$. By the continuity of S_b and S_w and the fact that $S_b(a_0) = S_w(a_0)$, we can infer that $S_b(a) > S_w(a)$ in a neighborhood of a_0. If $\hat{a}$ is the smallest value of a greater a_0 at which $S_b(a) = S_w(a)$, it must be true that $S_b'(\hat{a}) \leq S_w'(\hat{a})$, because $S_b(a)$ is converging to $S_w(a)$ from above. But by (20), this is possible only if $\lambda_b(s) = \lambda_w(s)$, which implies that $\hat{a} = a^*$. The proposition follows. ■

We can now show that at any education level (except the lowest) at which black productivity is observed less accurately than white productivity, the expected equilibrium earnings of blacks is less than that of whites with the same level of education.

Proposition 8 *In equilibrium, for $s \in (s_0, s^*)$, $E_{\varepsilon,u}[\hat{w}_b(p,s)] < E_{\varepsilon,u}[\hat{w}_w(p,s)]$.*

Proof. Equation (10) implies that in equilibrium we have

$$\ln E_{\varepsilon,u}[\hat{w}_i(p,s)] = \lambda_i(s)\, q(s, A_i(s)) + (1-\lambda_i(s))\, \hat{q}_i(s) + .5\sigma_e^2, \tag{21}$$

which reduces to

$$\ln E_{\varepsilon,u}[\hat{w}_i(p,s)] = \hat{q}_i(s) + .5\sigma_e^2, \tag{22}$$

because $\hat{q}_i(s) \equiv q(s, A_i(s))$. From the previous proposition, we know that $\hat{q}_b(s) < \hat{q}_w(s)$ for $s \in (s_0, s^*)$ and the theorem follows. ■

2.3 Empirical Implications of the Model

We have assumed that for low and intermediate levels of education ($s < s^*$), the productivity of black workers cannot be observed as accurately as the productivity of white workers with the same levels of education ($\lambda_b(s) < \lambda_w(s)$). In contrast, the productivity of workers with high levels of education ($s \geq s^*$), and thus ability are observed equally accurately for both races.

The primary implication of our model is that under these circumstances, black workers with low (except for the very lowest level) or intermediate levels of ability will obtain more education than their white counterparts, but black workers of high ability will obtain the same levels of educations as high-ability whites. Thus overall blacks will get more education than do whites of similar ability.

The model also has implications about the measured return to education. If we measure the return to education by comparing wages at an intermediate level of education with wages at the lowest level of education, our model predicts that the measured return to education should be lower for blacks than for whites. However, if we measure the return to education by comparing wages at an intermediate education level with wages at high levels of education, the measured return to education should be higher for blacks than it is for whites. This suggests that as the level of education increases, the measured return to education, not controlling for ability, should initially be lower for blacks than for whites and then become higher for blacks than for whites. Of course, this conclusion refers to the measured return. The actual private return to education is the common interest rate, r, for all workers.

Since, relative to whites with the same ability, blacks with intermediate levels of ability get more education, our model predicts that at these ability levels, blacks should earn more than whites do. At low and high ability levels, blacks and whites should have similar earnings. Put differently, the return to ability (not controlling for education) should be higher for blacks than for whites at relatively low levels of ability and lower for blacks than for whites at somewhat higher ability levels.

We note that the "*ability to learn*" example above demonstrates that our results apply more generally than simply to the case in which there is no asymmetric information beyond some level of education. In the case graphed in Figure 1, with imperfect information, the wage paid to workers with a given level of education is lower than it is with perfect information whenever the two education levels diverge. Relative to the case of perfect information, with imperfect information, the estimated return to education would be lower at low levels of education and higher at high levels of education.

3 Data

Although our initial focus is on differences in educational attainment not wages, later in the paper we will want to place our results in juxtaposition with those of Neal and Johnson. Therefore to a large extent, we mimic their procedures. Following NJ, we rely on data from the National Longitudinal Survey of Youth (NLSY79). Since 1979 the NLSY has followed individuals born between 1957 and 1964. Initially surveys were conducted annually. More recently, they have been administered every other year. In the period we use, the NLSY oversamples blacks and Hispanics (in the early years of the study, people from poor families and the military were oversampled as well). We use sampling weights to generate representative results.

The education variable is given by the highest grade completed as of 2000. For those missing the 2000 variable, we used highest grade completed as of 1998 and for those missing 1998 as well, we used the 1996 figure. Where available we used the 1996 weight. For observations missing the 1996 weight, we imputed the weight from the 1998 and 2000 weights using the predicted value from regressions of the 1996 weights on the 1998 and/or 2000 weights.

We determined race and sex on the basis of the sub-sample to which the individual belongs. Thus all members of the male-Hispanic cross-section sample were deemed to be male and Hispanic regardless of how they were coded by the interviewer.

In 1980, the NLSY administered the Armed Services Vocational Aptitude Battery (ASVAB) to members of the sample. A subset of the ASVAB is used to generate the Armed Forces Qualifying Test (AFQT) score. The AFQT is generally viewed as an aptitude test comparable to other measures of general intelligence. Like other such measures, it is generally regarded as reflecting a combination of environmental and hereditary factors. The AFQT was recalibrated in 1989. The NLSY data provide the 1989 AFQT measure. Following NJ, we regressed the AFQT score on age (using the 1981 weights) and adjusted the AFQT score by subtracting age times the coefficient on age. We then renormed the adjusted AFQT to have mean zero and variance one.

In the later part of the paper, we also examine wages. Because of the difficulties in addressing differential selection into labor force participation of black and white women (Neal, 2004), we limit our estimates using wages to men. In order to minimize the problem of missing data, we used hourly earnings data from the 1996, 1998 and 2000 waves of the survey. Next we took all observations with hourly wages between $1 and $100 in all three years and calculated (unweighted) mean hourly earnings for this balanced panel. We used the average changes in hourly wages to adjust 1996 and 2000 wages to 1998 wages. Note that this adjustment includes both an economy-wide nominal wage growth factor and an effect of increased experience. We then used the adjusted 1996, 1998 and 2000 wages for the entire sample to calculate mean adjusted wages for all respondents. We limited ourselves to observation/years in which the wage was between $1 and $100. If the respondent had three valid wage observations, we used the mean of those three. If the respondent had two observations, we used the average of those two. For those with only one observation, the wage measure corresponds to that adjusted wage. There were 237 observations of men who were interviewed in at least one of the three years but who did not have a valid wage in any of the three years. In the quantile regressions, these individuals are given low imputed wages except for five cases coded as missing for which the reported wage in at least one of the three years exceeded $100 per hour and for which there was no year with a valid reported wage.

4 Differences in Educational Attainment

Most labor economists are aware that the average education level is lower among blacks than among whites. In our sample blacks get about three-quarters of a year less education than do whites. It is less well known that conditional on AFQT, blacks get more education than whites do. This is shown in the first row of Table 1.[3] In the first and fifth columns, we show the difference in educational attainment between blacks and non-Hispanic whites among men and among women conditional on age and AFQT. Black men get about 1.2 years more education than do white men with the same AFQT. Among women the difference is about 1.3 years. There are also smaller but statistically significant differences between Hispanics and non-Hispanic whites (not shown).

4.1 School Quality

The difference in educational attainment, conditional on AFQT, between blacks and whites is predicted by our theoretical model, but there are, as always, a large number of other potential explanations for this finding. One is that AFQT is largely determined by schooling and that since blacks attend lower quality schools, on average they gain fewer cognitive skills from a given level of education. Under this view, blacks have more schooling given their AFQT because they require more schooling to reach a given level of cognitive skills. When we regress education on AFQT, we are, in effect, estimating a reverse regression.

To take a simple example, suppose that

$$a = q * s \tag{23}$$

where a is AFQT, q is school quality and s is years of education. Then if q is lower for blacks than for whites, the only way that a black and white can have the same a is if the black has higher s.

In the simple example, there are no innate ability differences that influence AFQT. More generally, lower school quality among blacks will only explain their greater educational attainment given their AFQT if the effect of schooling on AFQT is sufficiently large.

Consider now the opposite extreme. Suppose that AFQT is unaffected by school quality. How would school quality affect years of education conditional on AFQT? Put differently, of two people

[3] See Rivkin (1995) for findings from High School and Beyond that conditional on math and reading scores, blacks are more likely to remain in high school and begin college. Cameron and Heckman (2001) also use the NLSY and find that blacks get more education than whites conditional on measures of family background and note that AFQT has a particularly strong effect on reversing the education differential.

TABLE 1
Educational Attainment of Blacks Relative to Non-Hispanic Whites

Men				**Women**	
B/W	N	Birth Years	Other Controls	B/W	N
1.17 (0.10)	4060	All		1.30 (0.09)	4337
1.15 (0.14)	2302	All	school inputs	1.25 (0.14)	2326
1.11 (0.16)	2333	All	school composition	1.29 (0.16)	2385
1.20 (0.11)	3323	All	family background	1.39 (0.16)	3558
1.16 (0.20)	1603	All	school inputs, school composition, family background	1.33 (0.20)	1618
0.92 (0.14)	1719	1962+	grade in 1980	1.22 (0.15)	1665
0.95 (0.18)	1106	1963+	grade in 1980	1.31 (0.19)	1054
0.71 (0.26)	508	1964+	grade in 1980	1.43 (0.26)	474
0.97 (0.14)	1737	1962+		1.29 (0.15)	1683
0.99 (0.18)	1116	1963+		1.36 (0.19)	1062
0.77 (0.26)	514	1964+		1.40 (0.26)	478
0.86 (0.21)	913	1962+	grade in 1980, school inputs	1.27 (0.23)	862
1.04 (0.25)	914	1962+	grade in 1980, school composition	1.28 (0.26)	889
1.01 (0.17)	1385	1962+	grade in 1980, family background	1.36 (0.17)	1355
1.06 (0.31)	630	1962+	grade in 1980, school inputs, school composition, family background	1.28 (0.29)	592

Standard errors in parentheses. All estimates control for age. **School inputs**: log enrollment, log no. of teachers, log no. of guidance counselors, log library books, % teachers with MA/PhD, % teachers left during year, average teacher salary. **School composition**: % disadvantaged, daily attendance rate, dropout rate, % students Asian, % students black, % students Hispanic. **Family background**: mother's education, father's education, no. of sibling, born in US, lived in US at age 14, lives in urban area at age 14, mother born is US, father born in US.

with IQ's of 100 (or normalized AFQT's of 0), would we expect the one in a higher quality school to get more or less education than the one in the lower quality school?

Most labor economists would expect that holding other factors constant, lower school quality would lower years of education. Standard theoretical models do not offer us unambiguous results about the effect of school quality on years of schooling. In these models, the sign of the effect depends on second derivatives. The data, however, suggest a positive correlation between school quality and years of schooling (e.g. Card and Krueger 1992a&b; Hanushek, Lavy and Hitomi, 2006 and the discussion in Hanushek and Wößmann, 2007). In this case, the lower school quality faced by blacks should lead them to get fewer, not more, years of schooling than whites with the same AFQT.

To summarize, if AFQT is heavily influenced by education and if most sample members had completed their education at the time that they took the AFQT, then school quality differences would provide a plausible explanation for the higher education among blacks given their AFQT. If school quality has little effect on AFQT or if most sample members had not completed schooling, then we would expect blacks to get less education given their AFQT or given their AFQT and completed schooling at the time they took the test. Our own view is that the AFQT measures skills that are more heavily affected by preadolescent and early adolescent education so that the endogeneity of AFQT to ultimate educational attainment and the consequent link to school quality is not likely to be a major issue. However, others certainly disagree. Therefore we address the question empirically.

Our first approach is to measure the education differential conditional on measured secondary school inputs as measured by log enrollment, log number of teachers, log number of guidance counselors, log library books, the proportion of teachers with an MA or PhD, the proportion of teachers who left the school during the year and the average teacher salary. To conserve space, we do not report the coefficients on these controls. However, almost none of the individual coefficients is statistically significant. Among men, attending a school with more highly educated teachers is associated with greater educational attainment. Among women this variable and attending a school with more library books is associated with getting more education. In part, the paucity of individually significant factors reflects multicollinearity among the measured inputs. For both men and women, the coefficients on the school inputs are jointly significant. More importantly, controlling for these factors has almost no effect on the estimated education gaps, and some of the change is due to a change in the sample rather than to the effect of the additional controls.

Because inputs may be a very poor proxy for school quality, in the third row we control for measures of school composition and student behavior. These are designed to capture some of the elements that people think about when they think about struggling schools: high proportions of disadvantaged students, high dropout rates and poor attendance. We also control for the racial and ethnic composition of the school. The results are very similar to those we obtained in the first two rows.

Moreover, consistent with Cameron and Heckman (2002), as shown in the fourth row, the findings are robust to including measures of family background. Controlling for mother's and father's education, number of siblings, whether each parent was born in the United States, whether the respondent was born in the United States and whether he lived in an urban area at age fourteen has little or no effect on the coefficient on black.

Finally, the fifth row provides our most complete set of controls using the full sample. Controlling simultaneously for school inputs, school composition and family background costs us about sixty percent of the sample but has a negligible effect on the black/white education differential.

4.2 Younger Cohorts

Another way to control for the endogeneity of AFQT is to limit the sample to individuals who would not have completed their education at the time the NLSY administered the AFQT. Regardless of whether or not AFQT is endogenous to education, for these cohorts we have a measure of ability prior to their completion of schooling. We experiment with different age cutoffs and with using only AFQT and both AFQT and educational attainment at the time of testing as "ability" controls.

The sixth row shows the differential for individuals born after 1961. Only about 5% of this sample had completed schooling when they took the AFQT. While their AFQT may have been influenced by their education up to this point, future education should be caused by skills acquired up to this point and not the other way around. The estimated differential is somewhat smaller using the younger cohorts, particularly for men, but it remains substantial. The next two rows further restrict the cohorts we examine. When we restrict the sample to the youngest birth year in the data, the education differential is somewhat smaller for men and somewhat larger for women, but in neither case can we reject equality between the differential for the youngest cohort and the differential for the three youngest cohorts.

The next three rows repeat the exercise but do not control for grade completed at the time the exam was administered. Dropping this variable has only a small effect on the results.

The final four rows replicate rows 2-5 for the younger cohorts. As in the case of the full sample, controlling for measures of school quality and family background has little effect on the education differential between blacks and whites (conditional on AFQT and education in 1980). The last specification is the most restrictive we estimate. It limits the sample to those born in 1962 or later and controls for school inputs, school composition and family background. The estimated differentials are very close to those reported for the full sample with no controls (row 1).

In short, there is a robust result that among men of equal ability as measured by the AFQT, blacks get a year or more education than do non-Hispanic whites. Among women this differential is about 1.25 years.

Before we move on, it is important to make it clear what we are *not* claiming. As stated in the introduction, we are not claiming that AFQT is innate or even unaffected by education and school quality. And we are not claiming that school quality is unrelated to educational attainment or that school quality does not differ between blacks and whites. To the contrary, we believe that average school quality is lower for blacks and that individuals who attend higher quality schools both have higher AFQT's and get more education. It is beyond the scope of the paper to address whether these last two relations are causal. However, from our perspective, the simplest and most probable explanation for our results is that the effect of school quality on AFQT and the effect of school quality on educational attainment roughly cancel so that AFQT given educational attainment is roughly independent of school quality.

4.3 Affirmative Action in Education

Finally, we address an additional explanation that is frequently raised in seminars: affirmative action in college admissions. Since affirmative action is practiced at only a small number of elite colleges (Kane, 1998), this explanation cannot account for the large average difference in educational attainment that we observe. Moreover, we will see in the next section that the education difference is found throughout much of the range of AFQT scores including scores at which attendance at such elite colleges is highly unlikely. In fact, the education gap disappears at AFQT scores above roughly 1.5 standard deviations above the mean, the level at which we might expect such affirmative action to play a role.

5 Further Evidence

While our model predicts that conditional on ability, blacks will, on average, get more education than whites, it makes stronger predictions which we examine in this section.

5.1 Education and Ability

We begin with the prediction about the relation between educational attainment and ability. We have already seen that conditional on AFQT, blacks get more education relative to whites. Our model suggests that this should be true at intermediate levels of ability but not at very low or very high levels of ability. Note that this prediction stands in contrast with the affirmative action in education model which suggests that the education differential should be found only at relatively high levels of ability.

Table 2 shows the relation between education and AFQT, separately for men and women. Within each sex the younger cohorts, who had not completed school at the time that they took the AFQT, are shown separately, both with and without a control for their completed education at the time they took the test. In every specification, the interaction of race and the AFQT-squared term has its predicted negative sign. This is true for Hispanics as well as for blacks.

Although the individual interaction terms are generally not statistically significant when we limit the sample to the younger cohorts, in no case are the differences between the young and older cohorts in the three black interaction terms, the three Hispanic interaction terms or the six

TABLE 2
AFQT AND EDUCATIONAL ATTAINMENT BY RACE AND SEX

	Men			Women		
	All	Young Cohorts		All	Young Cohorts	
Constant	12.12 (0.23)	13.82 (0.78)	14.62 (0.79)	12.05 (0.23)	12.47 (0.81)	13.41 (0.81)
AFQT	1.64 (0.04)	1.61 (0.06)	1.43 (0.07)	1.67 (0.04)	1.70 (0.08)	1.43 (0.09)
AFQT2	0.57 (0.04)	0.43 (0.06)	0.48 (0.06)	0.32 (0.04)	0.24 (0.07)	0.37 (0.07)
Black Interactions						
Constant	1.28 (0.13)	1.02 (0.19)	0.96 (0.19)	1.40 (0.12)	1.32 (0.20)	1.20 (0.19)
AFQT	-0.13 (0.12)	-0.18 (0.19)	-0.15 (0.19)	-0.01 (0.13)	0.03 (0.21)	0.13 (0.21)
AFQT2	-0.41 (0.10)	-0.27 (0.16)	-0.29 (0.16)	-0.29 (0.11)	-0.17 (0.20)	-0.14 (0.20)
Interaction Equals 0	-1.94 1.63	-2.30 1.64	-2.09 1.57	-2.21 2.18	-2.67 2.91	-2.47 3.36
Hispanic Interactions						
Constant	0.68 (0.17)	0.29 (0.26)	0.30 (0.26)	0.99 (0.15)	0.73 (0.25)	0.78 (0.25)
AFQT	0.09 (0.13)	0.03 (0.21)	0.04 (0.21)	0.02 (0.16)	-0.07 (0.25)	-0.05 (0.25)
AFQT2	-0.48 (0.12)	-0.06 (0.21)	-0.09 (0.21)	-0.66 (0.12)	-0.38 (0.22)	-0.46 (0.23)
Interaction Equals 0	-1.10 1.28	-1.97 2.40	-1.59 2.03	-1.21 1.24	-1.47 1.29	-1.35 1.24
Other controls	Age	Age	Age, Education in 1980	Age	Age	Age, Education in 1980
N	4060	1737	1719	4337	1683	1665

Standard errors are in parentheses. Weights for education results are described in text.

interaction terms statistically significant.[4] Thus the results are not driven by the causal impact of education on AFQT.

For men, the black-white education differential is maximized at an AFQT about one-sixth standard deviation below the mean where it is about 1.3 years. Educational attainment is equal for blacks and whites at almost two standard deviations below the mean and at one and two-thirds standard deviations above the mean. For women, the black-white education differential is maximized just about at the mean AFQT where it is about 1.4 years. The education levels of black and white women are estimated to be equalized pretty much at the extremes of the AFQT distribution.

We note that the "affirmative action in education" explanation for higher education levels among blacks would imply that the differences in education would be largest at the high levels of AFQT associated with application to selective colleges. Thus the results are not consistent with that explanation.

Figure 3 shows the smoothed relation between education and AFQT for men. The nonparametric approach (which ignores the relation between age and education) confirms the parametric approach. Education levels for blacks and whites converge around a standardized AFQT of -2 and a little above 1.5. Figure 4, for women, is less consistent with the parametric estimates. It shows that education levels converge at a standardized AFQT between -2 and -2.5. However, education levels for black women remain higher than for white women even at very high AFQT levels. One potential explanation for this difference is the very high rate of labor force participation of high-skill black women relative to white women discussed in Neal (2004).

5.2 Education and Earnings

Because of the complications associated with differences in the selection of black and white women into the labor force, our discussion of the wage predictions is restricted to men. Our model implies that the wages of blacks and whites will be similar at low levels of education and at high levels but that blacks will have lower wages at intermediate levels of education. To test this prediction, we regress the log wage on education and its square and interactions with race and ethnicity as well as direct effects of age, race and ethnicity. Table 3 shows the results.

For all four specifications (all/young cohorts, with/without controls), as predicted, the return to education is initially lower for blacks than for whites and then turns more positive. With the full sample and no controls, wages for blacks and whites are estimated to be equal for those with a fifth grade education and those with nineteen years of completed education although these points of equality are estimated very imprecisely. In the most restrictive specification (only the young cohorts and a complete set of controls), we estimate that wages are equal for blacks and whites

[4] The interaction between Hispanic and $AFQT^2$ does differ at the .05 level for men. However, given that we are testing multiple equalities as well as some combinations, it is not surprising that we would find one "significant" test statistic.

Figure 3.

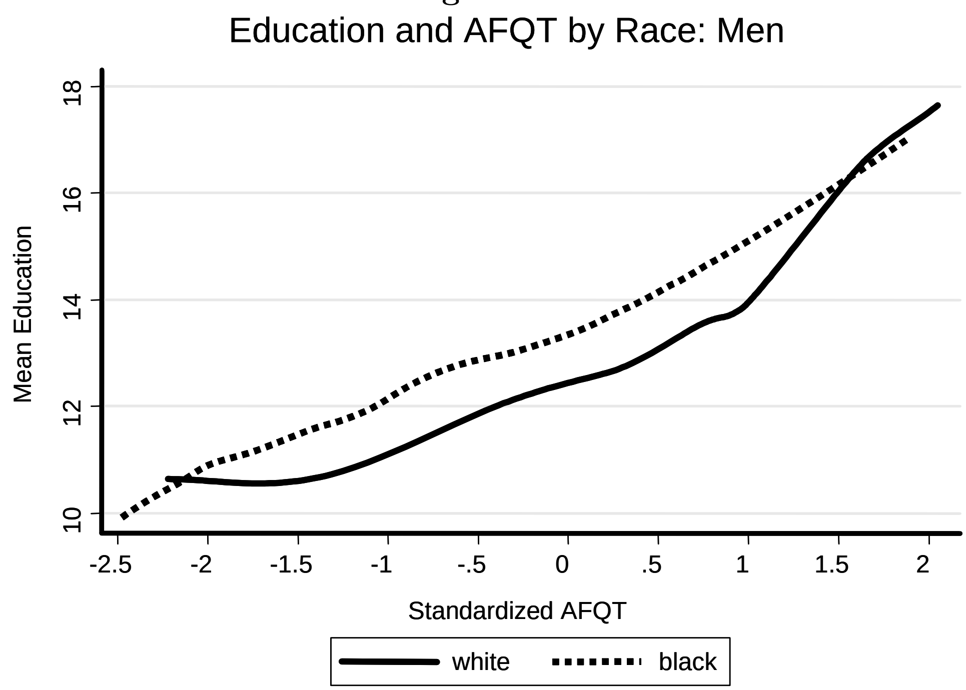


Figure 4.

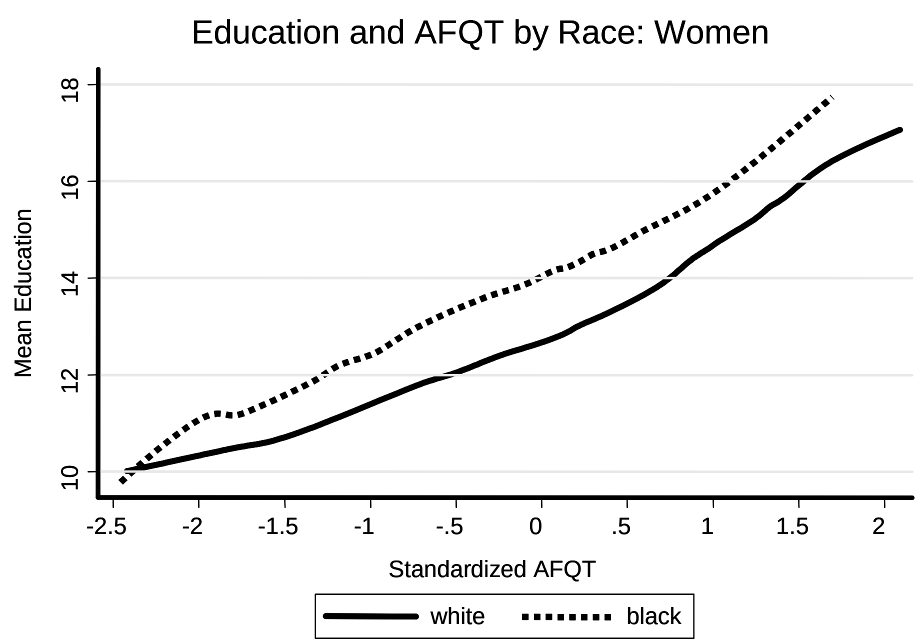
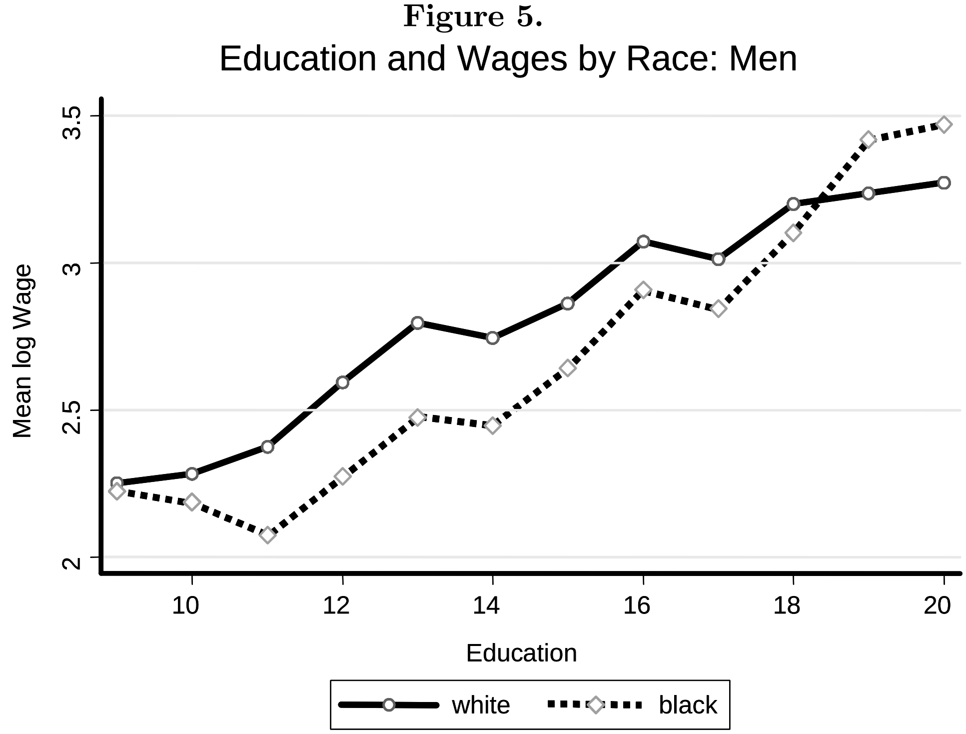


TABLE 3
BLACK-WHITE EARNINGS DIFFERENTIALS BY EDUCATIONAL ATTAINMENT

	All		Young Cohorts	
	No Controls	Controls	No Controls	Controls
Black	0.53 (0.49)	1.23 (1.27)	1.89 (0.93)	3.53 (2.17)
Black*Education	-0.14 (0.07)	-0.24 (0.19)	-0.34 (0.13)	-0.58 (0.32)
Black*Education^2/ 100	0.60 (0.26)	0.99 (0.66)	1.32 (0.50)	2.23 (1.16)
Grades at which wages are equal	5, 19	7, 17	8, 18	10,16
N	4041	1601	1663	607

Also controls for age, school inputs, school composition and family background (see table 1 for details) and, for young cohorts only, grade in 1980.

Figure 5.



who have completed college or who have completed only grade ten. However, these estimates are even more imprecise than those using the whole sample and no controls.

Figure 5 shows the relation between education and earnings nonparametrically for the full sample. It plots average wages (on a log scale) for men by education and race. There are very few individuals without any high school education and very few blacks with more than eighteen years of education. The estimates suggest that, as predicted by the model, wages are very similar for blacks and whites at low and high levels of education.[5]

TABLE 4
BLACK-WHITE EARNINGS DIFFERENTIALS BY AFQT SCORE

	All		Young Cohorts	
	No Controls	Controls	No Controls	Controls
Black	-0.08 (0.03)	-0.05 (0.07)	-0.13 (0.05)	-0.04 (0.10)
Black*AFQT	0.05 (0.03)	0.03 (0.05)	0.05 (0.05)	-0.02 (0.08)
Black*AFQT^2/100	0.64 (2.62)	0.03 (4.70)	2.78 (4.12)	0.68 (8.20)
F-test on all three coefficients	4.77	0.46	4.76	0.06
F-test on interaction terms	1.14	0.22	0.58	0.04
N	3841	1534	1637	600

Also controls for age, school inputs, school composition and family background (see table 1 for details) and, for young cohorts only, grade in 1980.

5.3 Wages and AFQT

Our model also predicts that, except at very high and very low levels of ability, blacks should earn more than equally able whites. The test of this prediction is shown in Table 4. The table follows the same format as Table 3. With no controls, this prediction is soundly rejected both for all cohorts and for the young cohorts. In both cases, there is no evidence of an interaction between race and AFQT although the point estimates suggest that black and white wages are equalized at an AFQT about 1.4 standard deviations above the mean. Below that level, blacks earn less, not more, than whites with the same AFQT. When we add controls for family background, all of the coefficients on black and its interaction terms become individually and jointly insignificant. Nevertheless, the

[5] The model also implies that at low levels of education, the variance of earnings will be lower for blacks than for whites but that this difference will disappear at higher levels of education. When we regress the squared residual from the regression of the lwage on education, education squared, age and race/ethnicity on race/ethnicity and interactions with education, the point estimates confirm the hypothesis but are so imprecisely estimated as to not be meaningful.

estimates using all cohorts imply that the black-white wage differential declines from about 10% at an AFQT two standard deviations below the mean to zero at 1.4 standard deviations above the mean. The estimates using only the young cohorts and controls are too imprecise to be meaningful. If one accepts the results with controls, then they are consistent with the model when education is a pure signal at the margin. If, instead, we rely on the results without controls, then the model must be supplemented with some other explanation for wage differentials.

If, as is generally accepted among labor economists, education is rewarded in the labor market, then in the absence of labor market discrimination, blacks should earn more than whites with the same AFQT. Given the education differential, the absence of a wage differential favoring blacks when we control only for AFQT suggests that blacks are not rewarded fully for their skills, a point to which we now turn.

6 Premarket versus Market Discrimination

There is a heated debate among labor economists about the extent to which black-white wage differentials can be ascribed to premarket factors (including discrimination outside the labor market) rather than to labor market discrimination.[6] One of the critical issues in this debate is for which factors we should control.

Table 5 presents estimates in the tradition of this research. In each case, we present ordinary least squares estimates using only the cohorts born in 1962 or later in the first column and using all cohorts in the second column. In no case are the estimates for the restricted and full samples statistically significantly different at even the .1 level, and, for the most part, the substantive interpretations of the results are similar. Therefore we concentrate on the more precisely estimated results in the second column while reminding readers who are concerned by the potential endogeneity of AFQT to either employment or schooling that the results for the younger cohorts are similar. In the third column, we present the results of median regressions using the full sample. This last set of estimates serves to address selection issues since black men are noticeably more likely than are white men not to have a wage.

[6] In addition to the papers by O'Neill and Neal and Johnson, discussed in the introduction, see Johnson and Neal (1998) and the critiques in Rodgers and Spriggs (1996, 2002) and Darity and Mason (1998) and the reply by Heckman (1998).

TABLE 5
BLACK-WHITE WAGE DIFFERENTIALS

			Other Controls		
Young Cohorts	All	Median Regression	AFQT	Education	family background, school inputs
-0.36 (0.04) [1637]	-0.36 (0.02) [3841]	-0.42 (0.03) [4055]			
-0.13 (0.04) [1637]	-0.09 (0.03) [3841]	-0.10 (0.03) [4055]	✓		
-0.17 (0.03) [1637]	-0.15 (0.02) [3841]	-0.18 (0.03) [4055]	✓	✓	
-0.07 (0.06) [732]	-0.06 (0.04) [1876]	-0.05 (0.03) [1955]	✓		✓
-0.11 (0.06) [732]	-0.11 (0.04) [1876]	-0.11 (0.04) [1955]	✓	✓	✓

All estimates control for age. Estimates for young cohorts control for education completed in 1980.

For purposes of comparison, the first row shows the very large differential that exists when we control only for age. However, our focus is on the wage differential that exists after we control for ability in the form of AFQT and its square. This is shown in the second row. Consistent with our findings in table 4 that the race/AFQT interaction terms are statistically insignificant, we drop these terms in table 5, which also simplifies interpretation. The second row suggests much more modest wage differentials although they are not trivial and are somewhat higher than in Neal and Johnson's study of the younger cohorts in the early 1990s.[7]

In the fourth row, we also control for family background and school inputs.[8] The estimated wage differential becomes noticeably smaller and statistically insignificant in all three specifications. Somewhat surprisingly, the median differential is slightly smaller than that obtained using standard

[7] Derek Neal was very helpful, supplying us with the code to replicate his and William Johnson's results. The modest difference in our results derives from a number of differences including NJ's use of the "class of worker" variable and our use of a later time period. Carneiro, Heckman and Mastrov (2004) explores the issue of time variation in the black-white wage differential using various specifications including those used by NJ. See also the discussion of this issue in Haider and Solon (2004).

[8] We dropped school composition after pretesting showed that these variables were highly insignificant but that their inclusion reduced the sample sufficiently to increase noticeably the imprecision of the estimate of the black-white wage differential. The decision to include or exclude school composition has no effect on the interpretation of table 4.

regression.

However, the important point in Table 5 is the comparison of the second and third rows and of the fourth and fifth rows. In each case the latter specification is identical to the former except that it also controls for educational attainment. In each case the estimated differential increases substantially when we include education in the equation.[9]

Although Neal and Johnson explore the effect of also controlling for education to some extent, they explicitly reject including education in their main estimating equation. They provide two arguments for their position. First, they maintain that we should examine black-white wage differentials without conditioning on education because education is endogenous. Their argument would be much more compelling if blacks obtained less education than equivalent whites. In that case, we might argue that blacks get less education because they expect to face discrimination in the labor market, and therefore controlling for education understates the importance of discrimination.

However, if blacks obtain *more* education because they anticipate labor market discrimination as we argue in this paper, *failing to control* for education understates the impact of discrimination. Consider the following example. Suppose that the market discriminates against blacks by paying them exactly what it would pay otherwise equivalent whites with exactly one less year of education. Then, to a first approximation,[10] all blacks will get one year more education than otherwise equivalent whites. Controlling only for ability, we find that blacks and whites will have the same earnings, but controlling for education as well as ability, we see that blacks earn less than whites by an amount equal to the return to one year of education.

Note that even if the higher educational attainment among blacks reflects premarket factors, it may still be appropriate to control for education when measuring discrimination in the labor market. After all, we would still anticipate that the labor market would compensate blacks for their additional education regardless of their reason for getting more education.

The second argument that NJ make is that education is a poor proxy for skills. In particular, on average, blacks attend lower quality schools than do whites. Whites will have more effective education than do blacks with the same nominal years of completed education. We have already noted that students who attend lower quality schools tend to get less education. Therefore if blacks attend lower quality schools, for any given level of education, they will have higher ability. If in the regression of wages on education and AFQT, AFQT only partially controls for ability, blacks will tend to have higher ability than do whites with the same education, and the coefficient on black will be spuriously positive. Therefore, differential school quality could lead to a spurious positive or negative coefficient on "black."

The last row in Table 5 shows that controlling for both family background and school inputs

[9] Carneiro et al (2004) use a specification similar to that in row (4) but adjust AFQT for schooling completed at the time the respondent took the AFQT. They find simialr results.

[10] This statement is precise if all workers maximize the present discounted value of lifetime earnings, lifetimes are infinite, there are no direct costs of education and the return to experience is zero.

somewhat reduces the estimated differential relative to the estimate in the third row (which controls for education but not for these additional variables). Although this seems to suggest that it is important to control for school quality, the reduction in the black-white wage differential actually results from the family-background controls.

Table 6 shows the full set of results when we control only for school quality as measured by either inputs or composition. The results are similar when we control for both simultaneously. Most of the remaining coefficients have the anticipated sign. Holding other resources constant, larger schools are associated with lower wages. Holding enrollment constant, having more guidance counsellors, more teachers and more library books are associated with higher wages. Having more educated teachers and higher paid teachers is associated with higher student earnings while teacher turnover has a negative effect.

Yet, controlling for inputs indicates that there is almost no effect on the measured black-white wage differential. The difference between the coefficients with and without school quality controls reflects differences in the sample rather than the effect of adding the controls. The coefficient on black using the observations for which we have school input measures is -0.14. At least as measured by inputs, differences in school quality do not account for the black-white wage differential.

The right-side of Table 6 controls for measures of student composition and behavior. Perhaps surprisingly, this effort is in some ways less successful than the estimation using school inputs. While higher fractions of disadvantaged students and dropouts are associated with lower wages, average absenteeism and the fraction of students who are black are not. The results are again quite similar to those obtained without controls for school quality.

Thus we find no evidence that the wage and education differentials are driven by differences in school quality. It is important to note that the absence of evidence for the role of these premarket factors does not depend on a causal interpretation of the relation between education quality and outcomes. It is entirely possible that attending a school with a higher dropout rate does not make any individual more likely to dropout. Students who attend schools with high dropout rates may have characteristics that make them more likely to dropout. Even if the dropout rate were merely a proxy for these unmeasured characteristics, we would expect including the dropout rate to lower the black-white education differential. The fact that it does not, supports the view that such premarket differences do not explain the wage and education differentials.

7 Discussion and Conclusion

While some of the principal predictions of the theory we presented are consistent with the data, it is important to recognize that the combination of statistical discrimination and educational sorting that we discuss cannot fully explain the data. Our model implies that, conditional on ability, relative to whites, blacks get more education. This, in turn, implies that conditional on AFQT, blacks should earn more than whites. But neither our results or any that we are aware of in the literature support that conclusion for men.

TABLE 6
DETERMINANTS OF LOG WAGES
USING CONTROLS FOR SCHOOL QUALITY. N&J Wages

Inputs		**Student Composition/Behavior**	
Black	-0.14 (0.04)	-0.14 (0.04)	Black
Hispanic	-0.03 (0.05)	-0.01 (0.06)	Hispanic
Age/10	0.13 (0.04)	0.14 (0.04)	Age/10
Education	0.06 (0.01)	0.06 (0.01)	Education
AFQT	0.14 (0.01)	0.15 (0.01)	AFQT
Log(Enrollment)	-0.08 (0.04)	-0.08 (0.06)	Proportion Disadvantage
Log(Teachers)	0.03 (0.05)	-0.05 (0.07)	Proportion Daily Attendance
Log(Guidance)	0.10 (0.04)	-0.10 (0.05)	Proportion Dropout
Log(Library books)	0.01 (0.01)	0.08 (0.06)	Proportion Students Black
Proportion Teachers MA/PhD	0.17 (0.05)	-0.08 (0.10)	Proportion Students Hispanic
Teacher Salary $0,000s	0.18 (0.09)	0.81 (0.43)	Proportion Students Asian
Teachers who left/100	-0.30 (0.13)		
N	2194	2223	N

Standard errors are in parenthesis. Weights were the same as the education results in Tables 3 and 4.

One possible explanation is that education is a pure signal at the margin. This is the case in our "ability to learn" example. In that example, while education is productive up to some point that depends on the worker's ability, it is unproductive beyond that point. In order to signal their ability, most workers invest in education beyond the point at which it increases their productivity. However, we view this model as extreme.

Our model and the supporting empirical evidence identifies statistical discrimination as one source of differences in outcomes for blacks and whites.[11] We have focused our attention on only one effect, increased investment in the observed signal. Blacks may also invest less in unobservable skills as in Lundberg and Startz which would lead to them have lower wages even conditional on AFQT. In addition, the work of Bertrand and Mullainathan (2004) on names and job applications suggests to us that statistical discrimination is of particular importance in the presence of search frictions. They find that applicants with African American names are less likely to receive calls for interviews than are similar applicants with names common among whites. If evaluating workers is costly, statistical discrimination may prevent large numbers of African American workers from consideration for many jobs. We expect that in this setting our principal results would hold: African Americans would have greater incentives to signal their productivity and would earn less conditional on their education. However, it is also likely that they would earn less conditional on their ability.

In our view, the results in this paper cast doubt on an emerging consensus that the origins of the black-white wage differential lie in premarket rather than labor market factors. Blacks earn noticeably less than whites with the same education and cognitive score. The evidence is not consistent with the view that the unexplained differential reflects differences in school quality, the principal premarket explanation. Thus, there are good grounds for believing that at least some of the black-white wage differential reflects differential treatment in the labor market.

[11] Our results stand somewhat in contrast to Altonji and Pierret (2001) who find that the black-white differential does not decline with experience when one controls for hard to observe measures of productivity. However, since our model implies that the estimated return to education changes differentially with experience for blacks and whites, it is not clear that the Altonji/Pierret evidence is inconsistent with our version of statistical discrimination.

References

Aigner, Dennis, and Cain, Glen, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 30 (1977): 175-87.

Altonji, Joseph G. and Pierret, Charles R., "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 116 (February 2001): 313-50.

Bertrand, Marianne and Mullainathan, Sendhil, "Are Emily and Brendan More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review,* 94 (September 2004): 991-1013.

Cameron, Stephen V. and Heckman, James J., " The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy*, 109 (June 2001): 455-99.

Card, David, and Krueger, Alan B., "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics*, 107 (February 1992a): 151-200.

______, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100(February 1992b): 1-40

Carneiro, Pedro, Heckman, James J. and Masterov, Dimitriy, "Labor Market Discrimination and Racial Differences in Pre-Market Factors," *Journal of Law and Economics* (2004), forthcoming.

Coate, S. and Loury, G. C., 1993. "Will affirmative-action policies eliminate negative stereotypes?" *American Economic Review* 83 (5), 1220–1240.

Darity, William A., Jr. and Mason, Patrick L., "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender," *Journal of Economic Perspectives*, 12 (Spring 1998): 63-90.

Heckman, James J. "Detecting Discrimination," *Journal of Economic Perspectives*, Spring 1998; 12(2): 101-16.

Johnson, William and Neal Derek, "Basic Skills and the Black-White Earnings Gap," in Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*, Washington, DC: Brookings Institution Press, 1998.

Kane, Thomas J., "Racial and Ethnic Preferences in College Admission," in Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*, Washington, DC: Brookings Institution Press, 1998.

Kirschenman, Joleen and Neckerman, Kathryn M., "'We'd Love to Hire Them, But...': The Meaning of Race for Employers," in Christopher Jencks and Paul E. Peterson, eds., *The Urban Underclass*, Washington, DC: Brookings Institution Press, 1991.

Lang, K. "A Language Theory of Discrimination," *Quarterly Journal of Economics*, 101 (May 1986): 363-382.

Lundberg, Shelly J. and Startz, Richard, "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review*, 73 (1983): 340-7

Neal, Derek, "The Measured Black-White Wage Gap Among Women is Too Small," *Journal of Political Economy*, 112 (February 2004): S1-28.

Neal, Derek A. and Johnson, William R., "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy* 104 (October 1996): 869-95.

Newman, Katherine S., *No Shame in My Game*, New York: Vintage Books, 1999.

O'Neill, June, "The Role of Human Capital in Earnings Differences Between Black and White Men," *Journal of Economic Perspectives,* 4 (Autumn 1990): 25-45.

Rivkin, Steven G., "Black/White Differences in Schooling and Employment," *Journal of Human Resources*, 30 (Fall 1995): 826-52.

Rodgers, William M., III, and Spriggs, William E. "The Effect of Federal Contractor Status on Racial Differences in Establishment-Level Employment Shares: 1979-1992," *American Economic Review*, 86 (May 1996): 290-93.

Rodgers, William M., III, and Spriggs, William E., "Accounting for the Racial Gap in AFQT Scores: Comment on Nan L. Maxwell, 'The Effect on Black-White Wage Differences of Differences in the Quality and Quantity of Education.' *Industrial and Labor Relations Review*, 55 (April 2002): 533-41.

Spence, A.M.: "Job Market Signaling," *Quarterly Journal of Economics* Vol. 87 (August 1973): 355-374.