NBER WORKING PAPER SERIES

EMPLOYER LEARNING AND THE
SIGNALING VALUE OF EDUCATION

Joseph G. Altonji
Charles R. Pierret

Working Paper 5438

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1996

EMPLOYER LEARNING AND THE
SIGNALING VALUE OF EDUCATION

ABSTRACT

If profit maximizing firms have limited information about the general productivity of new

workers, they may choose to use easily observable characteristics such as years of education to

"statistically discriminate" among workers. The pure credential value of education will depend

on how quickly firms learn. To obtain information on employer learning, we work with a wage

equation that contains both the interaction between experience and a hard-to-observe variable that

is positively related to productivity and the interaction between experience and a variable that

firms can easily observe, such as years of education. The time path of the coefficient on the

unobservable productivity variable provides information about the rate at which employers learn

about worker productivity. Using data from the NLSY we obtain preliminary estimates of the

rate at which employers learn about worker quality and use these, along with some strong

auxiliary assumptions, to explore the empirical relevance of the educational screening hypothesis.

We show that even if employers learn relatively slowly about the productivity of new workers,

the portion of the return to education that could reflect signaling of ability is limited.

Joseph G. Altonji   Charles R. Pierret
Department of Economics   Bureau of Labor Statistics
Northwestern University   2 Massachusetts Avenue, NE
Evanston, IL 60208   Suite 4945
and NBER   Washington, DC 20212

1. Introduction

If firms have limited information about productivity or the personal attributes that determine productivity (such as knowledge, aptitude, and motivation), they will have an incentive to "statistically discriminate" among young workers on the basis of easily observable variables that are correlated with productivity, such as education.[1] By the same token, the signaling value of education is likely to be an important part of the return to education only to the extent that firms lack good information about the productivity of new workers and learn slowly over time. In this paper, we provide some preliminary evidence on how much firms know about new workers and how quickly they learn over time and then use this information to address the issue of how much of the return to education could be due to signaling rather than to the direct effect of education on productivity.

The key difficulty for our investigation of signaling is how to measure how quickly employers learn. Our analysis uses an approach developed more fully in Altonji and Pierret (1995, herafter AP). It is based on a model in which firms have only limited information about the quality of workers in the early stages of their careers. They statistically discriminate among workers on the basis of easily observable variables that are correlated with productivity such as years of education or degree, the quality of the school the person attended, race, and gender. They weigh this information with other information about outside activities, work experience to date, and the information contained in references, the job interview, and perhaps formal testing by the firm. Each period, the firm observes noisy indicators of the worker's performance. Over

---

[1] We use the term "statistical discrimination" to mean that in the absence of full information, firms distinguish between individuals with different characteristics based on statistical regularities. That is, firms form rational expectations given the information they have. Many papers that use the term statistical discrimination analyze race or gender differentials that arise because firms have trouble processing the information they receive about the performance of minority group members. See Aigner and Cain (1977), Lundberg and Startz (1983), Lang (1986), and Coate and Loury (1993) and Oettinger (1996).

time, these make the information observed at the start redundant. Wages become more closely tied to actual productivity and less strongly dependent upon the information that was readily available at the beginning of a worker's career. We draw inferences about how quickly firms learn by observing the rate at which the weight in a wage equation shifts from variables that the firm can easily observe (such as schooling) to a variable that would be hard to observe.

Our investigation of employer learning builds on some previous work, particularly Farber and Gibbons (1994).[2] Farber and Gibbons investigate three implications of employer learning. Imagine a variable s (say schooling) which firms can observe directly and a second variable, z (say AFQT test scores) which firms cannot observe directly. They show first that employer learning does not imply that the coefficient on s in a wage regression will change with experience. This is because future observations, on average, simply validate the relationship between expected productivity and s for new entrants. This point has been made previously as a criticism of attempts to test screening/signaling models of the return to education based on changes in the education coefficient over time. Second, they show that the part of z that is orthogonal to information available to employers at the beginning of a worker's careers will have an increasingly large association with wages as time passes. Third, they note that wage growth will be a Martingale process, at least in the case in which productivity of the worker is constant.

In this paper and AP we make use of a different but related proposition. Specifically, the proposition concerns how controlling for the experience profile of the effect of z on wages alters

---

[2] Other relevant references are Gibbons and Katz (1991) which we discuss below and Parsons (1993). Foster and Rosenzweig (1993) use data on piece rate and time-rate workers to investigate several implications of imperfect information on the part of employers that are different from the one studied here. Their results imply that the incompleteness of employer information is an important issue. Parsons (1986), Weiss (1995) and Carmichael (1989) provide useful discussions of some of the theoretical issues on the link between wages and employer perceptions about productivity. Montgomery (1991) is part of a large literature on labor market networks. Albrecht (1982) conducts a test of screening models of education based on the idea that education will have less impact on the probability a worker will be hired if the worker was referred to the firm by another worker because some of the information contained in education will be transmitted through the referral.

the interaction between experience and s. We show that not only should the coefficient on z rise with time in the labor market, but the coefficient on s should fall. We also show that the time path of the coefficients provides information about employer learning In the current paper we present our basic framework and an initial set of empirical results.[3] We use our results to assess the signalling model of education.

The paper proceeds as follows. In Section 2 we use two simple examples to show that evidence on how much firms know about workers and how quickly they learn is highly relevant to assessing the potential importance of signaling in the return to education. In Section 3 we present our basic theoretical framework and our approach to assessing whether employers learn and the rate at which they learn. In Section 4 we discuss the NLSY data used in the study. In Section 5 we present estimates of the wage model and provide a preliminary assessment of the evidence that employers statistically discriminate among workers and learn over time. In Section 6 we use our estimates of the experience profile of the effect of AFQT scores on wages along with some auxilliary assumptions about employer learning to provide a range of estimates of what the internal rate of return to education would be if education has no direct effect on productivity and if a year of education raises the log of productivity by .05. We show that the "signalling component" of the return to education is probably only a small part of the percentage difference in wages associated with education. Thus, while we find evidence that information is imperfect and firms do statistically discriminate among young workers on the basis of education, our estimates suggest that they learn quickly enough to limit the return to a costly signal such as education.

In Section 7, we consider a potential role for other less costly signals of productivity in the labor market. Specifically, we demonstrate that interpreting our estimates of the time profile of the effect of AFQT on wages as the result of employer learning implies that high ability

---

[3] Our research on the theoretical and econometric issues surrounding employer learning and statistical discrimination and our empirical analysis using National Longitudinal Survey of Youth (NLSY) is ongoing and will be presented in full detail in a revised version of AP.

workers would have a substantial financial incentive to take the AFQT to differentiate themselves from those who are less able in this dimension. The fact that we do not observe this raises an issue for future research. We close the paper with a research agenda in section 8.

2. The Rate at which Employers Learn and the Signaling Value of Education

In this section we use two examples to establish the connection between the rate at which firms learn about worker quality and the quantitative significance of signaling models of the return to education. The examples make the point that if education does not raise skill and employers can observe skill after a short period of time, then the return to going to school would have to be small. The examples set the stage for our analysis in section 6, where we compute the internal rate of return to education conditional on assumptions about the direct effect of education on productivity and information about the rate at which firms learn.

Let Y denote the productivity of a particular worker. Ignore training considerations and assume that Y is time invariant and the same for all employers. Suppose wages W are

$$W = Y^e \, ,$$

where Y^e is the employer's estimate of Y for a particular worker. There are two types of workers, 0 and 1, who have productivity Y_0 and Y_1. Suppose a person has T (T>2) years to divide between work and school, the interest rate is 0, and persons can choose to go to school either 0 years or 1 year. School involves a nonpecuniary cost that is 0 for type Y_1 workers and large for type Y_0 workers. The wage differential associated with the year of school is such that only Y_1 workers attend school. Then in signaling equilibrium $(1+\kappa) = Y_1 / Y_0$ is the ratio of productivity of workers who choose school to workers who do not. Assume that the only information firms have about a new worker is his schooling decision, and they learn nothing for 2 years. Assume that after 2 year, firms know whether productivity is Y_1 or Y_0 independent of the schooling choice. Then the present value of earnings for a person of type Y_1 who chooses 1 year of school is

$$(T-1) \, (1+\kappa)Y_0$$

because the first year is spent in school. The present value for a type Y_1 who does not attend school is

$$2Y_0 + (T-2)(1+\kappa)Y_0$$

because after 2 years in the labor market productivity is known, regardless of whether the person went to school or not.

The above implies that a type Y_1 will choose school if

$$(T-1)(1+\kappa)Y_0 > 2Y_0 + (T-2)(1+\kappa)Y_0.$$

The parameter κ must be greater than or equal to 1 for the type Y_1 person to choose school. That is, for school to make sense for a type Y_1, employers must pay double for someone who they believe is type Y_1. Since a more realistic estimate of the percent increase in earnings associated with a year of school is .08 or .1, it is clear that if employers can observe productivity within a couple of years and the information is sufficiently public to force firms to adjust wages in response, then the signaling value of education cannot be the main reason why education is valued in the market.

A second example may help establish the inverse relationship between the rate at which employers learn about the skill of workers and the potential importance of signaling in the return to education. Let the average productivity of workers with skill level s' be $Y(s')$, and assume that education has no direct effect on labor market productivity. Suppose that the only information firms have about workers when they enter the labor market is their education level s. Suppose that the relationship between skill s' of a worker and the cost of acquiring education and between s and wages is such that in equilibrium a worker with skill level s' chooses s years of schooling. In this case the average productivity of a worker with s years of education is $Y_s = Y(s')$. Let the parameter $(1+\kappa)$ be equal to Y_s/Y_{s-1}, which is also equal to $Y(s')/Y(s'-1)$. Firms acquire information about the productivity of workers by observing their performance in the labor market. Suppose that in the absence of information on schooling, their estimate of the productivity of a type s' worker who has chosen s-1 years of schooling rather than the usual value

of s=s' for this type is

$$Y^e_{s-1,s',t} = Y^{[e^{-\psi t}]}_s Y(s')^{[1-e^{-\psi t}]} = Y_{s-1}(1+\kappa)^{[1-e^{-\psi t}]}$$

The above equation says that in the absence of information on productivity the firm uses the work force mean Y_{s-1} as an estimate of Y for a worker with s-1 years of education. As t goes to infinity, the estimate converges to $Y(s') = Y_{s-1}(1+\kappa)^{(1-[e^{-\psi t}]}$. The larger the value of ψ, the

more rapidly the firm learns. If ψ is .05 and $1+\kappa = 1.1$, then after the first year $Y^e_{s-1,s',t}/Y(s')$ is 1.005. When t is 10 the value is 1.038, and when t is 20 the value is 1.062. Thus, in this case firms learn relatively slowly about worker productivity in the absence of information about schooling. One may ask, "What is the value of the learning parameter ψ such that the present value of a year of education is 0 assuming a borrowing rate ρ and a career of 40 years?" If ρ is .05, the solution for ψ is the value that solves

$$0 = \int_0^T e^{-.05t} Y_{s-1}(1+\kappa)^{(1-[e^{-\psi t}]}dt - \int_1^T e^{-.05t} Y_{s-1}(1+\kappa)^1 dt$$

with $(1+\kappa) = 1.1$ and T set to 40. The solution is .0424, which implies a slow learning rate in that after 15 years $Y^e_{s-1,s',t}/Y(s')$ is still only 1.045. In this case, the internal rate of return to education is only .05, which is just half of the assumed percentage increase in earnings associated with a year of education. A faster rate of learning would imply an even lower rate of return to education.

We would like to be able to solve for ρ given empirical information about the rate at which information about skills is reflected in wages. In the next section, we examine the implications of employer learning for wage equations, and in section 5 we provide estimates of the parameters of the rate at which employers learn. These estimates enable us to estimate the

time path of employer learning up to scale. In Section 6 we will use these estimates with some auxilliary assumptions to perform some calculations of what the internal rate of return to schooling would be if school has no direct effect on productivity.

3. Implications of Statistical Discrimination and Employer Learning for Wage Growth.

In this section we outline a model of employer learning and wages. We use the model to show how coefficients in a wage equation on characteristics that employers can observe directly and on characteristics they cannot observe directly will change as employers become better informed about worker productivity.[4]

The basic setup of the model is similar to Farber and Gibbons (1994). Let y_i be the log of labor market productivity of worker i with t_i years of experience. y_i is determined by

(1) $y_{it} = rs_i + H(t_i) + \alpha_1 q_i + z_i + \eta_i$

where s_i is years of schooling, z_i is a correlate of productivity that is not observed directly by employers but is available to the econometrician, and $H(t_i)$ is the experience profile of productivity. The variable η_i consists of other determinants of productivity and is not directly observed by the employers or the econometrician. The variable z_i might be a test score, the income of an older sibling, or father's education. To simplify the notation but without loss of generality we scale z and η_i so that they have unit coefficients in the productivity equation. In addition to s_i, the employer observes a variety of other things about the worker that are relevant to productivity, which we denote by the vector q_i. For now we assume that the experience profile of productivity does not depend on s_i, z_i, q_i, or η_i, but we will briefly consider the consequences of relaxing this assumption below. In most of the analysis we suppress the i subscript. All variables are expressed as deviations from population means. In this paper we use years of schooling as our example of s, but the basic argument applies to any variable that employers can easily observe.

---

[4] See AP for a more complete development of the model.

For example, in AP we consider race as well.

Firms do not observe y_{it} and so must form an estimate of it. We assume that the conditional expectation of z given s and q, $E(z|s,q)$ and $E(\eta|s,q)$, are linear in q and s, so

(2) $z = E(z|s,q) + v = \gamma_1 q + \gamma_2 s + v$

(3) $\eta = E(\eta|s,q) + e = \alpha_2 s + e$,

where v and e have mean 0 and are uncorrelated with q and s by definition of an expectation.[5] Note that the link between z and η and s may be partially due to a causal effect of s.[6] The sum v+e is uncorrelated with q but in addition we assume that v +e is independent of q and s. The sum v + e is the error in the employer's belief about the log productivity of the worker at the time the worker enters the labor market.

Each period that a worker is in the labor market, firms observe a noisy signal of the productivity of the worker,

(4) $\xi_t = y + \epsilon_t$,

where y and ϵ_t reflects transitory variation in the performance of worker i and the effects of variation in the firm environment that are hard for the firm to control for in evaluating the worker. (We continue to suppress the i subscripts.) The term ϵ_t is independent of the other variables in the model.

Since the employers know q and s, observing ξ_t is equivalent to observing

---

[5] The exclusion of q from the conditional mean of η is innocuous, since we are simply defining η and the coefficient vector α_1 on q in (1) so that the mean of η does not depend on q.

[6] For example, below we use the Armed Forces Qualification Test (AFQT) as z and years of education as s, and Neal and Johnson (1995) present evidence that years of education have a sizeable positive effect on AFQT.

(5) $d_t = v + e + \epsilon_t = \xi_t - E(y|s,q)$.

The vector $D_t = \{d_1, d_2, ..., d_t\}$ summarizes the worker's performance history. Let μ_t be the difference between $v+e$ and $E(v+e|D_t)$. By definition μ_t is uncorrelated with D_t, q and s but in addition we assume μ_t is distributed independently of D_t, q and s.

We also assume that q, s, and D_t are known to all employers, as in Farber and Gibbons. As a result of competition among firms, the worker receives a wage W_t equal to the expected value of productivity Y_t ($Y_t = \exp(y_t)$) times the multiplicative error component $\exp(\varsigma_t)$ that reflects measurement error and firm specific factors outside the model.

(6) $W_t = E(Y_t|s,q,D_t) e^{\varsigma_t}$

Using (1) (2), and (3) and (6) the wage equation is

(7) $W_t = E(Y_t|s,q,D_t) e^{\varsigma_t} = e^{rs + H(t)} e^{(\alpha_1 + \gamma_1)q + (\alpha_2 + \gamma_2)s} e^{E(v+e|D_t)} E(e^{\mu_t}) e^{\varsigma_t}$

Taking logs and collecting terms leads to

(8) $w_t = (r + \gamma_2 + \alpha_2)s + H^*(t_i) + (\gamma_1 + \alpha_1)q + E(v+e |Dt) + \varsigma_t$

where $w_t = \log(W_t)$ and $H^*(t_i) = H(t_i) + \log(E(e^{\mu_t}))$.

Although some authors have purported to test screening models by testing whether the coefficient on s declines with experience (e.g., Layard and Psacharopoulos (1974)), Riley (1979) and others have noted that unless the relationship between schooling and actual productivity

changes, the coefficient on s will not change. This is true regardless of <u>why</u> s is related to productivity. Farber and Gibbons make this point by showing in a more general version of the model above that the expected value of the coefficient of an OLS regression of w_t on s does not depend on t. They estimate an equation of the form

$$(7a) \quad w_t = b_{st}s + H^*(t) + (\alpha_1 + \gamma_1)q + E(v + e|D_t)$$

with q treated as an error component. They find that b_{st} does not depend much on t.[7]

Farber and Gibbons also make a second point, which is that if one adds the component z' of (v + e) that is uncorrelated with the employer's initial information s and q to the wage equation and estimates

$$(8b) \quad w_t = b_{st}s + b_{z't}z' + H^*(t) + (\alpha_1 + \gamma_1)q + E(v + e|D_t),$$

the coefficient on s does not depend on t. They provide evidence from NLSY that b_{st} is relatively constant and $b_{z't}$ is increasing in t.

In AP we establish and make use of a third result, which is closely related to the second. Let the regression equation relating w_t to s, z and H(t) be

$$(9) \quad w_t = b_{st}s + b_{zt}z + H^*(t) + (\alpha_1 + \gamma_1)q + E(v + e|D_t) ,$$

where we have added the component z to (7) rather than followed Farber and Gibbons in adding only the part of z that is orthogonal to the firm's information set and where $(\alpha_1 + \gamma_1)q$ is part of the error term. When the individual starts work (t is 0) this equation is

---

[7] Farber and Gibbons formulate their model in terms of levels of productivity and wages rather than logs.

(9') $w_0 = b_{s0}s + b_{z0}z + H(0) + (\alpha_1 + \gamma_1)q$

Assume that the sample has been drawn so that s and t are uncorrelated. Also assume that z and t are uncorrelated, which is reasonable at least when t is potential experience and s and t are uncorrelated in the sample. Then it is easy to show (see AP) that

(12a) $b_{st} = b_{s0} + \theta_t \Phi s$

(12b) $b_{zt} = b_{z0} + \theta_t \Phi z$

where Φs and Φz are the coefficients of the regression of v+e on s and z and

$$\theta_t = \text{cov}(E(v+e|D_t), z)/\text{cov}(v+e, z) = \text{cov}(E(v+e|D_t),v)/\text{cov}(v+e, v)$$

is a parameter that is specific to experience level that summarizes how much the firm knows at time t.

To determine the behavior of $\theta_t \Phi s$ and $\theta_t \Phi z$ over time, note first that $\Phi s < 0$ and $\Phi z > 0$ if Cov $(v, v+e) > 0$ and cov(s,z) >0. The latter condition is true when s is schooling and z is AFQT. The condition cov(v,v+e) > 0 simply states that the unobserved productivity subcomponent v and composite unobserved productivity term v+e have a positive covariance. This seems plausible to us.

The time paths of b_{st} and b_{zt} are determined by θ_t. This parameter is bounded between 0 and 1. It is 0 in period 0, because in this period employers know nothing about v + e, so $E(v+e|D_0)=0$. The coefficient is 1 if $E(v+e|D_t)$ is v+e, since in this case the employer has learned what v+e is and thus knows productivity y. It is intuitive that θ_t is nondecreasing in t because the additional information that arrives as the worker's career progresses permits a tighter estimate of v+e. In AP we discuss conditions on the ϵ_t process that are sufficient for θ_t to converge to 1 as t

becomes large.

There are two conclusions, which we summarize in Proposition 1 and 2

Proposition 1: _Under the assumptions of the above model, the regression coefficient_ b_{zt} _is_ _nondecreasing in t. The regression coefficient_ b_{st} _is nonincreasing in t._[8]

Proposition 2: _If firms have complete information about the productivity of new workers, then_

$\partial b_{st}/\partial t = \partial b_{zt}/\partial t = 0.$

These results underlie our empirical analysis in AP (which also considers generalizations to vectors of s and z) and the present paper. We are also able to estimate the time profile of θ_t up to scale. Under the strong assumption that employers learn about v and e at the same rate, this enables us to estimate the time profile of employer learning about productivity up to scale.[9] In section 6 below we examine the implications of our estimates for pure signaling models of the return to education.[10]

---

[8] The coefficients on an unfavorable z characteristic, such as criminal involvement or alcohol use, will become more negative to the extent that these reflect permanent traits. Assuming s is negatively correlated with the unfavorable z, b_{st} will rise with t.

[9] We elaborate on this assumption in Section 6. The assumption is natural if firms only see v+e plus noise, as in the model. However, it seems more realistic to assume that firms observe a vector of indicators of productivity. Suppose that the firm observes an indicator d_{1t} of v and an indicator d_{2t} of e. Suppose that v and e are independent and the firm knows this. Then if d_{1t} is a less noisy indicator than d_{2t}, the firm will learn about v faster than e. This means that the time path of θ_t will depend upon the choice of z.

[10] Additional tests are possible if the econometrician has a set of variables B that are observed directly by employers, are negatively related to the direct costs and nonpecuniary costs of school and are unrelated to productivity. Both human capital and screening models imply that schooling s will rise with B. In the screening case, firms cannot directly observe y. Consider the earnings equation

$w_t = B\pi + rs + u_t$

In the screening case, π will be negative. The factor B contaminates the relationship between s and unobserved productivity y. Firms, by taking account of factors that affect schooling choices but not productivity, can form a better prediction of y than if they rely on s alone. Unfortunately, we have not been able to identify a set of variables that have the properties of B, and so we have not pursued this line of research. There is a large theoretical literature on screening/signaling models, but the empirical literature is far from conclusive. (See Weiss (1995) for a useful discussion of some of the evidence.) In future work it might be interesting to see if the "diploma effect" declines with t while the coefficients on hard to observe productivity characteristics that correlate with getting a diploma rise. See Frazis (1993) for a recent analysis of whether there is a diploma effect.

The model also implies a third result, which we state in proposition 3.

Proposition 3: _Under the assumptions of the above model, $\partial b_{st}/\partial t = -\partial b_{zt}/\partial t \, Cov(z,s)/Var(s)$._

Since $Cov(z,s)/Var(s)$ is simply the regression coefficient of z on s and can be estimated, the coefficient restriction in Proposition 3 may provide leverage in differentiating between the learning/statistical discrimination model and alternative explanations for the behavior of b_{st} and b_{zt}.

3.1 Modifications to the Model:

Proposition 1, 2, and 3 are not robust to allowing the effect of z and/or s in the productivity equation (1) to depend on t. For example, if s and/or z are complementary with learning by doing or enhance the productivity of training investments, then the productivity equation (net of training costs) might take the form

$$(13) \; y_t = r \, s + r_1 s \, t + r_2 \, z \, t + H(t) + (\alpha_1 + \gamma_1)q + z + \eta, \quad r_1 > 0, \; r_2 > 0 \,.$$

The interactions between t and s and z in the productivity equation would influence the interactions between t and s and z in the wage equation, although the precise effects depend on the rate at which the firm learns about z. Consequently, our estimates of the time path b_{zt} will be biased as estimates (up to the scale parameter Φ_z) of the rate at which employers learn about v + e.

In AP we are investigating how this modification affects our analysis of the hypothesis that employers learn about productivity over the career. Most discussions of human capital and most of the empirical evidence on employer provided training suggest that education makes workers more trainable and that educated workers receive more training. (See, for example, Altonji and Spletzer (1992), Bartel and Sicherman (1992), Lynch (1992) and Mincer (1993)). In AP we find that both highest grade completed (our s variable) and AFQT (the z variable in this paper) have strong positive effects in a probit model of whether a worker receives company

training during the year. In this case r_1 will be greater than 0. Below and in AP we find that the education slope of wages has a strong negative relationship with t, which is only consistent with a training interpretation if education reduces learning by doing, the productivity of training investments, and/or the quantity of training investments. The presence of r_2 z t in the productivity equation seems unlikely to lead a negative bias in $\partial b_{st}/\partial t$ when z_t is excluded from the model.[11] If both r_1 and r_2 are positive the introduction of z_t to the wage model that contains s_t could lead the coefficient on s_t to fall, but will not lead it to become negative unless r_1 is negative. This seems unlikely. However, our use of b_{zt} in section (5) to draw inferences about the rate at which employers learn implicitly assumes that the direct effect of z on skill accumulation is 0.

The issue of whether information about productivity is public or private deserves discussion. The equation that relates wages to expected productivity rests on the assumption that the information available to the employer leaks out to other employers. However, little is known about how much of an employer's, a supervisor's, or a co-workers' knowledge of the general productivity of a worker becomes known to other prospective employers. Theoretical papers by Greenwald (1986), Waldman (1984), Lazear (1986), and Gibbons and Katz (1991)) discuss whether information about productivity will be reflected in promotion paths and wage increases within firms. They also discuss the strategies firms might use to try to hide information about good workers. In AP we investigate some implications of these models for turnover and the relationship between quits, layoffs, and wage gains but our results to date are far from conclusive. In section 6 we briefly consider how the possibility that some of the information obtained by employers is private and not reflected in wages will effect our analysis of signaling.

4. Data

---

[11] Farber and Gibbons note that a training explanation of their finding that $\partial b_{z*t}/\partial t > 0$ is hard to reconcile with their finding that $\partial b_{st}/\partial t$ is close to 0 or negative.

The empirical analysis is based on the 1992 release of the NLSY. The NLSY is a panel study of men and women who were aged 14-22 in 1978. Sample members were surveyed annually since 1979. We restrict the analysis to men who have completed 8 or more years of education and have valid data on all variables used in the analysis. Hispanics are excluded from the analysis. We exclude labor market observations prior to the first time that a person leaves school and accumulate experience from that point. Each panel member contributes at most one observation for a particular year. If he is working at two or more jobs, we consider only the wage for the CPS job.

Actual experience is the number of weeks in which the person worked more than 30 hours divided by 50. Potential experience is defined as age minus years of schooling minus 6. The AFQT score is normalized to have a standard deviation of 1 in the population, but has a sightly larger standard deviation in our sample.[12] The means, standard deviations, minimum and maximums of the variables used in analysis are provided in Table 1, along with the variable definitions. The mean of actual experience is 4.9. The mean of potential experience is 7.3, and the mean of education is 12.7. All statistics in the paper are unweighted. Blacks are oversampled in the NLSY and contribute 29.1 percent of our observations.

5. Estimates of the Wage Equation

In Table 2-4 we report estimates of various specifications of the wage model. In table 2 we use potential experience as the experience measure and use OLS to estimate the model. The equations also control for a cubic in experience, a quadratic time trend, and residence in an urban

---

[12] The age of the sample members at the time the AFQT was administered varies somewhat in the NLSY sample. This induces some variation in schooling levels at the time the AFQT is taken. The standardization procedure uses the entire NLSY sample weighted so as to be nationally representative. Following Neal and Johnson (1995), each birth cohort year is standardized to a mean of 0 and a standard deviation of 1. (The lowest standard deviation is for the 1963 cohort (35.54) and the highest is for the 1960 cohort (36.82), so standard deviation adjustment makes little difference.) The fact that the mean of this variable is not 0 in our sample is due to the over sample of disadvantaged youths.

area. These variables are not reported in the tables.

In column 3 we present an equation that includes s, Black, and s*t. This corresponds to (7a) with b_{st} restricted to $b_{st} = b_{s0} + b_{s1}*t$. The coefficient on s*t is -.00075 (.00040), suggesting that the effect of education on wages declines slightly with experience. In column 4 we add AFQT. As has been well documented, AFQT has a powerful association with earnings even after controlling for education. A shift in AFQT from 1 standard deviation below the mean to 1 standard deviation above is associated with an increase in the log wage of .164. The coefficient on education declines to .0808, but note that b_{s1} falls only slightly, to -.00102.

In column 5 we add linear interactions between t and our z variable, AFQT. The resulting equation corresponds to (9) with the restriction that $b_{st} = b_{s0} + b_{s1}*t$ and $b_{zt} = b_{z0} + b_{z1}*t$. The estimates imply that the effect of AFQT on the wage increases greatly with experience t. b_{AFQT1}, which is the coefficient on AFQT*t, is .0090 (.0008). b_{AFQTt}, which is $\partial wt/\partial AFQT$, rises from only .0164 when experience is 0 to .1067 when experience is 10. The results imply that when experience is 10 and education is held constant, persons with a value of AFQT that is 1 standard deviation above the mean have a log wage that is .211 larger than persons 1 standard deviation below the mean, while the difference is only .033 when experience is 0.

Our results for AFQT parallel Farber and Gibbon's results in which they use the components of AFQT and an indicator for whether the family had a library card when the person is 14 that are orthogonal to the wage on the first job and education. The key new result is that the coefficient on s*t declines sharply (to -.00348 (.00046)) when AFQT*t is added. The implied effect of an extra year of education for a person with 10 years of experience is only .0586. Strikingly, the coefficient on s rises to .0987 which is almost exactly what we obtain when we exclude all terms involving AFQT from the model (columns 1 and 3).

These results are consistent with the hypothesis that employers have limited information about the productivity of labor force entrants. Early wages are based on expected productivity conditional on easily observable variables such as education. As experience accumulates, wages

become more strongly related to variables that are likely to be correlated with productivity but hard for the employer to observe directly. While one might argue that the positive coefficient on AFQT*t is due to an association between AFQT and training intensity, it is hard to reconcile this view with the negative coefficient on s*t. While measurement error in schooling may partially explain the decline in s between columns 1 and 2, it does not provide a simple explanation for the behavior of the interaction terms with experience in columns 3 and 4.

In Table 3 we present OLS results using actual experience in place of potential experience as the experience measure t. The main difference between this table and table 2 is that the return to education is lower and the s*t interaction is positive and fairly large in the equations that exclude AFQT*t. However, the coefficient on s*t declines from .0021 in column 5 to -.0004 when the interaction terms are added in column 6 of Table 3. This decline is similar to the decline that we obtain in Table 2.

The results in Table 3 are difficult to interpret, because the intensity of work experience may be conveying information to employers about worker quality. It is an outcome measure itself. Conditioning on actual work experience raises some of the issues that would arise if we conditioned on wages in t-1. On the other hand, the results based on potential experience are likely to be biased by the fact that potential experience mismeasures actual. For this reason, in Table 4 we report the results of re-estimating the models by instrumental variables (IV), treating all terms involving actual experience as endogenous with corresponding terms involving potential experience as the instruments. The results in columns 4 and 5 of Table 4 are basically consistent with those in Table 2. The coefficient on AFQT is .0155 (.0062) and the coefficient on AFQT*t is .0126 (.0011). These estimates imply that conditional on years of schooling, AFQT has only a small effect on initial wages, but when t is 10, a shift in AFQT from 1 standard deviation below to one standard deviation above the mean raises the log wage by .284. The coefficient on s*t declines from -.0018 when the interactions are excluded in column 4 to -.0055 in column 5.

It is interesting to ask whether the experience profiles of the education and AFQT

coefficients satisfy the restrictions in Proposition 3. Consider the specification with the linear interactions in column 5 of Table 2. Proposition 3 says that product of the value of -cov(s,z)/var(s) and the coefficient on the interaction between t and AFQT, which is equal to -.00253, should be close to the coefficient -.00348 on the interaction between s and t. The estimates differ but are in same neighborhood. The corresponding estimates for Table 4 are -.00357 and -.00552.[13]

Nonlinear Specifications of the Experience Profile of the Effects of AFQT and s on Wages

The above analysis assumes that the effects of AFQT and s are a linear function of experience. In this section we presents results for spline functions and fourth order polynomial specifications, in part to examine the sensitivity of estimates of the time path of the effects of AFQT to functional form. Specifically, we estimate models of the form

$$w_t = f(z,t;b_z) + h(s,t;b_s) + H(t) + \epsilon_{it}$$

where b_z and b_s are now vectors of parameters. Column 2 of table 5 is based on a model in which $f(z,t;b_z)$ is a linear spline with break points at t=2, t=4, t=7 and t=10. In that column the function $h(s,t;b_s)$ is simply the $b_{s0}s + b_{s1}s*t$. All of the models in the tables contain the other control variables discussed above.

Rows d to h report $\partial w_t/\partial AFQT$ at various experience levels. The coefficient on AFQT in row c is the derivative when t=0. The table shows that the derivative increases steadily from .0196 when t is 1 to .1193 when t is 12. For purposes of comparison, in column 1 we also report the corresponding derivatives when $f(z,t;b_z) = b_{z0}z + b_{z1} z t$ as in Table 2-4. The pattern is similar, suggesting that to a first approximation, the effect of AFQT rises linearly with

---

[13]   [13] The numbers are calculated as -.00253= .00896*-.2828 and -.00357=.01264*-.2828.

experience. As noted earlier, employer learning implies that $\partial w_t/\partial AFQT$ is nondecreasing in t (

i.e., $\partial^2 w_t/\partial AFQT, \partial t \geq 0$), with a strict inequality likely if some new information arrives each

period on y. If the noise in observations of y_t are iid, then the rate of increase is declining, as

shown in the expression θ_t above. In this case, $\partial^3 w_t/\partial AFQT, \partial^2 t < 0$. The rate of increase must

decline eventually because the amount of additional information in additional observations of

labor market performance is declining. (θ_t is bounded at 1.) However, it is possible that the first

two or three observations on a worker are particularly noisy because of factors that we have left

out of the model. For example job specific or occupation specific match quality may be more

variable for new workers than more experienced ones.

In rows i to m of the table we report $\partial^2 w_t/\partial AFQT, \partial t$ for various experience levels. We

only go out to t=12 because sample information becomes thin at higher values. In the linear case

in column 1, the values of $\partial^2 w_t/\partial AFQT, \partial t$ is constant. In column 2, the value rises from -.0003

(.0154) when t is 1 to .0135 when t is 3, increases slightly to .0154 when t is 5.5 and then declines

to .0065 when t is 8.5 and to .0037 when t is 12. These results are reasonably consistent with a

decline in the amount of new information with experience after the first few years in the labor

market.[14]

Column 5 reports results when $f(z,t;b_z)$ and $h(s,t;b_s)$ are both fourth order polynomials in

t, with coefficients so that $\partial^2 w_t/\partial AFQT\ \partial t = 0$ and $\partial^2 w_t/\partial s\ \partial t = 0$ when t is 25. This is the

specification that underlies our calculations of the internal rate of return to education in the next

section. The restrictions captures that idea that firms learn whatever they are going to learn about

the information about productivity that is contained in AFQT by the time t is 25. It is important

to point out that the 99th percentile value for potential experience is 16 and the 95th percentile

---

[14] In column 3 we estimate a model with a spline for the education/experience interaction
$h(s,t;b_{st})$. The results in rows a and n-r show an increase in $\partial w_t/\partial s$ in the initial years followed by a
decline. Given the standard errors we would not want to make too much of the initial increase. For
the same specification the effect of AFQT declines slightly during the first 2 years in the labor
market before increasing. The initial decline is not consistent with the pattern implied by theory.
However, given the standard errors on AFQT terms we are not sure how much to make of this.

value is 14, so there is little information in the sample beyond t=15 or so. The results based on the restricted quartic polynomial are similar to those based on the spline functions. The effect of AFQT increases monotonically. The rate of increase rises at first but then declines between t=8.5 and t=12. In column 7 we report results for unrestricted quartic polynomials. These estimates are very similar to those for the restricted polynomials but the standard errors are bigger.[15]

6. The Rate at which Employers Learn and the Signaling Value of Education

We are now ready to examine the implications of employer learning for the quantitative significance of signaling models of the return to education. The internal rate of return ρ to an additional year of schooling is implicitly defined as the solution to the equation

$$(15) \quad 0 = pv(\rho, \psi_1, \dots, \psi_T, r, \kappa W_{s-1}, T) = \int_0^T e^{-\rho t} W_{s-1} e^{r+\kappa \psi_t} dt - \int_1^T e^{-\rho t} W_{s-1} e^{r+\kappa} dt \ .$$

In the above equation W_{s-1} is the earnings of the average worker with s-1 years of education, T is number of years until retirement, which we set to 40, r is the direct effect of education on productivity and we redefine κ so that $\kappa \equiv \alpha_2 + \gamma_2$ is the relationship between s and components of productivity v+e that the firms cannot observe when new workers enter the labor market. The total difference in productivity associated with an extra year of school is $r + \kappa + \Upsilon$, where Υ is the slope of the relationship between s and the productivity component $\alpha_1 q$ that is observable to the firm when the individual enters the labor market. In our calculations we assume that $r + \kappa$ is .10,

---

[15] In Table 6 we replace potential experience with actual experience, and treat actual experience as endogenous. The 99th percentile value for this variable is only 13.33, so there not much sample information on t beyond this point. Focussing on the model with unrestricted quartic specifications for both f( ) and h( ) (column 7) we find that the effect of AFQT increases monotonically with experience. The rate of increase rises at first from .01042 when t = 1 to .0144 when t=5.5, but declines to .0059 when t = 12. However, the standard errors on these derivatives are quite large. These results are loosely consistent with the proposition that the rate at which new information about initial productivity arrives declines with experience, but the estimates are not sufficiently precise to say much about this. As the NLSY sample ages, it will be interesting to revisit the issue. The return to education declines slightly in the first year or two but more rapidly after that.

which is a bit above the estimate of the relationship between education and the log wage in our sample. (See Table 2, column 1) This estimate will overstate $r + \kappa$ by Υ. The parameter ψ_t is the coefficient relating the firm's expectation of v+e given the information set D_t to v+e, with

$$\psi_t = Cov(E(v+e|Dt),v+e)/Var(v+e)$$

If the firm is fully informed by period t, ψ_t is 1.

We wish to solve for ρ given empirical information about the rate at which information about skills is reflected in wages. The models underlying Tables 5 and 6 allow us to compute the time profile of the effect of AFQT out to about 15 years. Assume, perhaps heroically given that AFQT is positively related to receipt of company training, that none of the increase in the effect of AFQT reflects training.[16] Assume firms learn about all components of the productivity of the worker at the same rate. This assumption rules out the possibility that the firm might learn about competence on the job more rapidly than it learns about absenteeism and the propensity to quit. Specifically, we need to assume that $\theta_t = \psi_t$, or, using the definitions of these parameters,

$$Cov(E(v+e|D_t),v)/Cov(v+e,v) = Cov(E(v+e|Dt),v+e)/Var(v+e) .$$

This assumption is consistent with the assumption in our model that firms receive information about v + e rather than about separate components of productivity. If it is correct, then $b_{zt}-b_{z0}$, the difference between the derivative of the wage with respect to z at time t and time 0, will be an

---

[16] We believe that AFQT is positively related to receipt of general training and firm specific training, but it it worth pointing out that if the NLSY training measure captures training that is financed by the firm (perhaps because it is highly firm specific), then the association between training and AFQT might not bias our analysis. However, the introduction of match specific capital into the analysis (either through heterogeneity in match quality or firm specific training) complicates our simple competitive model of wage determination because it leads to a gap in the value of the worker to the current employer and outside employers. The effect this would have on our analysis is unclear.

estimate of $\psi_t\Phi_z$. We pin down the scale parameter Φ_z by making specific assumptions about how much the firm knows when t is 15. For example, suppose that by the time t is 15 the information available to employers about workers explains 75% of the variation in v+e, which is the component of productivity that the firms do not know when the worker enters the market. Then this would imply that $\psi_{15} = \theta_{15} = .75$. Our estimate of ψ_t is $(b_{zt}-b_{z0}/b_{z15}- b_{z0})\psi_{15}$ when t is less than 15. We will report results setting this weight ψ_{15} to various values. Some discipline on the appropriate value for ψ_{15} is provided by the fact that $(b_{z15}-b_{z0})/\psi_{15}$ is an estimate of the scale parameter Φ_z, which is the regression coefficient relating v+e to z. For example, assuming that ψ_{15} is only .5 implies that controlling for s, persons with 15 years experience who are 1 standard deviation above the mean in the AFQT are 29.2 percent more productive but are only paid 14.6 percent more. Finally, we assume that between t=15 and T=40 the weight on AFQT rises linearly to the level ψ_{40}, so that $\psi_t=\psi_{15} + (t-15)(\psi_{40}-\psi_{15})/(40-15)$ when t is between 15 and 40. We vary the assumption about ψ_{40}.

In table 7 we report values of the internal rate of return ρ to education for various values of ψ_{15} and ψ_{40}. The estimates of $b_{zt} - b_{z0}$ are based on the wage equation underlying Table 5, column 5, which uses interactions between AFDC and the first four powers of potential experience. We set r to 0 in the left panel, which corresponds to the pure signalling case. If r is 0 and employers are fully informed after 15 years, with $\psi_{15}=1$, then ρ is negative. If ψ_{15} is .75 the internal rate of return ρ is only .032 even if firms never learn anything else about y after t=15 ($\psi_{40}=.75$). The value of ρ is .017 if they are fully informed by the time t=40. Even if the weight on κ is as small as .5 when t=15 and firms do not learn anything else, ρ is a maximum of .062. When we use values of ψ_t based on the IV estimates using actual experience as the experience measure (Table 6, column 5), we obtain even lower estimates of the internal rate of return that could be due to signalling.

In the right panel we consider a human capital/signalling model by setting r to .05. The results for the mixed model show that if the "human capital" component of the return to education

is .05 then the total return is about .08 if the weight on v+e is .5 when experience is 15 and firms never learn anything else. The signalling component of the return is only .03, which is only 60% of what one obtains by subtracting .05 from the total education differential $(r + \kappa)$ of .1

We conclude from this that even a relatively slow rate of learning would eliminate much of the economic return to using education to signal productivity. We have emphasized that our use the time path of b_{zt} to infer the path of ψ_t requires a strong assumption about the flow of information to firms which implies that they are learn about all components of productivity at the same rate. If this is not the case, then the time profile of learning might be sensitive to our use of AFQT as the z variables. We also stress that our estimates of b_{zt} might be affected by other factors. This would affect our rate of return calculations.

In addition, there are two other important caveats that deserve discussion. First, it is likely that the type of job an employee is in influences the type of information the employer receives. For example, employers probably do not learn much about the managerial ability, technical training, or communications skills from observations on the performance of janitors. Education may influence the initial assignment, and information flows may be limited from some jobs. While our estimates of the profiles of the AFQT and s variables are not very sensitive to adding controls for 1 digit occupation, more analysis is needed. (Results not reported.) Second, to the extent that information is private, workers may need to resort to education because it is a public signal of ability. While the two caveats may limit employer learning, the empirical results suggest that learning does take place.

7. The Potential for Testing Services to Certify Skill

While education may be too expensive to serve as a means for able workers to certify themselves to employers, perhaps other mechanisms could perform this function, at least for some determinants of productivity. Here we point out that interpreting our estimates of the time profile of the effect of AFQT on wages as the result of employer learning implies that high ability

workers would have a substantial financial incentive to take the AFQT to differentiate themselves from those who are less able in this dimension.

Suppose that a third party were to administer the AFQT and certify the results to outside employers, in much the same way that the Educational Testing Service administers the SAT exams. Using our estimates of the learning profile and the same range of assumptions about the fraction of information contained in AFQT that is known to firms by the time experience is 15 that we used in Table 7, we have computed how much a person who believes that he is 1 standard deviation above the mean for the AFQT would pay to take the test at the time he enters the workforce.[17] The OLS estimates using potential experience (Table 5, column 5) imply that if firms become fully informed about productivity by the time experience is 15 and the interest rate is .1, then the person would be willing to pay .559 of the first year's salary for the test.[18] The corresponding value when we use the IV estimates in Table 6, column 5 is .330.

These calculations raise the issue of why such a testing service has not emerged if information is initially imperfect. One answer is that firms are not aware that the AFQT captures characteristics that have a strong association with productivity. It is only recently, with the availability of the NLSY, that labor economists have become aware of this. Another is that it would be difficult for a testing firm to become established at a national level. A third is that, given race differences in distribution of AFQT scores, firms who make use of AFQT information in hiring for a specific job would have the burden of establishing that they are relevant to productivity in that job or run the risk of violating discrimination laws. This would be true even if individuals provided firms with the test results. However, we do not find these answers to be

---

[17] If a worker did not know his ability, he could take a practice test on his own. Presumably, this would not raise the total cost of the test very much.

[18] Here we are assuming that only 1 worker takes the test and ignoring the fact that the composition of the pool of workers who choose to take the test in equilibrium would influence return for a particular type of worker.

fully satisfactory.[19] Analyses based on variables such as the wage rates of siblings or father's education may be less vulnerable to this objection. We report qualitatively similar findings for these variables in AP.

8. Conclusion

This paper presents some preliminary evidence on the rate at which employers learn about productivity and uses that evidence to guide an exploration of the extent to which the return to education could be due to signaling. Our basic idea is that if employers learn quickly, then the signaling component of the return to education must be small. To get evidence on how quickly employers learn, we work with a model that is based on the premise that firms use the information they have available to them to form judgements or beliefs about the productivity of the workers and then revise these beliefs as additional information becomes available. Building upon some previous work, particularly Farber and Gibbons (1994), we show that as firms acquire more information about a worker, pay may become more dependent on productivity and less dependent on easily observable characteristics or credentials. Our result that the effect of AFQT on the wage rises with experience and the effect of education declines is consistent with the model.

Taken at face value, our estimates identify the rate at which employer knowledge of worker quality rises with experience up to a scale parameter. We use these estimates along with some strong auxiliary assumptions to provide a range of estimates of what the internal rate of return to education would be if education has no direct effect on productivity. Our calculations suggest that the "signaling component" of the return to education is probably only a small part of the percentage difference in wages associated with education. Thus, while we find evidence that

---

[19] Note also that in the absence of an institution such as the Educational Testing Service, a firm might provide the test. Some firms perform their own testing.. However, if the results were available to the employees or other firms know that a particular firm tests its employees, then the firm would not be able to capture the full return to testing.

information is imperfect and firms do statistically discriminate among young workers on the basis of education, our estimates suggest that they learn quickly enough to limit the return to a costly signal such as education. This does not mean, of course, that none of the return of education is a return to signaling.

We wish to stress that we are still in a relatively early stage in our analysis of employer learning and statistical discrimination on the basis of education and other characteristics, and so the estimates underlying our analysis of the signaling return to education are somewhat preliminary. Among the issues we are pursuing is the possibility that differences among workers in training rather than statistical discrimination with learning explains our results. We are also extending the analysis to other variables that are in the NLSY but would be hard for employers to observe, such as characteristics of the father, mother and siblings. In preliminary work we have experimented with both the wages of siblings with 5 to 8 years of experience and with father's education. It will be interesting to repeat our calculations of the signaling value of education with the learning profiles suggested by other measures. Finally, we are investigating the assumption that information about labor market performance is public.

References

Aigner, D. and Cain, G., "Statistical Theories of Discrimination in the Labor Market," Industrial and Labor Relations Review, 30 (January 1977), 175-87.

Albrecht, J., "A Procedure for Testing the Signaling Hypothesis,." Journal of Public Economics, (1981), 123-32.

Altonji, J. G. and Pierret, C. R., "Employer Learning and Statistical Discrimination," (September 1995), revision in progress.

Altonji, J. G. and Spletzer, J, "Worker Characteristics, Job Characteristics, and the Receipt of On-the-Job Training",Industrial and Labor Relations Review, 45(1) (October 1991): 58-79.

Bartel, A. P. and Sicherman, N., "Technological Change and On-The-Job Training of Young Workers," unpublished paper, Columbia University (July 1993).

Borjas, G., "Ethnic Capital and Intergenerational Mobility," Quarterly Journal of Economics, 117 (1) (February 1992), 123-150.

Carmichael, H. L.. "Self-Enforcing Contracts, Shirking, and Life Cycle Incentives," The Journal of Economic Perspectives, 3 (Fall, 1989), 65-84.

Coate, S. and Loury, G. C., (1993) "Will Affirmative Action Policies Eliminate Negative Stereotypes?" American Economic Review, 83 (5) (December 1993): 1220-1240

Devine, T.J. and Kiefer, N.M., Empirical Labor Economics, New York: Oxford University Press, 1991.

Farber, H., and Gibbons, R. , "Learning and Wage Dynamics," Industrial Relations Section Working Paper, Princeton University (1994).

Foster, A. D. and Rosenzweig, M. R., "Information, Learning, and Wage Rates in Low Income Rural Areas," Journal of Human Resources (Fall 1993).

Frazis, H., "Selection Bias and the Degree Effect," Journal of Human Resources, 28 (3): (Summer 1993), 538-554.

Gibbons, R., and Katz, L., "Layoffs and Lemons," Journal of Labor Economics, 9 (January, 1991), 351-80.

Greenwald, B., "Adverse Selection in the Labor Market," Review of Economic Studies, 53 (1986), 325-47.

Holzer, H., "Search Methods Used by Unemployed Youth," Journal of Labor Economics, 6 (1988), 1-20.

Jovanovic, B.,"Job Matching and the Theory of Turnover," Journal of Political Economy, 87 (1979), 972-90.

Lang K.,"A Language Theory of Discrimination," Quarterly Journal of Economics, 101 (May 1986): 363-82.

Lazear, E., "Raids and Offer Matching" Research in Labor Economics 8(part A):(1986) 141-165.

Layard, R., and G. Psacharopoulos, "The Screening Hypothesis and the Returns to Education," Journal of Political Economy 82 (September 1974), 985-998.

Lundberg, S., and Startz, R., "Private Discrimination and Social Intervention in Competitive Labor Markets," American Economic Review, 73 (June 1983), 340-347.

Lynch, L., "Private Sector Training and the Earnings of Young Workers," American Economic Review, 82 (March 1992), 299-312.

Medoff, J. and Abraham, K., "Experience, Performance and Earnings," Quarterly Journal of Economics, (December 1980), 703-736.

Mincer, J., Studies in Human Capital, E. Elgar, 1993.

Montgomery, J., "Social Networks and Labor Market Outcomes: Toward and Economic Analysis," American Economic Review, 81 (1991), 1408-18.

Neal, D. A., and Johnson, W. R.,"The Role of Pre-Market Factors in Black-White Wage Differences," NBER working paper no. 5124, (May 1995).

Oettinger, G. S., "Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap," forthcoming Journal of Labor Economics, (September 1993).

Parsons, D. O., "The Employment Relationship: Job Attachment, Worker Effort, and the Nature of Contracts," in Ashenfelter, O. and Layard, R., eds. Handbook of Labor Economics, Amsterdam: North-Holland Publishing Co., 1986,

Parsons, D. O., "Reputational Bonding of Job Performance: The Wage Consequences of Being Fired," unpublished paper. (July 1993).

Rees, A.,"Information Networks in Labor Markets," American Economic Review, 56 (1966) 559-66.

Riley, J. G., "Testing the Educational Screening Hypothesis," Journal of Political Economy, 87 (October 1979), S227-52.

Waldman, M., "Job Assignment, Signaling, and Efficiency," <u>Rand Journal of Economics</u>. 15 (1984), 255-67.

Weiss, A., "Human Capital vs. Signaling Explanations of Wages," <u>Journal of Economic Perspectives</u>, 9 (Fall 1995), 133-54.

Table 1: Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Real Hourly Wage	8.366	4.763	2.01	96.46
Log of Real Hourly Wage (w)	2.005	0.475	0.70	4.57
Potential Experience (t)	7.325	3.657	0.00	21.00
Actual Experience (t)	4.914	3.413	0.00	18.26
Education (s)	12.715	2.140	8.00	18.00
Black dummy (Black)	0.291	0.454	0	1
Standardized AFQT Score (AFQT)	-0.138	1.042	-2.780	1.922
Dummy for Urban Dweller	0.780	0.414	0	1
Year	86.643	3.548	79	92

Sample size = 26,651 observations.

Table 2: The Effects of Standardized AFQT and Schooling on Wages

Dependent Variable: Log Wage.   Experience Measure: Potential Experience.

OLS estimates (standard errors)

	Model:	(1)	(2)	(3)	(4)	(5)
(a)	Education (s)	0.0949 (0.0016)	0.0735 (0.0018)	0.1002 (0.0033)	0.0808 (0.0033)	0.0987 (0.0037)
(b)	Black	-0.1812 (0.0057)	-0.1012 (0.0066)	-0.1811 (0.0057)	-0.1008 (0.0066)	-0.0993 (0.0066)
(c)	Standardized AFQT (z)		0.0819 (0.0035)		0.0822 (0.0035)	0.0164 (0.0069)
(d)	Education * Experience (s*t)			-0.00075 (0.00040)	-0.00102 (0.00040)	-0.00348 (0.00046)
(e)	AFQT * Experience (z*t)					0.00896 (0.00081)
	R-squared	.2315	.2471	.2316	.2473	.2507

Note: All equations control for a quadratic time trend, a cubic experience profile, and urban residence. The sample size is 26,651.

Table 3: The Effects of Standardized AFQT and Schooling on Wages

Dependent Variable: Log Wage.   Experience Measure: Actual Experience.

OLS estimates (standard errors)

	Model:	(1)	(2)	(3)	(4)	(5)
(a)	Education (s)	0.0810 (0.0012)	0.0609 (0.0015)	0.0713 (0.0022)	0.0510 (0.0023)	0.0629 (0.0027)
(b)	Black	-0.1396 (0.0056)	-0.0658 (0.0065)	-0.1399 (0.0056)	-0.0661 (0.0065)	-0.0661 (0.0065)
(c)	Standardized AFQT (z)		0.0762 (0.0034)		0.0762 (0.0034)	0.0368 (0.0053)
(d)	Education * Experience (s*t)			0.00206 (0.00038)	0.00208 (0.00038)	-0.00035 (0.00045)
(e)	AFQT * Experience (z*t)					0.00811 (0.00085)
	R-squared	.2695	.2829	.2703	.2838	.2862

Note: All equations control for a quadratic time trend, a cubic experience profile, and urban residence. The sample size is 26,651.

Table 4: IV Estimates of the Effects of Standardized AFQT and Schooling on Wages

Dependent Variable: Log Wage.   Experience Measure: Actual Experience with Potential Experience as Instruments .
IV estimates (standard errors)

	Model:	(1)	(2)	(3)	(4)	(5)
(a)	Education (s)	0.0817 (0.0013)	0.0616 (0.0016)	0.0889 (0.0031)	0.0702 (0.0032)	0.0884 (0.0035)
(b)	Black	-0.1390 (0.0058)	-0.0636 (0.0066)	-0.1389 (0.0058)	-0.0634 (0.0066)	-0.0632 (0.0066)
(c)	Standardized AFQT (z)		0.0770 (0.0034)		0.0771 (0.0035)	0.0155 (0.0062)
(d)	Education * Experience (s*t)			-0.00153 (0.00056)	-0.00183 (0.00056)	-0.00552 (0.00065)
(e)	AFQT * Experience (z*t)					0.01264 (0.00106)
	R-squared	.2376	.2530	.2364	.2515	.2554

Note: All equations control for a quadratic time trend, a cubic experience profile, and urban residence. The instrumental variables are the corresponding terms involving potential experience and the other variables in the model. The sample size is 26,651.

Table 5: The Effects of Standardized AFQT and Schooling on Wages

Nonlinear Interactions of AFQT and Education with Experience
Dependent Variable: Log Wage.   Experience Measure: Potential Experience.

OLS estimates (standard errors)

		Linear Interaction	Splines		Quartic Polynomial Restricted Derivative		Quartic Polynomial	
		AFQT & Ed	AFQT only	AFQT & Ed	AFQT only	AFQT & Ed	AFQT only	AFQT & Ed
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(a)	Education (s)	0.0987 (0.0037)	0.0992 (0.0037)	0.0815 (0.0066)	0.0993 (0.0037)	0.0798 (0.0074)	0.0993 (0.0037)	0.0791 (0.0083)
(b)	Black	-0.0993 (0.0066)	-0.0993 (0.0066)	-0.0992 (0.0066)	-0.0994 (0.0066)	-0.0992 (0.0066)	-0.0994 (0.0066)	-0.0993 (0.0066)
(c)	Standardized AFQT (z)	0.0164 (0.0069)	0.0199 (0.0277)	0.0480 (0.0306)	-0.0018 (0.0175)	0.0226 (0.0206)	-0.0030 (0.0207)	0.0211 (0.0223)

Derivative of Wages With Respect to AFQT, by experience level (yrs.)

(d)	1	0.0253 (0.0070)	0.0196 (0.0317)	0.0361 (0.0349)	0.0098 (0.0198)	0.0246 (0.0233)	0.0094 (0.0244)	0.0242 (0.0260)
(e)	3	0.0433 (0.0071)	0.0327 (0.0357)	0.0345 (0.0393)	0.0358 (0.0239)	0.0378 (0.0284)	0.0359 (0.0311)	0.0380 (0.0329)
(f)	5.5	0.0657 (0.0072)	0.0693 (0.0364)	0.0660 (0.0400)	0.0685 (0.0286)	0.0641 (0.0342)	0.0686 (0.0389)	0.0640 (0.0408)
(g)	8.5	0.0925 (0.0073)	0.1021 (0.0368)	0.0992 (0.0404)	0.1013 (0.0341)	0.0979 (0.0408)	0.1012 (0.0478)	0.0977 (0.0499)
(h)	12	0.1239 (0.0075)	0.1193 (0.0372)	0.1220 (0.0407)	0.1235 (0.0401)	0.1266 (0.0482)	0.1236 (0.0580)	0.1271 (0.0603)

Slope of AFQT Effect, by experience level (yrs.)

(i)	1	0.0090 (0.0008)	-0.0003 (0.0154)	-0.0119 (0.0168)	0.0123 (0.0092)	0.0038 (0.0112)	0.0128 (0.0132)	0.0045 (0.0137)
(j)	3	0.0090 (0.0008)	0.0135 (0.0060)	0.0102 (0.0062)	0.0134 (0.0097)	0.0089 (0.0117)	0.0134 (0.0142)	0.0089 (0.0147)
(k)	5.5	0.0090 (0.0008)	0.0154 (0.0033)	0.0142 (0.0033)	0.0125 (0.0103)	0.0115 (0.0124)	0.0124 (0.0153)	0.0113 (0.0159)
(l)	8.5	0.0090 (0.0008)	0.0065 (0.0031)	0.0079 (0.0031)	0.0091 (0.0110)	0.0104 (0.0133)	0.0091 (0.0167)	0.0105 (0.0173)
(m)	12	0.0090 (0.0008)	0.0037 (0.0025)	0.0055 (0.0026)	0.0035 (0.0117)	0.0055 (0.0142)	0.0036 (0.0183)	0.0057 (0.0188)

Table 5 (cont.)

Derivative of Wages With Respect to Education, by experience level (yrs.)								
(n)	1	0.0952 (0.0037)	0.0956 (0.0037)	0.0841 (0.0069)	0.0957 (0.0037)	0.0834 (0.0080)	0.0957 (0.0037)	0.0831 (0.0091)
(o)	3	0.0883 (0.0038)	0.0884 (0.0038)	0.0865 (0.0073)	0.0885 (0.0038)	0.0864 (0.0092)	0.0885 (0.0038)	0.0865 (0.0107)
(p)	5.5	0.0796 (0.0039)	0.0795 (0.0039)	0.0824 (0.0075)	0.0796 (0.0039)	0.0834 (0.0105)	0.0796 (0.0039)	0.0836 (0.0124)
(q)	8.5	0.0692 (0.0039)	0.0688 (0.0039)	0.0719 (0.0075)	0.0688 (0.0039)	0.0723 (0.0121)	0.0688 (0.0039)	0.0722 (0.0145)
(r)	12	0.0570 (0.0040)	0.0563 (0.0040)	0.0531 (0.0077)	0.0562 (0.0040)	0.0529 (0.0139)	0.0563 (0.0040)	0.0527 (0.0167)
Slope of Education Effect, by experience level (yrs.)								
(s)	1	-0.0035 (0.0005)	-0.0036 (0.0005)	0.0026 (0.0021)	-0.0036 (0.0005)	0.0029 (0.0031)	-0.0036 (0.0005)	0.0032 (0.0038)
(t)	3	-0.0035 (0.0005)	-0.0036 (0.0005)	-0.0003 (0.0013)	-0.0036 (0.0005)	0.0002 (0.0032)	-0.0036 (0.0005)	0.0003 (0.0040)
(u)	5.5	-0.0035 (0.0005)	-0.0036 (0.0005)	-0.0025 (0.0006)	-0.0036 (0.0005)	-0.0025 (0.0034)	-0.0036 (0.0005)	-0.0025 (0.0041)
(v)	8.5	-0.0035 (0.0005)	-0.0036 (0.0005)	-0.0045 (0.0006)	-0.0036 (0.0005)	-0.0047 (0.0035)	-0.0036 (0.0005)	-0.0048 (0.0043)
(w)	12	-0.0035 (0.0005)	-0.0036 (0.0005)	-0.0060 (0.0009)	-0.0036 (0.0005)	-0.0061 (0.0037)	-0.0036 (0.0005)	-0.0060 (0.0046)

All equations control for a quadratic time trend, cubic experience profile, and urban residence. The spline function in column 2 consists of interactions between AFQT and a variable equal to the minimum of experience and 2, the product of a dummy variable equal to 1 when experience is greater than 2 and the minimum of experience minus 2 and 2, the product of a dummy variable equal to 1 when experience is greater than 4 and the minimum of experience minus 4 and 3, the product of a dummy variable equal to 1 when experience is greater than 7 and the minimum of experience minus 7 and 3, and the product of a dummy variable equal to 1 when experience is greater than 10 and experience. Column 3 contains similar interactions involving s and experience. Columns 4-7 contain the interaction between AFQT and a fourth order polynomial in t. Column 5 and 7 contain similar interactions between s and a fourth order polynomial in t. In columns 4 and 5 the coefficients of the polynomial are constrained so that $\partial^2 w / \partial \mathrm{AFQT}, \partial t$ is 0 when t is 25.

Table 6: The Effects of Standardized AFQT and Schooling on Wages

Nonlinear Interactions of AFQT and Education with Experience

Dependent Variable: Log Wage.

Experience Measure: Actual Experience Instrumented with Potential Experience.

IV estimates (standard errors)

		Linear Interaction	Splines		Quartic Polynomial Restricted Derivative		Quartic Polynomial	
		AFQT & Ed	AFQT only	AFQT & Ed	AFQT only	AFQT & Ed	AFQT only	AFQT & Ed
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(a)	Education (s)	0.0884 (0.0035)	0.0885 (0.0036)	0.0766 (0.0103)	0.0886 (0.0035)	0.0827 (0.0079)	0.0887 (0.0035)	0.0884 (0.0091)
(b)	Black	-0.0632 (0.0066)	-0.0633 (0.0067)	-0.0641 (0.0069)	-0.0633 (0.0066)	-0.0630 (0.0066)	-0.0632 (0.0066)	-0.0633 (0.0066)
(c)	Standardized AFQT (z)	0.0155 (0.0062)	0.0474 (0.0302)	0.0945 (0.0422)	0.0207 (0.0144)	0.0284 (0.0168)	0.0285 (0.0181)	0.0180 (0.0191)
Derivative of Wages With Respect to AFQT, by experience level (yrs.)								
(d)	1	0.0281 (0.0063)	0.0274 (0.0449)	0.0131 (0.0643)	0.0283 (0.0197)	0.0277 (0.0232)	0.0278 (0.0284)	0.0274 (0.0290)
(e)	3	0.0534 (0.0065)	0.0419 (0.0704)	0.0566 (0.1029)	0.0504 (0.0282)	0.0462 (0.0337)	0.0471 (0.0446)	0.0512 (0.0448)
(f)	5.5	0.0850 (0.0067)	0.0989 (0.0925)	0.0905 (0.1380)	0.0849 (0.0376)	0.0875 (0.0450)	0.0872 (0.0625)	0.0861 (0.0625)
(g)	8.5	0.1229 (0.0069)	0.1188 (0.1116)	0.1298 (0.1741)	0.1269 (0.0481)	0.1340 (0.0577)	0.1294 (0.0832)	0.1277 (0.0828)
(h)	12	0.1672 (0.0072)	0.1626 (0.1275)	0.1404 (0.2141)	0.1650 (0.0599)	0.1523 (0.0719)	0.1582 (0.1073)	0.1619 (0.1064)
Slope of AFQT Effect, by experience level (yrs.)								
(i)	1	0.0126 (0.0011)	-0.0200 (0.0333)	-0.0814 (0.0485)	0.0089 (0.0137)	0.0033 (0.0165)	0.0036 (0.0227)	0.0104 (0.0226)
(j)	3	0.0126 (0.0011)	0.0345 (0.0427)	0.1248 (0.0641)	0.0127 (0.0149)	0.0139 (0.0180)	0.0141 (0.0258)	0.0131 (0.0257)
(k)	5.5	0.0126 (0.0011)	0.0150 (0.0345)	-0.0606 (0.0540)	0.0144 (0.0164)	0.0176 (0.0198)	0.0164 (0.0296)	0.0144 (0.0293)
(l)	8.5	0.0126 (0.0011)	-0.0017 (0.0374)	0.0868 (0.0677)	0.0130 (0.0181)	0.0120 (0.0218)	0.0111 (0.0338)	0.0127 (0.0334)
(m)	12	0.0126 (0.0011)	0.0232 (0.0293)	-0.0598 (0.0658)	0.0084 (0.0200)	-0.0023 (0.0241)	0.0070 (0.0385)	0.0059 (0.0380)

Table 6 (cont.)

Derivative of Wages With Respect to Education, by experience level (yrs.)							
(n)	1	0.0829 (0.0036)	0.0830 (0.0036)	0.0837 (0.0131)	0.0830 (0.0036)	0.0832 (0.0101)	0.0831 (0.0036)	0.0839 (0.0121)
(o)	3	0.0718 (0.0037)	0.0718 (0.0037)	0.0766 (0.0160)	0.0719 (0.0037)	0.0748 (0.0137)	0.0719 (0.0037)	0.0712 (0.0169)
(p)	5.5	0.0580 (0.0039)	0.0579 (0.0039)	0.0563 (0.0167)	0.0580 (0.0039)	0.0562 (0.0177)	0.0579 (0.0039)	0.0558 (0.0221)
(q)	8.5	0.0415 (0.0040)	0.0412 (0.0040)	0.0322 (0.0182)	0.0413 (0.0040)	0.0366 (0.0222)	0.0411 (0.0040)	0.0428 (0.0279)
(r)	12	0.0222 (0.0042)	0.0218 (0.0042)	0.0459 (0.0295)	0.0218 (0.0042)	0.0330 (0.0272)	0.0216 (0.0042)	0.0239 (0.0343)
Slope of Education Effect, by experience level (yrs.)							
(s)	1	-0.0055 (0.0006)	-0.0056 (0.0006)	0.0071 (0.0080)	-0.0056 (0.0006)	-0.0014 (0.0063)	-0.0056 (0.0006)	-0.0054 (0.0081)
(t)	3	-0.0055 (0.0006)	-0.0056 (0.0006)	-0.0141 (0.0044)	-0.0056 (0.0006)	-0.0063 (0.0068)	-0.0056 (0.0006)	-0.0067 (0.0087)
(u)	5.5	-0.0055 (0.0006)	-0.0056 (0.0006)	-0.0041 (0.0018)	-0.0056 (0.0006)	-0.0078 (0.0074)	-0.0056 (0.0006)	-0.0053 (0.0094)
(v)	8.5	-0.0055 (0.0006)	-0.0056 (0.0006)	-0.0120 (0.0057)	-0.0056 (0.0006)	-0.0046 (0.0081)	-0.0056 (0.0006)	-0.0037 (0.0102)
(w)	12	-0.0055 (0.0006)	-0.0056 (0.0006)	0.0158 (0.0156)	-0.0056 (0.0006)	0.0029 (0.0088)	-0.0056 (0.0006)	-0.0091 (0.0111)

All equations control for a quadratic time trend, cubic experience profile, and urban residence. The spline function in column 2 consists of interactions between AFQT and a variable equal to the minimum of experience and 2, the product of a dummy variable equal to 1 when experience is greater than 2 and the minimum of experience minus 2 and 2, the product of a dummy variable equal to 1 when experience is greater than 4 and the minimum of experience minus 4 and 3, the product of a dummy variable equal to 1 when experience is greater than 7 and the minimum of experience minus 7 and 3, and the product of a dummy variable equal to 1 when experience is greater than 10 and experience. Column 3 contains similar interactions involving s and experience. Columns 4-7 contain the interaction between AFQT and a fourth order polynomial in t. Column 5 and 7 contain similar interactions between s and a fourth order polynomial in t.

In columns 4 and 5 the coefficients of the polynomial are constrained so that $\partial^2 w / \partial \text{AFQT}, \partial t$ is 0 when t is 25.

Weight on Productivity when t=40, ψ_{40}	Pure Signalling Model: r=0.0					Mixed Model: r=0.05				
	Weight on Productivity when t=15 ψ_{15}					Weight on Productivity when t=15 ψ_{15}				
	0.1	0.25	0.5	0.75	1	0.1	0.25	0.5	0.75	1
0.1	.092					.095				
0.25	.091	.082				.094	.091			
0.5	.088	.079	.062			.094	.089	.082		
0.75	.086	.076	.057	.032		.093	.088	.081	.073	
1	.083	.072	.050	.017	.00[b]	.091	.087	.079	.071	.062

a) When $t \leq 15$, the weight $\psi(t)$ of actual productivity in employer expectations of productivity is $\psi_{15}=(b_{z1}t + b_{z2}t^2 + b_{z3}t^3 + b_{z4}t^4)/(b_{z1}15 + b_{z2}15^2 + b_{z3}(15^3) + b_{z4}(15^4))$, and $b_{z1},..., b_{z4}$ are the coefficients on the interactions between AFDC and the first four powers of experience t in the regression model underlying column 5 in Table 5 in the case of Table 7 and column 5 in Table 6 in the case of Table 8. ψ_{15} is set to the values in the column heading. When t is greater than 15, $\psi(t)$ is $\psi_{15} + (\psi_{40}-\psi_{15})(t-t_1)/(40-t_1)$, where 40 is the length of the career. The internal rate of return should be compared to the $Y_s/Y_{s-1} - 1$, which is exp(.1)-1.
b) The internal rate of return to a year of school is negative.

Weight on Productivity when t=40, ψ_{40}	Pure Signalling Model: r=0.0					Mixed Model: r=0.05				
	Weight on Productivity when t=15 ψ_{15}					Weight on Productivity when t=15 ψ_{15}				
	0.1	0.25	0.5	0.75	1	0.1	0.25	0.5	0.75	1
0.1	.091					.094				
0.25	.089	.078				.094	.089			
0.5	.087	.075	.053			.093	.087	.078		
0.75	.084	.071	.046	.015		.092	.086	.077	.066	
1	.081	.067	.038	.00[b]	.00[b]	.091	.084	.075	.064	.053

Table 8: The Internal Rate of Return to Education in a Pure Signalling Model and a Mixed Model under various assumptions about the Rate that Employers Learn about Productivity. Coefficients from IV Estimate of Wage Equation Using Actual Experience [a]

a, b) See Table 7.