Employer Learning and Statistical Discrimination

Author(s): Joseph G. Altonji and Charles R. Pierret

Source: *The Quarterly Journal of Economics*, Vol. 116, No. 1 (Feb., 2001), pp. 313-350
Published by: Oxford University Press
Stable URL: http://www.jstor.org/stable/2696451
Accessed: 17-01-2016 15:34 UTC

EMPLOYER LEARNING AND STATISTICAL DISCRIMINATION*

JOSEPH G. ALTONJI AND CHARLES R. PIERRET

We show that if firms statistically discriminate among young workers on the basis of easily observable characteristics such as education, then as firms learn about productivity, the coefficients on the easily observed variables should fall, and the coefficients on hard-to-observe correlates of productivity should rise. We find support for this proposition using NLSY79 data on education, the AFQT test, father's education, and wages for young men and their siblings. We find little evidence for statistical discrimination in wages on the basis of race. Our analysis has a wide range of applications in the labor market and elsewhere.

I. INTRODUCTION

People go through life making an endless stream of judgments on the basis of limited information about matters as diverse as the safety of a street, the quality of a car, the suitability of a potential spouse, and the skill and integrity of a politician. When hiring, employers must assess the value of potential workers with only the information contained in resumes, recommendations, and personal interviews. Do employers "statistically discriminate" among young workers on the basis of easily observable variables such as education, race, and other clues to a worker's labor force preparation? As they learn over time, do they rely less on such variables? These questions are directly relevant for many issues in labor economics including the signaling model of education [Spence 1973; Weiss 1995], statistical theories of discrimination [Aigner and Cain 1977; Lundberg and Startz 1983], the interpretation of earnings dynamics, and the design of institutional mechanisms for hiring and firing workers.

* This research was supported by the Institute for Policy Research, Northwestern University, the Bureau of Labor Statistics, U. S. Department of Labor under contract B9J34633, and the National Science Foundation under grant SBR-9512009. We owe a special debt to Nachum Sicherman for assisting us with the NLSY79 data. We thank Paul Devereux, Henry Farber, Robert Gibbons, Judith Hellerstein, Lawrence Katz, Derek Neal, Bruce Weinberg, three anonymous referees, and participants in seminars at the University of California, Berkeley, Boston University, Brigham Young University, the Bureau of Labor Statistics, Columbia University, the Federal Reserve Bank of Chicago, Indiana University, the University of Maryland, McMaster University, National Bureau Economic Research, Northwestern University, Tel Aviv University, the University College London, the UpJohn Institute, and the University of Western Ontario for helpful comments. We are responsible for all errors and omissions. The opinions stated in this paper do not necessarily represent the official position or policy of the U. S. Department of Labor.

The Quarterly Journal of Economics, February 2001

In this paper we explore the implications of a hypothesis that we refer to as Employer Learning with Statistical Discrimination, or EL-SD. Our working hypothesis is that firms, with only limited information about the quality of workers in the early stages of their careers, distinguish among workers on the basis of easily observable variables that are correlated with productivity. These might include years of education or degree, the quality of the school the person attended, race, and gender. (To avoid misunderstanding, we wish to stress that part of the relationship between wages and race and gender may reflect biased inferences on the part of employers or forms of discrimination that have nothing to do with productivity or information.) Firms weigh this information with other information about outside activities, work experience to date, references, the job interview, and perhaps formal testing by the firm. Each period, the firm observes noisy indicators of the worker's performance. Over time, these observations make the initial information redundant.

The main contribution of the paper is to provide a way to test for whether firms statistically discriminate on the basis of readily available information such as education and race. Under some strong assumptions our econometric model also provides a way to estimate the learning profile of firms up to a scale parameter, an issue that we pursue in more detail in Altonji and Pierret [1998] (hereinafter AP [1998]).

Our research builds on some previous work, particularly Farber and Gibbons [1996] (hereinafter FG).[1] FG investigate

1. Other relevant references are Gibbons and Katz [1991] which we discuss below and Parsons [1993]. Glaeser [1992] uses variances in wage innovations as a measure of learning. His work is somewhat closely related to FG. However, he attempts to distinguish between information that is specific to the job match and information about general productivity. 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. Studies following performance evaluations within firms based on the EOPP data, or studies using firm personnel files [Medoff and Abraham 1980] are also relevant, but have a very different focus than the present paper. 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. Albrecht [1981] 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. Montgomery [1991] presents a model in which employers obtain valuable information on the productivity of new employees through referrals and is part of a large literature on labor market networks. For empirical evidence see Holzer [1988].

three implications of employer learning when information is common across firms and the labor market is competitive, key (and strong) assumptions that we also make. Imagine two variables that affect productivity, s (say schooling) that firms can observe directly and z (say AFQT test scores) that 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. Their empirical evidence is generally supportive of this result, although they note that a positive interaction could arise if schooling is complementary with training. Second, they establish and obtain empirical support for the proposition that the part of z (say $\tilde{z}$) that is orthogonal to information available to employers at the beginning of a worker's career 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. FG reject this stark prediction in favor of a more general wage growth model that includes a stationary component.

In this paper we establish a different but related proposition that allows us to examine the issue of statistical discrimination. The proposition concerns how controlling for the experience profile of the effect of all of z (not just $\tilde{z}$) on wages alters the interaction between experience and s. We show that if s and z are positively correlated, so that s is informative about z, then statistical discrimination in the presence of employer learning implies both that the coefficient on z will rise with experience and that the coefficient on s will fall.[2] Our proposition provides a solution to a fundamental identification problem that has blocked tests of statistical discrimination—one cannot tell whether a correlation between the wage and an easily observed variable arises because imperfectly informed firms use the variable to statistically discriminate or because the variable happens to be correlated with information about productivity that is used directly by the firm but not by the econometrician.

We use the proposition to study statistical discrimination on the basis of education using the AFQT test, father's education,

2. Our analysis is fully consistent with FG's analysis of the orthogonal component $\tilde{z}$. In particular, introducing the interaction between z and experience into the wage model affects the interaction between experience and s only if z and s are correlated.

and wage rates of older siblings as the hard-to-observe z variables for a sample of young men from NLSY79. We find that the wage coefficients on the z variables rise with experience while the coefficient on education falls. These results provide support for the hypothesis that firms statistically discriminate on the basis of education. We also explore the implications of statistical discrimination on the basis of race, which is also easily observable to employers and is correlated with hard-to-observe background variables that influence productivity.[3] Subject to some important caveats our estimates suggest that statistical discrimination on the basis of race plays a relatively minor role in the race gap in wages. We do not address the issue of discrimination in employment.

In Section II we present our basic theoretical framework. We also consider alternate hypotheses for the interactions between s, z, and experience. In Section III we discuss the NLSY79 data and the econometric specification used in the study. In Sections IV and V we present our results for education and race. In Section VI we present results in which we control for job training. In Section VII we close the paper with an extended discussion of some of the additional implications of our analysis and a research agenda.

II. Implications of Statistical Discrimination and Employer Learning for Wages

II.1. A Model of Employer Learning and Wages

In this section we show how the wage coefficients on characteristics that employers can observe directly and on characteris-

3. We are using the term "statistical discrimination" as synonymous with the use of the term "rational expectations" in the economics literature. We mean that in the absence of full information, firms distinguish between individuals with different characteristics based on statistical regularities. That is, firms form stereotypes that are rational given their information. 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. This difficulty may lead to negative outcomes for minorities because it lowers their incentives to make unobservable investments that raise productivity or if the productivity of a job match depends on the fit between the worker and the job. Some papers also consider whether firms that start with incorrect beliefs about the relationship between personal characteristics and productivity (inaccurate stereotypes) would correct them, and, in models with worker investment, whether the priors held by firms may be self-fulfilling. See Aigner and Cain [1977], Lundberg and Startz [1983], Lang [1986], Coate and Loury [1993], and Oettinger [1996]. See Altonji and Blank [1999] for a recent survey of this research.

tics they cannot observe directly will change with experience if employers statistically discriminate and become better informed about workers over time.

Our model is similar to FG. Let y_{it} be the log of labor market productivity of worker i with t_i years of experience:

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

In (1) we separate the determinants of productivity into four categories: s_i represents variables that are observed by both the employer and the econometrician; q_i includes variables observed by the employer but not seen (or not used) by the econometrician; z_i consists of correlates of productivity that are not observed directly by employers but are available to and used by the econometrician; and η_i is an index of other determinants of productivity and is not directly observed by the employers and not observed (or observed but not used) by the econometrician. We normalize z_i so that all the elements of the conformable coefficient vector Λ are positive. In addition, $H(t_i)$ is the experience profile of productivity. For now we assume that the experience profile of productivity does not depend on s_i, z_i, q_i, or η_i. In subsection II.C we discuss the sensitivity of our analysis to this assumption. To simplify the exposition, all variables are expressed as deviations from population means, and we abstract from economywide trends in the link from z and s to y, although we control for them in the empirical work. Additionally, in most of the analysis we suppress the i subscript.

In the absence of knowledge of z and η, firms form the conditional expectations $\mathrm{E}(z|s,q)$ and $\mathrm{E}(\eta|s,q)$, which we assume are linear in q and s. Consequently,

$$\begin{aligned} z &= \mathrm{E}(z|s,q) + v = \gamma_1 q + \gamma_2 s + v \\ \eta &= \mathrm{E}(\eta|s,q) + e = \alpha_2 s + e, \end{aligned} \tag{2}$$

where the vector v and the scalar e have mean 0 and are uncorrelated with q and s by definition of an expectation.[4] The links from s to z and η may be due in part to a causal effect of s.[5] Equations (1) and (2) imply that $\Lambda v + e$ is the error in the

4. 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.

5. For example, below we use the Armed Forces Qualification Test (AFQT) as z and years of education as s, and Neal and Johnson [1996] present evidence that years of education have a sizable positive effect on AFQT.

employer's belief about the log of productivity of the worker at the time the worker enters the labor market. The sum $\Lambda v + e$ is uncorrelated with s and q. We make the additional assumption that $\Lambda v + e$ is independent of q and s.

Firms do not see y_t, but each period that a worker is in the labor market, firms observe a noisy signal of the productivity of the worker, $\xi_t = y + \epsilon_t$, where $y = y_t - H(t)$. 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. It is assumed to be independent of the other variables in the model.[6] Since the employers know q and s, observing ξ_t is equivalent to observing $d_t = \xi_t - \mathrm{E}(y|s,q) = \Lambda v + e + \epsilon_t$, which is the sum of the noise ϵ_t and the error $\Lambda v + e$ in the employer's belief about initial log productivity. The vector $D_t = \{d_1, d_2, \ldots, d_t\}$ summarizes the worker's performance history. Let μ_t be the difference between $\Lambda v + e$ and $\mathrm{E}(\Lambda v + e|D_t)$. By definition μ_t is uncorrelated with D_t, q, and s, but in addition we assume that μ_t is distributed independently of D_t, q, and s. We assume that q, s, and D_t are known to all employers, as in FG.

As a result of competition among firms, the worker receives a wage W_t equal to $\mathrm{E}(Y_t|s,q,D_t)\ \exp^{\zeta_t}$, where Y_t is the level of productivity $\exp^{y_t}$, $\mathrm{E}(Y_t|s,q,D_t)$ is expected productivity conditional on s, q, and D_t, and $\exp^{\zeta_t}$ reflects measurement error and firm-specific factors that are outside the model and are unrelated to s, z, and q. Substituting and taking logs, we arrive at the log wage process:

$$(3)\quad w_t = (r + \Lambda\gamma_2 + \alpha_2)s + H^*(t) + (\alpha_1 + \Lambda\gamma_1)q + \mathrm{E}(\Lambda v + e|D_t) + \zeta_t,$$

where $w_t = \log(W_t)$ and $H^*(t) = H(t) + \log(\mathrm{E}(\exp^{\mu_t}))$. The presence of $\mathrm{E}(\Lambda v + e|D_t)$ in (3) shows that wages change over time not just because productivity changes with experience, but

6. We are also implicitly assuming that the component of ϵ_t that reflects temporal variation in productivity from sources specific to worker i is serially uncorrelated. Otherwise, firms would have an incentive to base compensation in $t + 1$ on what they know about the worker-specific component of ϵ_t. However, ϵ_t may be serially correlated as a result of the other factors. The firm's knowledge of a serially correlated productivity component would imply serially correlated transitory variation in the wage error of the type found by FG, but would not have much effect on our analysis.

also because firms learn about errors in their initial assessment of worker productivity.

In the context of the debate over signaling models of education, Riley [1979] and others have noted that unless the relationship between schooling and actual productivity changes with experience, the coefficient on s will not change. This is true regardless of *why* s is related to productivity. FG make this point by showing in a similar model that the expected value of the coefficient of an OLS regression of the wage level W_t on s does not depend on t.

FG also make a second point. If one adds $\tilde{z}$, the part of z that is uncorrelated with the employer's initial information, to the wage equation, the coefficient on s remains constant (adding a variable to a regression has no effect on the coefficient of an uncorrelated variable) but the coefficient on $\tilde{z}$ rises with t. This is because $\tilde{z}$ will be positively correlated with the change over time in $\mathrm{E}(\Lambda v + e|D_t)$ that arises if employers learn. They provide evidence from NLSY79 that the effect of s on W_t is relatively constant while the effect of $\tilde{z}$ is increasing in t.[7]

Our contribution is to study the experience profiles of s and z rather than s and $\tilde{z}$. By examining the change with experience in the coefficients on s and z when s is informative about z and employers learn, we can study statistical discrimination. We proceed by examining the parameters of the conditional expectation of w_t given s, z, t, and the experience profile $H^*(t)$. We begin with the case in which z and s are scalars and then turn to the more general cases.

Consider the conditional expectation function when $t = 0, \ldots, T$, with

$$\mathrm{E}(w_t|s,z,t) = b_{st}s + b_{zt}z + H^*(t). \tag{4}$$

To simplify the algebra but without any additional assump-

7. It may be helpful to briefly summarize the specifics of how our model differs from FG, which is more general. First, we specify the production function (1) as linear and measure output y in logs, while FG specify output in levels and work with essentially any conditional distribution of output given the variables on the right side of (1). Second, in (2) and (3) we specify that conditional expectations of the worker characteristics z and η are linear in q and s with independent error, while FG do not place this restriction on the joint distribution. Our formulation allows us to work with log wages, which facilitates comparison to the large literature that works in logs. One could obtain results similar to ours using a linear production function. As we have already noted, the main substantive difference is that we analyze the behavior of the coefficients in a wage equation containing s and z when s and z are correlated. In footnote 18 we discuss differences in the specifics of sample choice, etc.

tions, we reinterpret s, z, and q as the components of s, z, and q that are orthogonal to $H^*(t)$.[8] Given that the wage evolves according to (3), the omitted bias formula for least squares regression implies that

$$\begin{aligned} b_{st} &= b_{s0} + \Phi_{st} = [r + \Lambda\gamma_2 + \alpha_2] + \Phi_{qs} + \Phi_{st} \\ b_{zt} &= b_{z0} + \Phi_{zt} = \Phi_{qz} + \Phi_{zt}, \end{aligned} \tag{5}$$

where Φ_{qs} and Φ_{qz} denote the coefficients of the auxiliary regressions of $(\alpha_1 + \Lambda\gamma_1)q$ on s and z, respectively, and Φ_{st} and Φ_{zt} are the coefficients of the regression of $\mathrm{E}(\Lambda v + e|D_t)$ on s and z. Note that $\mathrm{E}(\Lambda v + e|D_0) = 0$ because there is no work history when $t = 0$, so Φ_{s0} and Φ_{z0} equal 0. The coefficients b_{s0} and b_{z0} pick up part of the effect of q, which is used by employers to estimate productivity but is omitted from the regression.

Using the facts that cov $(s, \mathrm{E}(\Lambda v + e|D_t)) = 0$ and cov $(z, \mathrm{E}(\Lambda v + e|D_t)) =$ cov $(v, \mathrm{E}(\Lambda v + e|D_t))$ and the least squares regression formula, one may express Φ_{st} and Φ_{zt} as

$$\begin{aligned} \Phi_{st} &= \theta_t \Phi_s \\ \Phi_{zt} &= \theta_t \Phi_z, \end{aligned} \tag{6}$$

where Φ_s and Φ_z are the coefficients of the regression of $\Lambda v + e$ on s and z and

$$\theta_t = \frac{\mathrm{cov}\,(\mathrm{E}(\Lambda v + e|D_t),\, z)}{\mathrm{cov}\,(\Lambda v + e,\, z)} = \frac{\mathrm{cov}\,(\mathrm{E}(\Lambda v + e|D_t),\, v)}{\mathrm{cov}\,(\Lambda v + e,\, v)}. \tag{7}$$

Equations (6) and (7) say that the experience paths of b_{st} and b_{zt} depend on the signs of Φ_s and Φ_z and the experience path of θ_t. It can be shown that $\Phi_s < 0$ and $\Phi_z > 0$ if cov $(\Lambda v + e, v) > 0$ and cov $(s,z) > 0$. The latter condition is true when s is schooling and the scalar z is AFQT, father's education, or the wage rate of an older sibling. The condition cov $(\Lambda v + e, v) > 0$ simply states that the unobserved (by the firm) productivity subcomponent v and composite unobserved productivity term $\Lambda v + e$ have a

8. Estimates of the experience profile $H^*(t)$ and the economywide trend will be affected if the mean of q depends on t through the age cohort of the individual, but this has no bearing on our analysis. We are making the implicit assumption that the other parameters of the model do not depend on the age cohort of the sample members conditional on experience t and the s and z specific economywide trends that we introduce in the empirical specification of wages.

positive covariance. This seems plausible to us for the z variables we consider.

The parameter θ_t summarizes how much the firm knows about $\Lambda v + e$ at experience t. θ_t is bounded between 0 and 1. It is 0 in period 0, because in this period employers know nothing about $\Lambda v + e$. The coefficient is 1 if $E(\Lambda v + e|D_t)$ is $\Lambda v + e$, since in this case the employer has learned what $\Lambda v + e$ is and thus knows productivity y. It is also 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 $\Lambda v + e$.[9] This is the basis for Proposition 1.

PROPOSITION 1. Under the assumptions of the above model, a) the regression coefficient b_{zt} is nondecreasing in t, and b) the regression coefficient b_{st} is nonincreasing in t.[10]

The intuition for the decline in b_{st} is that as employers learn the productivity of workers, s will get less of the credit for an association with productivity that arises because s is correlated with z, provided that z is included in the wage equation with a time-dependent coefficient and can claim the credit. It immediately follows that if firms learn nothing new about the worker, then $E(\Lambda v + e|D_t)$ does not change with t, θ_t does not change, and b_{st} and b_{zt} are constants.[11]

It is also easy to show using the least squares regression formula that the model implies that $\Phi_s = -\Phi_z\Phi_{zs}$ where Φ_{zs} is the coefficient of the regression of z on s, which is the basis for the next proposition.

PROPOSITION 2. Under the assumptions of the above model,

$$\frac{\partial b_{st}}{\partial t} = -\Phi_{zs}\frac{\partial b_{zt}}{\partial t}.$$

9. To establish this, note that since D_{t-1} is a subset of the information in D_t, $[\text{cov}(v, E(\Lambda v + e|D_t) - E(\Lambda v + e|D_{t-1}))]/\text{cov}(v, \Lambda v + e) = \theta_t - \theta_{t-1} \geq 0$.

10. 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 that s is negatively correlated with the unfavorable z, b_{st} will fall with t. As noted earlier, we have normalized z so that $\Lambda > 0$.

11. Note also that the experience path of the parameter b_{zt} provides an estimate of the time profile of θ_t up to the scale parameter Φ_z. This means that under the assumption that employers learn about v and e at the same rate, one can estimate the time profile of employer learning about productivity up to scale. In AP [1998] we examine the implications of our estimates for pure signaling models of the return to education. The faster firms learn, the less relevant signaling is.

Since Φ_{zs} is simply the regression coefficient of z on s and can be estimated, the coefficient restriction in Proposition 2 provides some leverage in differentiating between the learning/statistical discrimination model and alternative explanations for the behavior of b_{st} and b_{zt}. Proposition 2 holds because s is part of the firm's initial information set, which means that the effects of learning on b_{st} arise solely out of the relationship between s and z. The fact that the effects of learning on the coefficient on z will spill over to the coefficients on variables that firms use to statistically discriminate among new workers is the essence of our test for statistical discrimination.

When s and z are vectors (and we reinterpret related variables and parameters as vectors or matrices accordingly), we can no longer make the strong statement in Proposition 1. Consider the case where z and s are $K \times 1$ and $J \times 1$ vectors, b_{zt} is a $1 \times K$ vector with $b_{z_k t}$ as the kth element, and b_{st} is a $J \times 1$ vector with $b_{s_j t}$ as the jth element. One cannot in general sign $\partial b_{s_j t}/\partial t$ and $\partial b_{z_k t}/\partial t$ even if all the elements of Λ are positive, each element of cov $(z, \Lambda v + e)$ is positive, and all coefficients of the regression of s on z are positive.[12] However, a matrix version of Proposition 2 still holds

$$\frac{\partial b_{st}}{\partial t} = -\frac{\partial b_{zt}}{\partial t}\Phi_{zs},$$

where Φ_{zs} is now the $K \times J$ matrix of coefficients of the regression of z on s. This places J restrictions on the parameters on s and z. It also indicates that if $\partial b_{z_k t}/\partial t > 0$ and $\Phi_{z_k s} > 0$ for all z_k used in the analysis, then $\partial b_{st}/\partial t < 0$.[13] These conditions hold in our sample when s is education and the z vector consists of the AFQT test, the sibling wage rate, and father's education.[14]

Note that the time paths of the elements of b_{zt} will reflect the rate at which firms learn about the productivity components that they are correlated with. This is an important result, because it means that differences in the effects of particular variables on

12. The intuition is that in the multivariate regression of $\Lambda v + e$ on z the coefficient on the kth element z_k can be negative even if all of the coefficients of the simple regressions relating $\Lambda v + e$ to the elements of z are positive.

13. See AP [1997] for more details.

14. In the empirical work we include controls for some additional variables, such as location dummies, which we do not interact with experience. These can be viewed as s variables, although they may also capture demand side factors related to productivity or compensating differentials. Their presence does not alter the predictions of the model for the s variables we do interact with experience.

wage growth may reflect differences in the rate at which firms learn about the variables. Thus, EL-SD provides an alternative or a complement to the standard view that the differential effects on growth rates reflect differences in the relationship between the variables and other sources of wage growth such as on-the-job training.

II.2. Statistical Discrimination on the Basis of Race

By almost any measure, young black men are disadvantaged relative to whites in the United States. On average, black males have poorer, less educated parents, are more likely to grow up in a single-parent household, live in more troubled neighborhoods, attend schools with fewer resources, and have fewer opportunities for teenage employment than white males. Many of these factors are correlated with educational attainment and labor market success. They are likely to lead to a black/white differential in the average skills of young workers. Discrimination in various forms may further hinder the development of human capital in black children and add to a gap in skills that is due to the race difference in socioeconomic background. The gap in some indicators of skill is very large. In our regression sample, the unweighted mean of the standardized AFQT score for blacks is 1.11 standard deviations below the mean for whites. Neal and Johnson [1996] and others have shown that in the NLSY79 sample of men a substantial part of the race gap in wages is associated with the race gap in AFQT.

If premarket discrimination is an important factor in the gap between the average skills of black and white workers, then it seems likely that various forms of current labor market discrimination contribute to race differences in wages that are unrelated to skill. However, it is nevertheless interesting to examine the possibility that a correlation between race and skill might lead a rational, profit-maximizing employer to use race as a cheap source of information about skills. Such statistical discrimination along racial lines can have very negative social consequences and is against the law. However, it would be hard to detect.

A statistically discriminating firm might use race, along with education and other information to predict the productivity of new workers. With time, the productivity of the worker would become apparent, and compensation would be based on the larger information that accumulates with experience rather than the limited information available at the time of hire.[15] In this case, race can be thought of as an s variable. Our model implies almost immediately that the coefficient on race does not vary over time if the interaction between z and t is excluded from the model. This is because the initial wage already incorporates information the employer has about how race is related to productivity. If the interaction between z and t is included, then the model implies that the coefficient on race will rise over time. The intuition is that firms initially pay less to blacks because race is negatively correlated with productivity conditional on the firms' information set, s and q. As experience accumulates, firms base pay on s, q, and D_t. This leads the coefficient on z to rise, which in turn leads to a lower weight on race.

In contrast, if firms obey the law and do not use race as information, then in the econometric model, race has the properties of a z variable. First consider the case where race is the only z variable in the equation. In this case our model implies that if (i) race is negatively related to productivity ($\Lambda < 0$), (ii) firms do not statistically discriminate on the basis of race, and (iii) firms learn over time, then (a) the race gap when experience is 0 will be smaller than if firms illegally use race as information and (b) the race differential will widen as experience accumulates. The intuition for (b) is that firms are acquiring additional information about performance that may legitimately be used to differentiate among workers. If race is negatively related to productivity, then the new information will lead to a decline in wages. If education is negatively related to race, then the coefficient on education should fall with experience.

Now consider what happens when one adds a second z variable (one that is positively related to productivity) and its interaction with t to a model that contains race and an s variable. In Appendix 1 we show that if the coefficient on this new z variable in the regression of race on s and this variable is negative, then the coefficient on the interaction between race and t will be less

15. The element of r corresponding to the race indicator s_1 in the productivity equation (1) and the wage equation (3) is 0 unless consumer or employee tastes for discrimination reduce profitability of employing members of the minority group, as in Becker [1971]. (Even if r is 0, race may be negatively related to productivity if it is correlated with elements of z, q, or η that affect productivity.) Presumably, firms that violate the law and discriminate in response to their own prejudice or the prejudice of consumers or other employees might also be willing to use race as information. Employers who harbor prejudice against certain groups may be especially unlikely to form beliefs about the productivity of those groups that are rational in the statistical sense used in this paper.

negative when the new z variable is included in the wage equation than when it is excluded. We conclude that if firms do not statistically discriminate on the basis of race and race is negatively related to productivity, then (1) the race gap will widen with experience, and (2) adding a favorable z variable to the model will reduce the race difference in the experience profile. We examine this below as well as some additional implications for the race intercept. We wish to stress that other factors that determine race differences in experience profiles as well as other forms of discrimination will also influence the wage results. We discuss some of these in the next subsection.

II.3. Alternative Explanations for Variation in the Wage Coefficients with Experience

The analysis so far assumes that the effects of z and s on the log of productivity do not depend on t. Human capital accumulation is included in the model through the $H(t)$ and $H^*(t)$ functions but is assumed to be "neutral" in the sense that it does not influence the experience paths of the effects of s and z on productivity.[16] In the more general case, the links between productivity and s and z may depend on experience. This would affect the b_{st} and b_{zt}.

One potential mechanism for such impacts is differential access to or benefits from on-the-job training. Most discussions of human capital and most of the empirical evidence on employer-provided training suggest that education and ability make workers more trainable and that more educated and more able workers receive more training. If this is the case, one might expect the effect of education and AFQT on wages to increase over time. We would not expect, however, that the effect of education would decrease over time as is predicted by EL-SD. As it turns out, we find that b_{st} does decrease over time, which is only consistent with a training interpretation if education *reduces* learning by doing, the productivity of training investments, or the quantity of training investments.

Having a measure of employee training does not by itself allow us to disentangle the effects of learning from those of training. To see why, consider the following extension to our basic

16. One may easily modify the theoretical framework to allow for this form of human capital accumulation. For example, the $H(t)$ function may reflect learning by doing in all jobs that is observable to firms, or worker-financed investments in human capital that are observable to firms.

model. Assume that units of training in period t, R_t, is determined by employer beliefs about productivity given D_t, q, s, and t, as well as by D_t, q, s, and experience, that productivity is a linear function of the sum $\Sigma R_\tau = \Sigma_{\tau=1..t} R_\tau$, of current and past training, and that a unit of R_t costs c in current productivity. There are two points to emphasize. First, even if the training profile depends only on information that is known to the firm when t is 0, the relationship between q and R_t and ΣR_τ may change with t, leading the coefficients on s and z to depend on t even if there is no learning.

The second point is that training may depend on D_t. To see the implications of this possibility, suppose that (1) learning is important, (2) variation with s and z in the rate of skill accumulation is not, and (3) variation in our measure of training is driven by worker performance (which leads to promotion into jobs that offer training) rather than by exogenous differences in the level of human capital investment. Even under this hypothesis one would expect the introduction of the training measures to lead to a reduction in the growth with t in the coefficient on z and a reduction in the impact of z on the experience path of the coefficient on s.

For both reasons, we cannot separate the effects of training from the effects of statistical discrimination with learning if, as seems plausible, the quantity of training is influenced by the employer beliefs about productivity. With an indicator of y_t, the identification problem is easily solved, but we lack such an indicator. Despite the absence of a clear structural interpretation, we think it is important in this initial study to see how introducing measures of training alters b_{st} and b_{zt}, and we do so below.

Training may also affect our findings concerning statistical discrimination with respect to race. On one hand, ability differences that are correlated with race and that influence the productivity of training may lead the race gap to widen with experience because of differential human capital formation rather than labor market discrimination. On the other hand, discrimination-related differences in access to networks or to mentors may affect training, and promotions may also cause wages for African-Americans to decrease over time relative to whites.

If taste-based racial discrimination and "social distance" between blacks and whites become more important in higher level positions, a widening of the race gap with experience may be a reflection of increased discrimination rather than employer learning. Perhaps most importantly, we model statistical discrimination in wages and do not analyze the implications of an extended model in which statistical discrimination influences the decision to hire. Statistical discrimination in employment is likely to have effects on the wage/experience profiles that we estimate. In light of these and other possible alternative explanations, our results concerning statistical discrimination based on race should be interpreted cautiously.

III. Data and Econometric Specification

The empirical analysis is based on the 1992 release of NLSY79. The NLSY79 is a panel study of men and women who were aged 14–21 in 1978. Sample members have been surveyed annually since 1979. (In 1994 the NLSY79 moved to a biennial survey schedule.) We restrict the analysis to men who are white or black and who have completed eight or more years of education. We exclude labor market observations prior to the first time that a person leaves school and accumulate experience from that point. When we analyze wage changes, we further restrict the sample to persons who do not change education between successive years. 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 six. To reduce the influence of outliers, father's education is set to 4 if father's education is reported to be less than 4. AFQT is standardized by age of the individual at the time of the test. The means, standard deviations, minimum, and maximums of the variables used in analysis are provided in Table VI in Appendix 2. 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 NLSY79 and contribute 28.8 percent of our observations. Appendix 2 provides more details about how the sample was selected and how key variables were constructed.[17]

17. Although we use different sample selection rules than FG and use the log of wages rather than the level, in preliminary work our results were not sensitive to these differences. FG use both men and women, include Hispanics, and restrict their sample to persons who have worked at least three consecutive years since attending school. We also experimented with another variable FG use—an indicator for whether any person in the respondent's household had a library card at the time the respondent was fourteen. Like FG, we used the residual obtained from a regression of the library card variable on the initial real wage, education,

Our basic econometric specification is an OLS regression of log wage on our s variable, years of schooling, and various z variables, including the AFQT score, the log of the wage of a sibling, and father's education. The coefficients on these explanatory variables are allowed to vary with experience. Our wage equations control for a cubic in experience, residence in an urban area, and dummy variables for whether the sibling's wage is missing, whether father's education is missing, and whether the sibling whose wage is being used is a female. We add interactions between the dummy variables for missing data and experience when interactions between sibling's wage and experience and father's education and experience are added to the model. These variables are not reported in the regression tables.

One possible objection to our theoretical formulation is that it assumes that the flow of information to employers is independent of the type of job in which the worker begins his career. This is contrary to the idea that some jobs are "dead-end" jobs. Perhaps education (and high AFQT) enables a worker to gain access to jobs in which firms have the ability to observe whether the worker has higher level skills that are strongly related to productivity. For this reason, we include controls for the two-digit occupation of the first job.[18]

Murphy and Welch [1992], Katz and Murphy [1992], Taber [1996], and Chay and Lee [1998] are among a large number of recent studies of economywide changes in the structure of wages in the United States. (See Katz and Autor [1999] for a recent survey.) Failure to control for secular change in the wage structure could bias our estimates of the effect of experience on the wage structure.[19] Our wage equations control for calendar year

part-time status, an interaction between education and part-time status, race, sex, age, and calendar year. We confirm FG's finding that the wage coefficient on the library card residual increases with experience. We also confirm that the results for the library card and AFQT residuals are weakened substantially when these residuals are interacted with calendar time. However, when we use the library card variable rather than the residual in the wage equation, the effect of the library card variable falls rather than rises with experience. We thank Henry Farber for assisting us in reconstructing the FG sample.

18. AP [1997] report qualitatively similar results with the occupation dummies excluded. An interesting project for future research would be to use information from the Dictionary of Occupational Titles on skill requirements of occupations and trace how easy-to-observe and hard-to-observe productivity characteristics are related to changes over a career in the skill requirements of the job a worker holds.

19. Since calendar time is positively correlated with experience t in a panel data set, EL-SD implies that estimates of secular changes in the return to

dummies, education interacted with a cubic time trend, and Black interacted with a cubic time trend. Where appropriate, models also include AFQT, sibling wage rates, and father's education interacted with a cubic in calendar time. (As a general rule, any s or z variable or missing value indicator for such a variable that is entered in a model is also interacted with a cubic in calendar time.) The time trend interactions are normalized so that the "main effects" of education, Black, AFQT, father's education, and the sibling wage reported in the tables refer to 1992 for a person with 0 experience. An unavoidable consequence of having to control for secular change in the wage structure is that we are relying on variation across age cohorts in NLSY79 in experience to identify the interactions between experience t and our s and z variables. This reduces the precision of our estimates substantially.[20] It also raises the possibility of bias if, for example, the younger cohorts in NLSY79 are different from the older cohorts. We are maintaining that these differences are minor. See AP [1997] for detailed results with the time trend interactions excluded. They are qualitatively similar to those with the time trends but are much more precise and provide stronger support for EL-SD.

IV. Results for Education

IV.1. AFQT as a z Variable

In Panel 1 of Table I we report OLS estimates of (4) using potential experience as the experience measure t. Throughout the paper the reported standard errors and test statistics are based on White/Huber standard errors that account for arbitrary forms of heteroskedasticity and correlation among the multiple observations for each person.

In column (1) we present an equation that includes education, AFQT, Black, and education $\times$ $t/10$. This corresponds to (4) with s equal to education and b_{st} restricted to $b_{st} = b_{s0} + b_{s1} \times t$ and $b_{zt} = b_{z0}$. Throughout the paper we normalize the interactions between s and z variables with experience to represent the change in the wage slope between $t = 0$ and $t = 10$. The

education and AFQT will be biased in opposite directions if one fails to add the interaction between these variables and experience t to the model.

20. Cawley, Heckman, and Vytlacil [1998] stress the difficulty of identifying models in which the returns to both ability and education depend on both age and time.

TABLE I

THE EFFECTS OF STANDARDIZED AFQT AND SCHOOLING ON WAGES

Dependent Variable: Log Wage; OLS estimates (standard errors).

Panel 1—Experience measure: potential experience

Model:	(1)	(2)	(3)	(4)
(a) Education	0.0586	0.0829	0.0638	0.0785
	(0.0118)	(0.0150)	(0.0120)	(0.0153)
(b) Black	−0.1565	−0.1553	0.0001	−0.0565
	(0.0256)	(0.0256)	(0.0621)	(0.0723)
(c) Standardized AFQT	0.0834	−0.0060	0.0831	0.0221
	(0.0144)	(0.0360)	(0.0144)	(0.0421)
(d) Education *	−0.0032	−0.0234	−0.0068	−0.0193
experience/10	(0.0094)	(0.0123)	(0.0095)	(0.0127)
(e) Standardized AFQT *		0.0752		0.0515
experience/10		(0.0286)		(0.0343)
(f) Black * experience/10			−0.1315	−0.0834
			(0.0482)	(0.0581)
R^2	0.2861	0.2870	0.2870	0.2873

Panel 2—Experience measure: actual experience instrumented by potential experience

Model:	(1)	(2)	(3)	(4)
(a) Education	0.0836	0.1218	0.0969	0.1170
	(0.0208)	(0.0243)	(0.0206)	(0.0248)
(b) Black	−0.1310	−0.1306	0.0972	0.0178
	(0.0261)	(0.0260)	(0.0851)	(0.1029)
(c) Standardized AFQT	0.0925	−0.0361	0.0881	0.0062
	(0.0143)	(0.0482)	(0.0143)	(0.0572)
(d) Education *	−0.0539	−0.0952	−0.0665	−0.0889
experience/10	(0.0235)	(0.0276)	(0.0234)	(0.0283)
(e) Standardized AFQT *		0.1407		0.0913
experience/10		(0.0514)		(0.0627)
(f) Black * experience/10			−0.2670	−0.1739
			(0.0968)	(0.1184)
R^2	0.3056	0.3063	0.3061	0.3064

Experience is modeled with a cubic polynomial. All equations control for year effects, education interacted with a cubic time trend, Black interacted with a cubic time trend, AFQT interacted with a cubic time trend, two-digit occupation at first job, and urban residence. For these time trends, the base year is 1992. For the model in Panel 1 column (1) the coefficient on AFQT and Black are .0312 and −.1006, respectively, when evaluated for 1983. In Panel 2 the instrumental variables are the corresponding terms involving potential experience and the other variables in the model. Standard errors are White/Huber standard errors computed accounting for the fact that there are multiple observations for each worker. The sample size is 21,058 observations from 2976 individuals.

coefficient on education $\times$ $t/10$ is $-.0032$ (.0094), suggesting that the effect of education on wages declines very slightly with experience. As had been well documented, AFQT has a powerful association with earnings even after controlling for education. Since AFQT is normalized to have a standard deviation of 1, the estimates imply that a one-standard-deviation increase is associated with an increase in the log wage of .0834.

In column (2) we add linear interactions between t and a z variable, AFQT, to the equation. The resulting equation corresponds to (4) with the restriction that $b_{st} = b_{s0} + b_{s1} \times t$ and $b_{zt} = b_{z0} + b_{z1} \times t$. The coefficients of $-.0060$ (.0360) on AFQT and the coefficient of .0752 (.0286) on AFQT $\times$ $t/10$ imply that the effect of a one-standard-deviation shift in AFQT rises from essentially 0 when experience is 0 to .0692 when experience is 10. Our result for AFQT supports the hypothesis that employers learn about productivity. It is consistent with FG'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 fourteen that are orthogonal to the wage on the first job and education.

The key result in the table relating to statistical discrimination is that the coefficient on education $\times$ $t/10$ declines sharply to $-.0234$ (.0123) when AFQT $\times$ $t/10$ is added between columns (1) and (2). The implied effect of an extra year of education declines from .0829 (.0150) to .0595 (.0071) during the first ten years in the labor market. These results provide support for the hypothesis that employers have limited information about the productivity of labor force entrants and statistically discriminate on the basis of education. 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. When we condition the experience profile of earnings on both an easy-to-observe variable, such as education, and a hard-to-observe variable, such as AFQT, we find the partial effect of the easy-to-observe variables declines substantially with experience. While one might argue that the positive coefficient on AFQT $\times$ $t/10$ is due to an association between this variable and training intensity, it is hard to reconcile this view with the negative coefficient on education $\times$ $t/10$. While measurement error in schooling may enhance the effect of AFQT and may partially explain the decline in the magnitude of the coefficient on education $\times$ $t/10$ between columns (1) and (2), it does not provide a simple explanation for the signs of the interaction terms with experience.

In Panel 2 of Table I we report two-stage least squares estimates using actual experience as the experience measure t. We treat all terms involving actual experience as endogenous and use the corresponding terms involving potential experience as the instruments.[21] The results are basically consistent with those using potential experience. In Panel 2, column (2) the coefficient on AFQT is $-.0361$ (.0482), and the coefficient on AFQT $\times$ $t/10$ is .1407 (.0514). These estimates imply that conditional on years of schooling, AFQT has only a small effect on initial wages, but when t is 10, a one-standard-deviation shift in AFQT is associated with a wage differential of .1046. The coefficient on education $\times$ $t/10$ declines from $-.0539$ (.0235) when the interactions are excluded in column (1) to $-.0952$ (.0276) in column (2), a swing of $-.0413$. The substantial negative coefficient on education $\times$ $t/10$ in column (1) is disconcerting but is much smaller when calendar time interactions are excluded. (Results are not reported.)

While these results give general support for Proposition 1, we may want to know whether the experience profiles of the education and AFQT coefficients satisfy Proposition 2. One complication in performing these tests is the place of race within our model—should we treat race as an s variable or a z variable? The answer hinges on the extent to which employers violate the law and use race as an indicator of productivity. We discuss this at length in Section V below. For now we will sidestep the issue by running separate tests on the white and black samples. Proposition 2 says that the product of $-\text{cov}\,(s,z)/\text{var}\,(s)$—the negative of the coefficient of the regression of z on s—times the coefficient on the interaction between AFQT and experience ($z \times t$) should equal the coefficient on the interaction between education and experience ($s \times t$). In the white sample, the product is $-.0005$, and the coefficient on $s \times t$ is $-.0014$. In the black sample the corresponding numbers are $-.0040$ and $-.0049$. These numbers

21. The results based on potential experience are biased as estimates of the effect of actual experience. We instrument actual work experience because the intensity of work experience may be conveying information to employers about worker quality. It is an outcome measure itself. The implications of employer learning for the wage equation are changed if one conditions on information that becomes available to employers as the worker's career unfolds. When we treat actual experience as exogenous, we obtain a positive interaction between schooling and experience, but the impact of adding $z \times t$ to the models is similar to the pattern in Table I. See AP [1997, Table 2].

are very close, and in both samples a Wald test fails to reject Proposition 2.

IV.2. The Sibling Wage and Father's Education as z Variables

In columns (1) and (2) of Table II we use the log wage of siblings with five to eight years of experience as a hard-to-observe background characteristic. The coefficient on education $\times$ $t/10$ is .0107 (.0131) in column (1), which includes the log wage of the oldest sibling. When we add sibling wage $\times$ $t/10$ in column (2), the coefficient on the education interaction falls to .0012 (.0136), and the coefficient on the interaction between the sibling wage and $t/10$ is .1796 (.0749). The effect of the sibling wage rises from $-.0260$ (.0913) upon labor force entry to .1536 (.0345) after ten years of experience—a very large increase. Our interpretation of these results begins with the premise that the labor market productivities of siblings are correlated. As a worker acquires experience, this correlation is reflected in the performance record D_t and in wage rates. The sibling wage is positively correlated with education, and so the effect of education on the wage declines with experience because firms are estimating productivity with a bigger information set than at the time of labor force entry.

In models (5)–(8) of the table we replace the sibling wage with father's education. The effect of father's education also increases with experience. The main effect of father's education is actually slightly negative, and the experience interaction term is positive. Adding father's education $\times$ $t/10$ to the model leads to a reduction in the coefficient on education $\times$ $t/10$ from .0023 (.0104) to $-.0029$ (.0113). Consequently, the results for father's education conform to the predictions of the model, but none of the coefficients are statistically significant. We obtain much stronger results for father's education when calendar time interactions with father's education are excluded (see AP [1997]).

In Table III we simultaneously include AFQT, father's education, and the sibling wage rate in the same model. The interactions with experience of all three variables are positive, and, in the case of AFQT $\times$ $t/10$ and sibling wage $\times$ $t/10$, large and statistically significant. The coefficient on education $\times$ $t/10$ declines from .0005 (.0093) to $-.0269$ (.0123) when the interactions of the z variables are added. We tested the vector analog of Proposition 2 on models that include AFQT, father's education, and the sibling wage. We also considered as z variables the dummy variables indicating whether these quantities were

TABLE II

THE EFFECTS OF FATHER'S EDUCATION, SIBLING WAGES, AND SCHOOLING ON WAGES
Dependent Variable: Log Wage; Experience Measure: Potential Experience.
OLS estimates (standard errors)

Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(a) Education	0.0511	0.0630	0.0568	0.0659	0.0666	0.0730	0.0704	0.0734
	(0.0160)	(0.0166)	(0.0163)	(0.0167)	(0.0129)	(0.0140)	(0.0130)	(0.0140)
(b) Black	−0.2074	−0.2076	−0.0509	−0.0878	−0.2212	−0.2209	−0.0705	−0.0793
	(0.0276)	(0.0276)	(0.0846)	(0.0871)	(0.0250)	(0.0250)	(0.0668)	(0.0692)
(c) Log of sibling's wage	0.1802	−0.0260	0.1817	0.0010				
	(0.0328)	(0.0913)	(0.0329)	(0.0940)				
(d) Father's education/10					0.0826	−0.0187	0.0829	0.0314
					(0.0366)	(0.1000)	(0.0364)	(0.1030)
(e) Education * experience/10	0.0107	0.0012	0.0065	−0.0008	0.0023	−0.0029	−0.0002	−0.0027
	(0.0131)	(0.0136)	(0.0133)	(0.0136)	(0.0104)	(0.0113)	(0.0105)	(0.0113)
(f) Log of sibling's wage * experience/10		0.1796		0.1571				
		(0.0749)		(0.0770)				
(g) Father's education * experience/100						0.0867		0.0441
						(0.0813)		(0.0841)
(h) Black * experience/10			−0.1311	−0.1004			−0.1270	−0.1194
			(0.0686)	(0.0704)			(0.0541)	(0.0563)
R^2	0.3183	0.3196	0.3191	0.3200	0.2748	0.2750	0.2755	0.2756
Observations	10746	10746	10746	10746	18523	18523	18523	18523
Individuals	1441	1441	1441	1441	2594	2594	2594	2594

Experience is modeled with a cubic polynomial. All equations control for year effects, education interacted with a cubic time trend, Black interacted with a cubic time trend, two-digit occupation at first job, and urban residence. Columns (1)–(4) control for sibling's gender and the log of sibling's wage interacted with a cubic time trend. Columns (5)–(8) control for father's education interacted with a cubic time trend. For these time trends, the base year is 1992. For the models in columns (1) and (5), the coefficients on log of sibling wage and father's education are .1680 and .0357, respectively, when evaluated for 1983. Standard errors are White/Huber standard errors computed accounting for the fact that there are multiple observations for each worker.

TABLE III
THE EFFECTS OF STANDARDIZED AFQT, FATHER'S EDUCATION, SIBLING WAGE, AND SCHOOLING ON WAGES
Dependent Variable: Log Wage; Experience Measure: Potential Experience.
OLS estimates (standard errors)

Model:	(1)	(2)	(3)	(4)
(a) Education	0.0505	0.0832	0.0563	0.0780
	(0.0118)	(0.0151)	(0.0120)	(0.0155)
(b) Black	−0.1333	−0.1296	0.0454	−0.0284
	(0.0255)	(0.0257)	(0.0609)	(0.0704)
(c) Standardized AFQT	0.0792	−0.0206	0.0789	0.0065
	(0.0145)	(0.0361)	(0.0144)	(0.0413)
(d) Log of sibling's wage	0.1602	0.0560	0.1617	0.0604
	(0.0208)	(0.0352)	(0.0207)	(0.0351)
(e) Father's education/10	0.0362	0.0154	0.0385	0.0295
	(0.0356)	(0.0963)	(0.0354)	(0.0968)
(f) Education * experience/10	0.0005	−0.0269	−0.0035	−0.0220
	(0.0093)	(0.0123)	(0.0094)	(0.0128)
(g) Standardized AFQT * experience/10		0.0843		0.0614
		(0.0285)		(0.0333)
(h) Log of sibling wage * experience/10		0.1194		0.1151
		(0.0393)		(0.0393)
(i) Father's education * experience/100		0.0176		0.0055
		(0.0789)		(0.0794)
(j) Black * experience/10			−0.1500	−0.0861
			(0.0474)	(0.0570)
R^2	0.2991	0.3014	0.3002	0.3016

Experience is modeled with a cubic polynomial. All equations control for year effects, education interacted with a cubic time trend, Black interacted with a cubic time trend, AFQT interacted with a cubic time trend, father's education interacted with a cubic time trend, sibling wage interacted with a cubic time trend, two-digit occupation at first job, and urban residence. Also included are sibling's gender and dummy variables to control for whether father's education is missing and whether sibling's wage is missing, and interactions between these dummy variables and experience when experience interactions are included. Standard errors are White/Huber standard errors computed accounting for the fact that there are multiple observations for each worker. The sample size is 21,058 observations from 2976 individuals.

known. This test amounts to a t-test of whether the sum of the products of $-\text{cov}\,(s,z)/\text{var}\,(s)$ and the coefficient on $z \times t$ for each z variable is equal to the coefficient on $s \times t$. For whites, the sum of the products equals $-.0021$ and the coefficient on $s \times t$ is $-.0020$. For blacks, we obtain $-.0042$ and $-.0049$. In both cases we fail to reject the proposition.

IV.3. *The Experience Profile of the Effects of AFQT and Education on Wages*

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 if some new information arrives each period on y. If the noise in observations of y_t are iid, then the rate of increase $\partial^2 w_t/\partial AFQT, \partial t$ should decline with t, as shown in expression (14) for θ_t above. To investigate this, we replaced the linear interactions between education and t and $AFQT$ and t in column (2) of Table I with quartic interactions. $\partial w_t/\partial AFQT$ increases steadily from $-.0025$ (.0400) when t is 0 to .0874 (.0796) when t is 12. The values of $\partial^2 w_t/\partial AFQT, \partial t$ increase from .0038 (.0159) when t is 0 to .0082 (.0192) when t is 5, to .0089 (.0210) when t is 8, and then decline to .0064 (.0234) when t is 12. These estimates suggest that the flow of new information is relatively constant after an initial period of noisy observations, but they are too imprecise for us to draw conclusions.

V. Do Employers Statistically Discriminate on the Basis of Race?

As we discussed in Section II, a statistically discriminating firm might use race along with education and other information to predict the productivity of new workers. With experience, the productivity of the worker would become apparent, and compensation would be based on all the information available rather than just the information available at the time of hire. Consequently, if statistical discrimination on the basis of race is important, then adding interactions between t and z variables such as AFQT and father's education to the wage equations should lead to a positive (or less negative) coefficient on Black $\times$ $t/10$ and should lead to an increase in the race intercept. As noted in Section II, if firms use race as information, then Black behaves as an s variable in the model, and the logic is the same as in our analysis of the effect of education. On the other hand, if firms do not use or only partially use race as information, then Black behaves as a z variable. In this case the race intercept when experience is 0 will be smaller than when firms use race to discriminate. The gap should widen with experience if race is negatively related to productivity, and adding a second z variable that is negatively related to race will reduce the race gap in experience slopes and possibly make the race intercept more negative.[22]

The race differential in our basic specification in column (1) of

22. The learning model in Section II implies that differences across groups in the association between s and the z variable will lead to group differences in the

Table I is −.1565 (.0256).[23] When Black × $t/10$ is added in column (3), it enters with a coefficient of −.1315 (.0482). The coefficient on Black in column (3) is .0001, although the standard error is large, (.0621). The hypothesis that firms do not statistically discriminate on the basis of race does not imply that coefficient on Black will be 0, since race may be correlated with information in q that can legally be used. It does imply, however, that the coefficient will be smaller when firms do not use race to discriminate than when they do. The fact that the race gap when t equals 0 is essentially 0 and that the gap rises sharply with experience is consistent with the hypothesis of no or very limited statistical discrimination on the basis of race. It is inconsistent with the hypothesis that firms make full use of race as information. The fact that the coefficient on Black × $t/10$ rises to −.0834 (.0581) when AFQT × t is added to the equation (column (4)) is not informative about whether or not firms make full use of race as information.

We obtain similar results using actual experience measures in Panel 2 columns (1), (3), and (4) of Table I. In Table II we obtain qualitatively similar but less dramatic results when we use the sibling wage or father's education as the z variable. Finally, in Table III we obtain results that are similar to those in Table I when we simultaneously use AFQT, father's education, and the sibling wage as z variables. When the interactions between these variables are excluded from the model, the coefficient on Black is .0454 (.0609), and the coefficient on Black × $t/10$ is −.1500 (.0474). The latter coefficient declines to −.0861 (.0570) when interactions between t and AFQT, father's education, and the sibling wage are introduced.

We wish to stress that the simple model of statistical discrimination cannot explain the large negative coefficient on Black × t unless firms do not make full use of race as information. The fact that the race gap is so small at low experience levels suggests either that there is not much difference in the productivity of black and white men at the time of labor force entry or that firms do not statistically discriminate very much. The accu-

b_{st} and b_{zt} coefficients. We have not explored this empirically, in part because the results might be sensitive to the linearity assumptions that we have made.

23. It should be kept in mind that this estimate refers to the race gap in 1992, conditional on AFQT, education, potential labor market experience, and two-digit occupation of the first job after leaving school for the first time. The coefficient on Black is −.2362 (.0214) when AFQT is excluded from the model in column (1).

mulation of additional information during a career that can legally be used to differentiate among workers would imply a widening of the race gap with experience (again, if there is a productivity gap) and is fully consistent with our results. However, there are other discrimination-related explanations of the race differences in the experience slope that may be at work here, as we emphasize in subsection III.3. It is also important to point out that the coefficients on Black and Black $\times$ $t/10$ alone (i.e., ignoring the behavior of the coefficients on education and education $\times$ t) are potentially consistent with a story in which firms are fully informed, AFQT is positively associated with on-the-job training, and the race difference in AFQT is partially responsible for a race differential in wage growth. Adding AFQT $\times$ t would reduce a negative bias in Black $\times$ t associated with differential training levels. The increase in Black $\times$ t when AFQT $\times$ t is added to the model would lead to a fall in the coefficient on Black. As we report below, we obtain qualitatively similar results when we add controls for employer training, but these controls reduce the magnitude of the coefficient on Black $\times$ t and the effect of adding AFQT $\times$ t on the coefficient on Black $\times$ t.

Another potential test of whether race is used to statistically discriminate or not is to see whether Proposition 2 holds either when race is treated as an s variable or when it is treated as a z variable. To do this, we use the model in column (4) of Table III. With race treated as an s variable, we regress the z variables (AFQT, the log of sibling's wage, father's education, and the dummies for not knowing these quantities) on the two s variables. We sum the product of these coefficients and the coefficients on the z $\times$ t interactions in the main regression and compare them with the coefficients on the s $\times$ t interactions. We can then conduct a joint test of whether these two quantities are equal. For the education interactions the sum of the products equals $-.0024$ while the model coefficient is $-.0022$. For the race interaction, the two terms have opposite signs; the sum is .0088 while the model coefficient is $-.0086$. Not surprisingly, the proposition is soundly rejected.

When we treat race as a z variable, we begin our test by regressing the six z variables on education, our s variable. Here, we have only one restriction to test. The sum of the products equals $-.0028$ while the model coefficient equals $-.0022$. The proposition cannot be rejected, providing further evidence that employers are not using race as information, or at least not fully.

VI. Models with Training

In Table IV we report estimates of equation (8). The model in column (1) is the same as the model in Table III, column (1), but with current training R_t and cumulative training $\Sigma\ R_\tau$ added. There are two problems in using the training data. First, the measure R_t^* of R_t is almost certain to contain measurement error. Second, the quality of the training data prior to 1988 is too poor to be used, which means that the data required for $\Sigma\ R_\tau$ are missing for persons who left school prior to that year. We do not have a solution for the first problem. We deal with the second problem by estimating a flexible model relating R_t^* to s, z, and t using data from 1988–1992 and then using this model to impute values in the earlier years.[24] We estimate (8) in first differences as well as in levels. The first difference specification exacerbates measurement error but has the advantage of only requiring data on R_t and R_{t-1} and eliminates bias from unobserved person-specific effects that are known to firms (q) and are correlated with both training and wages.

Adding the training measures to the models in Table III leads to only slight changes in the coefficients on education, AFQT, sibling's wage, and father's education. The variable R_t has the expected negative sign of $-.1143$ (.0200), while $\Sigma\ R_\tau$ has a coefficient of .1881 (.0139). Adding the training leads to a decrease in the coefficient on education $\times\ t/10$ from essentially 0 (Table III, column (1)) to $-.0231$ (.0095). The substantial negative experience slope on education is consistent with a human capital story in which knowledge obtained in school depreciates over time unless one receives training. However, it is also consistent with a model in which the correlation between cumulative training and employer beliefs about productivity grows stronger with experience, inducing a decline in the education coefficient because education has a strong positive correlation with cumulative training. In column (2) we add the interactions between the z variables and experience t. The coefficient on education $\times\ t/10$ drops from $-.0231$ (.0095) to $-.0392$ (.0123), and $AFQT \times t/10$ and the

24. Spletzer and Lowenstein [1996] provide means of dealing with measurement error in the training data but these are beyond the scope of our study.

TABLE IV
THE EFFECTS OF STANDARDIZED AFQT, FATHER'S EDUCATION, SIBLING WAGE, SCHOOLING, AND TRAINING ON WAGES
Dependent Variable: Log Wage; Experience Measure: Potential Experience.
Training Measure: Predicted before 88, Actual After.
OLS estimates (standard errors)

Model:	(1)	(2)	(3)	(4)
(a) Education	0.0606	0.0802	0.0651	0.0746
	(0.0119)	(0.0151)	(0.0121)	(0.0155)
(b) Black	−0.1159	−0.1135	0.0241	−0.0028
	(0.0265)	(0.0267)	(0.0616)	(0.0722)
(c) Standardized AFQT	0.0334	−0.0199	0.0338	0.0102
	(0.0150)	(0.0363)	(0.0150)	(0.0420)
(d) Log of sibling's wage	0.1594	0.0716	0.1611	0.0759
	(0.0213)	(0.0357)	(0.0213)	(0.0356)
(e) Father's education/10	0.0460	0.0211	0.0482	0.0353
	(0.0356)	(0.0974)	(0.0354)	(0.0977)
(f) Education *	−0.0231	−0.0392	−0.0260	−0.0339
experience/10	(0.0095)	(0.0123)	(0.0096)	(0.0128)
(g) Standardized AFQT *		0.0460		0.0207
experience/10		(0.0287)		(0.0339)
(h) Log of sibling's wage *		0.1041		0.1001
experience/10		(0.0402)		(0.0402)
(i) Father's education *		0.0205		0.0084
experience/100		(0.0803)		(0.0805)
(j) Black * experience/10			−0.1180	−0.0945
			(0.0476)	(0.0583)
(k) Training: R_t	−0.1143	−0.1095	−0.1115	−0.1091
	(0.0200)	(0.0199)	(0.0199)	(0.0199)
(l) Cumulative training: Σ	0.1881	0.1830	0.1854	0.1827
R_T	(0.0139)	(0.0139)	(0.0139)	(0.0139)
R^2	0.3188	0.3199	0.3195	0.3202

Experience is modeled with a cubic polynomial. All equations control for year effects, education interacted with a cubic time trend, Black interacted with a cubic time trend, AFQT interacted with a cubic time trend, father's education interacted with a cubic time trend, sibling wage interacted with a cubic time trend, two-digit occupation at first job, and urban residence. Also included are dummy variables to control for whether father's education is missing and whether sibling's wage is missing, and interactions between these dummy variables and experience when experience interactions are included. For these time trends, the base year is 1992. R_t is the predicted probability of training in year t if before 1988 and actual training if year t is after 1987. Predictions are based on a probit model containing years of schooling, potential experience, Black, AFQT, schooling times potential experience and potential experience squared, AFQT times potential experience and potential experience squared, and the product of AFQT, schooling, and potential experience. Cumulative training is aggregated over the individual's entire career, using estimated training before 1988 and actual training thereafter. Standard errors are White/Huber standard errors computed accounting for the fact that there are multiple observations for each worker. The sample size is 19,785 observations from 2912 individuals.

sibling wage $\times$ $t/10$ enter with coefficients of .0460 (.0287) and .1041 (.0402), respectively. These changes are consistent with EL-SD. If we reverse the order in which the variables are added

TABLE V

ESTIMATES OF THE EFFECTS OF AFQT, FATHER'S EDUCATION, SIBLING WAGE, AND SCHOOLING ON WAGE GROWTH WITH CONTROLS FOR TRAINING

Dependent Variable: Δ log Wage; Experience Measure: Potential Experience. Coefficient estimates (standard errors)

Model:	(1)	(2)	(3)	(4)
Education *	−0.0060	−0.0694	−0.0106	−0.0729
Δexperience/10	(0.0833)	(0.0960)	(0.0832)	(0.0959)
AFQT * Δexperience/10		0.3025		0.2975
		(0.1613)		(0.1614)
Log of sibling wage *		0.2153		0.2107
Δexperience/10		(0.1477)		(0.1477)
Father's education *		−0.4306		−0.4215
Δexperience/10		(0.5034)		(0.5034)
Black * Δexperience/10	−0.0504	−0.0425	−0.0503	−0.0426
	(0.0484)	(0.0485)	(0.0483)	(0.0484)
Training: $R_t/10$			0.2468	0.2429
			(0.1024)	(0.1025)
Lag training: $R_{t-1}/10$			−0.0194	−0.0230
			(0.1108)	(0.1108)
S.E.E.	.2965	.2965	.2965	.2964

All equations control for year effects education, AFQT, sibling wage, and father's education all interacted with the change in the square and cube of time, the change in urban residence, and dummy variables to control for whether father's education is missing and whether AFQT is missing, and interactions between these dummy variables and the change in experience when change in experience interactions are included. Standard errors are White/Huber standard errors computed accounting for the fact that there are multiple observations for each worker. The sample size is 14,938 observations from 2703 individuals.

by adding AFQT $\times$ $t/10$ before the training measures, the marginal effect of the training measures on education $\times$ t is much smaller.

In columns (3) and (4) we investigate the effect of introducing the training measure on the race gap in wage slopes. These columns correspond to columns (3) and (4) in Table III with R_t and cumulative training $\Sigma R_{i\tau}$ added. The coefficient on Black $\times$ $t/10$ declines from −.1500 (.0474) (Table III, column (3)) to −.1180 (.0476) when we add the training measures. Adding the experience interactions with the z variables leads to a further decline to −.0945 (.0583).

To reduce the difficulties associated with the lack of data on training in the early years of the study and individual heterogeneity that is correlated with both training and wages, we turn to a first differenced wage model. In the first differenced version R_t and its lag R_{t-1} enter. These results are in Table V. The coefficient on education $\times$ $\Delta t/10$ declines from −.0060 (.0833) to

−.0106 when the training measures are added. The coefficient on Black × $\Delta t/10$ rises very slightly from −.0505 (.0484) to −.0503 (.0483). However, the coefficient on R_t is large and positive while the coefficient on R_{t-1} is small and negative. These signs are inconsistent with a simple human capital model but are consistent with an EL-SD model in which training opportunities are given to more productive workers, and learning about productivity occurs over time.[25] Adding the training variables to a model that contains $AFQT \times \Delta t/10$, log sibling wage × $\Delta t/10$, and father's education × $\Delta t/10$ has little impact on the coefficients on these variables. (Compare columns (2) and (4).) Imprecision in the training measures may partially explain this fact, but does not provide an explanation for the size and the sign pattern of the training coefficients. The coefficients on education × $\Delta t/10$ and Black × $\Delta t/10$ decline in absolute value when the z variables interacted with Δt are added, as is predicted by the EL-SD. The wage change results are quite consistent with an important role for EL-SD.

Overall, the evidence incorporating training suggests a role for both human capital and EL-SD. In view of the econometric problems and very serious data problems discussed above, we cannot make a precise statement about the relative importance of these two factors.

VII. Conclusions and a Research Agenda

This paper provides a way to test for statistical discrimination based on the premise that firms use the information they have available to form judgments about the productivity of workers and then revise these beliefs as additional information becomes available. We show that as firms acquire more information about a worker, pay will become more dependent on productivity and less dependent on easily observable characteristics or credentials. This basic proposition is quite general and provides a way to test for statistical discrimination in

25. In the EL-SD model the component of R_t that reflects new information about the workers will induce a positive sign on R_t. In this model R_{t-1} enters with a negative coefficient because it is positively correlated with the component of R_t that is not new information. Note that in the wage growth equations the coefficients on the interactions of a variable with $\Delta t/10$ coefficients are the combined effect of the interaction between the variable and experience and the variable and the linear term in a cubic secular time trend.

the labor market and elsewhere in situations in which agents learn, such as credit markets.

We investigate it empirically by estimating a wage equation that contains interactions between experience and hard-to-observe correlates of productivity such as AFQT, the wage of a sibling, and father's education, and between experience and more easily observed characteristics such as years of education. We find that the wage effects of the unobservable productivity variables rise with time in the labor market and the wage effect of education falls. These results match the predictions of our model of statistical discrimination with employer learning.

We use a similar methodology to investigate whether employers statistically discriminate on the basis of race. If our model is taken literally, the small race differentials for new workers and the widening of the race gap with experience is most consistent with the view that race is negatively correlated with productivity and the productivity gap becomes reflected in wages as firms acquire additional information that can legally be used to differentiate among workers. We wish to stress, however, that other factors are probably as or more important in differences between whites and blacks in wage profiles and that race differences in human capital accumulation account for at least part of our findings.

We feel that this study has broad applicability to many areas of labor economics and hope that it will lead to more research in a number of areas:

— Studies of statistical discrimination on the basis of other easily observable characteristics such as gender, country of origin, neighborhood, and rank of college or professional school attended.

— The incorporation of additional "hard-to-observe measures," particularly those related to noncognitive skills, effort levels, and labor force attachment.

— A reinterpretation of previous studies of wage determination containing interactions between experience and productivity correlates of different degrees of observability.

— An analysis of information and price determination in markets where a measure of productivity may be available, such as matched firm-worker studies or mortgage lending studies.

— Consideration of models in which employer information is private, and distinguishing between learning with experience and learning with firm seniority.[26]
— An examination of the effect on group differences in wage dynamics of group differences in the accuracy of information firms have.
— An inquiry into why methods to determine hard-to-observe correlates of productivity (e.g., testing using the AFQT) are not widely used by firms given their economic value.[27]

Appendix 1: Proof That Adding a Second z Variable Reduces the Race Gap in Experience Profiles

Let z_1 denote black and z_2 denote a second z variable. Assume that z_2 is a favorable characteristic (such as AFQT) in the sense that it has a positive coefficient in a wage equation. Let $b^*_{z_1t}$ be the coefficient on z_1 when $w_t - w_0$ is regressed on s and z_1. Let b_{z_1t} be the regression coefficient on z_1 when $w_t - w_0$ is regressed on s, z_1, and z_2. Assume that $\theta_{z_1t} = \theta_{z_2t} = \theta_t$, where θ_{z_kt} is defined in (7) above with z_k substituted for z. This assumption means that the rates at which firms learn about the productivity components associated with each z are equivalent. Note that $0 \leq \theta_t \leq 1$ and $\partial\theta_t/\partial t \geq 0$. Then

$$b^*_{z_1t} = \Phi^*_{z_1}\theta_t$$

$$b_{z_1t} = \Phi_{z_1}\theta_t,$$

where $\Phi^*_{z_1}$ is the coefficient on z_1 in the regression of $\Lambda v + e$ on s and z_1, and Φ_{z_1} is the coefficient on z_1 in the regression of $\Lambda v + e$ on s, z_1, and z_2. From the omitted variables formula, we know that $\Phi^*_{z_1} = \Phi_{z_1} + \Phi_{z_2}\Phi_{z_2z_1}$, where Φ_{z_2} is the coefficient on z_2 in the regression of $\Lambda v + e$ on s, z_1, and z_2 and $\Phi_{z_2z_1}$ is the coefficient on z_1 in the regression of z_2 on z_1 and s. Since z_2 is a favorable characteristic, $\Phi_{z_2} > 0$. Assume that z_2 is negatively related to z_1 given s ($\Phi_{z_2z_1} < 0$):

26. Key references include Greenwald [1986], Waldman [1984], Lazear [1986], and Gibbons and Katz [1991]. In AP [1997] we present some very preliminary evidence that hard-to-observe variables like AFQT, father's education, and the wage of an older sibling are positively related to the layoff probability but have only a weak relationship with quits. We did not find much evidence that these variables are negatively related to wage growth conditional on a layoff and positively related to wage growth in the case of quits, as some private information models imply. Our results suggest that information flows in the labor market are sufficient to force a firm to differentiate among workers as the firm obtains better information about their productivity.

27. In AP [1997], using plausible assumptions about how fast employers learn about employee productivity, we estimate that a person who believes that he is one standard deviation above the mean for the AFQT would be willing to pay a substantial fraction of his first year salary to take the test.

$$\frac{\partial b_{z_1t}}{\partial t} = \Phi_{z_1}\frac{\partial \theta_t}{\partial t}, \text{ and}$$

$$\frac{\partial b^*_{z_1t}}{\partial t} = \Phi^*_{z_1}\frac{\partial \theta_t}{\partial t} = (\Phi_{z_1} + \Phi_{z_2}\Phi_{z_2z_1})\frac{\partial \theta_t}{\partial t}.$$

Taking the difference leads to

$$\frac{\partial b_{z_1t}}{\partial t} - \frac{\partial b^*_{z_1t}}{\partial t} = -\frac{\partial \theta_t}{\partial t}\cdot[\Phi_{z_2}\Phi_{z_2z_1}] \geq 0.$$

APPENDIX 2

The National Longitudinal Survey of Youth 1979 (NLSY79) is comprised of 12,686 respondents who were born between January 1, 1957, and December 31, 1964. Our study focuses on the 5403 non-Hispanic males in the NLSY79. When first interviewed in 1979, these youths were between 14 and 22 years of age. We use data through the 1992 wave of the survey at which point respondents ranged in age from 27 to 35.

We limit our analysis to jobs after a person leaves school for the first time. The first time a person leaves school is the month and year of the most recent enrollment at the first interview where the respondent is not currently enrolled in school. We calculate actual experience as the cumulative number of weeks in which the respondent, after leaving school for the first time, has worked 30 or more hours a week (divided by 50 in order to approximate years of experience).[28] If he returns to school, valid jobs are still included in our analysis, and any experience that meets the 30 hours per week rule is accumulated.

We consider only employment for the current or most recent employer (the CPS job) and only if the respondent is working at the job in the interview week. If the respondent is holding two jobs at the time of the interview, only the job with the greatest number of hours worked is considered. Both full- and part-time jobs are used. Military jobs are excluded from the wage analysis and the accumulation of work experience. We include all valid data for all individuals, including those who fail to respond in certain years or eventually leave the NLSY79 sample due to the elimination of certain subsamples (military, economically disad-

28. We looked at other cutoff points (ten or twenty hours per week) and other measures (total hours or total years working at least 1500 hours) but found the initial results insensitive to the definition of actual experience.

vantaged whites). Since the work history format picks up all employment activity since the previous interview, failure to respond in one year does not necessitate dropping respondents if they return for subsequent interviews. Our wage measure is the hourly wage from the work history file. We divide by the 1987 fixed-weighted price index for GNP personal consumption expenditures to obtain real wages. Observations where the real wage is below $2.00 or above $100.00 are eliminated from the analysis.

Within the NLSY79 sample, there are 5403 non-Hispanic males. We exclude 120 individuals who never left school by their 1992 interview (or their last previous interview if they were nonrespondents in 1992). Among the remaining 5283 respondents, 3783 first left school between 1978 and 1992. For this group we could calculate experience for all but 162 individuals using the work history file. For the 1500 who left school before 1978, we construct work history prior to 1978 using three sets of questions from the 1979 survey. The first set asks about a respondent's first job after leaving school including starting and stopping dates. The second set asks about military service. If a respondent reported being in full-time military service, we assumed he was not employed. The third set asks the number of weeks and hours per week that a respondent worked in 1977, 1976, and 1975. For the 695 individuals for whom the number of weeks we could not account for was five or less, we calculated experience using the data we had available. We dropped 805 possible respondents who left school before 1978 for whom we could not completely determine their work history.

Subsequent to the calculation of experience, 1340 individuals were eliminated from our sample: 121 who did not have eight years of education, 124 who had no valid jobs or wages, 936 who had no first occupation, 130 who had no AFQT score, and 29 who were missing other variables. Our wage analysis sample contains 2976 individuals.

As one would expect, the less educated were more likely to be dropped from the analysis, especially among the oldest cohorts. A youth born in 1957, if he attended school continuously, would normally graduate from high school in 1975 and from college in 1979. It is harder to track a person who stopped his education after high school four years prior to the survey than one who continued on to college. Table VI shows, by birth year, the number of non-Hispanic, male NLSY79 respondents overall and in our sample, and their education level when they first appear in our sample. As expected,

TABLE VI
NLSY79 NON-HISPANIC MALE SAMPLE USED IN ANALYSIS, BY BIRTH YEAR

Birth year	Number in NLSY79	Number in sample	Percent in sample	Years of education	Ave. number of obs.
57	743	230	31.0%	14.26	7.19
58	740	272	36.8%	13.93	7.12
59	715	328	45.9%	12.88	7.42
60	711	433	60.9%	12.77	7.88
61	615	416	67.6%	12.80	7.58
62	690	494	71.6%	12.63	6.97
63	646	436	67.5%	12.36	6.59
64	541	366	67.7%	12.57	5.86
65	2	1	50.0%	15.00	9.00
Total	5403	2976	55.1%	12.90	7.08

sample members born in 1957 have on average a whole year more education than cohorts born after 1960 do.

Our education variable is simply the number of grades completed with a maximum of 20. We assume that the initially reported education level is correct and require that educational attainment remain constant or increase after that point. Thus, education is nondecreasing over time for each sample member. Those with education levels below eighth grade were eliminated from the analysis. Likewise, father's education is measured in years, with reports below four years set to 4.

Because the age of the sample members at the time the AFQT was administered varies somewhat in the NLSY79 sample, AFQT scores are standardized to account for the difference in schooling levels across ages. To calculate standardized AFQT, we adjust the raw AFQT score by subtracting the mean score for a person of that age and dividing by the standard deviation for that age. For individuals with siblings in the sample, the coefficients of the regression of the unadjusted test score of the older sibling on the test score of the younger sibling and of the regression of the test score of the younger sibling on the score of the older sibling are very similar after one also controls for age. This suggests that the information in the test is not very sensitive to age at the time of the test.

In the sibling analysis we use the oldest available sibling for whom we have a wage. The wage measure we use is an average wage over the period between the fifth and eighth year after the sibling has left school. Only 1881 of the 4042 individuals in the main analysis have a sibling with a valid wage and can be used in these models.

In our analysis we also include dummies for whether we know the wage of a sibling and whether we know father's education.

In our models that control for training, our training variable R_t is simply a dummy variable for whether the respondent received company training in the previous year. However, training data within the NLSY79 are weak before 1988. Until that year, training spells of less than four weeks were not counted. As such, the incidence of company training more than doubled between 1986 and 1988. Thus, we ignore the training data before 1988, and estimate the probability that one received company training during that time based on post-1987 data. Specifically, we run a probit of the receipt of company training in the years after 1987 on years of schooling, potential experience, Black, AFQT, and interactions between these variables. We then use actual data on these variables for the pre-1988 period to estimate the probability that one received company training in this period. Cumulative training is aggregated over the individual's entire career, using estimated training before 1988 and actual training thereafter.

Table VII contains descriptive statistics for observations used in the analysis.

TABLE VII
DESCRIPTIVE STATISTICS

Variable	Mean	Standard deviation	Minimum	Maximum
Real hourly wage	8.23	4.71	2.01	96.46
Log of real hourly wage (w)	1.99	0.47	0.70	4.57
Potential experience (t)	7.09	3.61	0	20.00
Actual experience (t)	4.74	3.39	0	14.92
Education (s)	12.75	2.13	8.00	18.00
Black dummy (Black)	0.29	0.45	0	1
Standardized AFQT Score (AFQT)	−0.14	1.04	−2.78	1.92
Do not know sibling's wage	0.49	0.50	0	1
Log of sibling's wage	1.95	0.45	0.73	3.47
Do not know father's education	0.12	0.33	0	1
Father's education	11.68	3.28	4.00	20.00
Training (R_t)	0.10	0.20	0	1
Cumulative Training: (ΣR_τ)	0.48	0.40	0	2.65

Sample size = 21,058 observations except for sibling wage (10,746 observations), father's education (18,523 observations), and the training measures (19,785 observations).

NORTHWESTERN UNIVERSITY AND NATIONAL BUREAU OF ECONOMIC RESEARCH
BUREAU OF LABOR STATISTICS

References

Aigner, D., and G. Cain, "Statistical Theories of Discrimination in the Labor Market," *Industrial and Labor Relations Review,* XXX (1977), 175–187.
Albrecht, J., "A Procedure for Testing the Signaling Hypothesis," *Journal of Public Economics,* XV (1981), 123–132.
Altonji, J. G., and R. M. Blank, "Race and Gender in the Labor Market," in *Handbook of Labor Economics, Volume 3C,* O. Ashenfelter and D. Card, eds. (Amsterdam: North-Holland, 1999), pp. 3143–3259.
Altonji, J. G., and C. R. Pierret, "Employer Learning and Statistical Discrimination," National Bureau of Economic Research Working Paper No. 6279, November 1997.
Altonji, J. G., and C. R. Pierret, "Employer Learning and the Signaling Value of Education," in *Internal Labour Markets, Incentives and Employment,* I. Ohashi and T. Tachibanaki, eds. (London: Macmillan Press Ltd., 1998), pp. 159–195.
Becker, G. S., *The Economics of Discrimination,* 2nd ed. (Chicago: University of Chicago Press, 1971).
Carmichael, H. L., "Self-Enforcing Contracts, Shirking, and Life Cycle Incentives," *Journal of Economic Perspectives,* III (1989), 65–84.
Cawley, J., J. Heckman, and E. Vytlacil, "Cognitive Ability and the Rising Return to Education," NBER Working Paper No. 6388, January 1998.
Chay, K. Y., and D. S. Lee, "Changes in Relative Wages in the 1980s: Returns to Observed and Unobserved Skills and Black-White Wage Differentials," unpublished paper, February 1998.
Coate, S., and G. Loury, "Will Affirmative Action Policies Eliminate Negative Stereotypes?" *American Economic Review,* LXXXIII (1993), 1220–1240.
Farber, H., and R. Gibbons, "Learning and Wage Dynamics," *Quarterly Journal of Economics,* CXI (1996), 1007–1047.
Foster, A. D., and M. R. Rosenzweig, "Information, Learning, and Wage Rates in Low Income Rural Areas," *Journal of Human Resources,* XXVIII (1993), 759–790.
Gibbons, R., and L. Katz, "Layoffs and Lemons," *Journal of Labor Economics,* IX (1991), 351–380.
Glaeser, E., "Who Knows What about Whom: An Empirical Analysis of Heterogeneous Information in the Labor Market," unpublished paper, University of Chicago, 1992.
Greenwald, B. C., "Adverse Selection in the Labour Market," *Review of Economic Studies,* LIII (1986), 325–347.
Holzer, H., "Search Methods Used by Unemployed Youth," *Journal of Labor Economics,* VI (1988), 1–20.
Katz, L. F., and D. H. Autor, "Changes in the Wage Structure and Earnings Inequality," forthcoming in *Handbook of Labor Economics,* Vol. *3A,* O. Ashenfelter and D. Card, eds. (Amsterdam: North-Holland, 1999). pp. 1463–1555.
Katz, L. F., and K. M. Murphy, "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *Quarterly Journal of Economics,* CVII (1992), 35–78.
Lang, K., "A Language Theory of Discrimination," *Quarterly Journal of Economics,* CI (1986), 363–382.
Lazear, E., "Raids and Offer Matching," in *Research in Labor Economics 8,* R. G. Ehrenberg, ed. (Greenwich, CT: JAI Press, 1986), pp. 141–165.
Lundberg, S., and R. Startz, "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review,* LXXIII (1983), 340–347.
Medoff, J., and K. Abraham, "Experience, Performance, and Earnings," *Quarterly Journal of Economics,* C (1980), 703–736.
Montgomery, J., "Social Networks and Labor Market Outcomes: Toward an Economic Analysis," *American Economic Review,* LXXXI (1991), 1408–1418.
Murphy, K., and F. Welch, "The Structure of Wages," *Quarterly Journal of Economics,* CVII (1992), 285–326.
Neal, D. A., and W. R. Johnson, "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy,* CIV (1996), 869–895.
Oettinger, G. S., "Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap," *Journal of Labor Economics,* XIV (1996), 52–78.
Parsons, D. O., "The Employment Relationship: Job Attachment, Worker Effort,

and the Nature of Contracts," in *Handbook of Labor Economics,* Vol. 2, O. Ashenfelter and R. Layard, eds. (Amsterdam: North-Holland, 1986), pp. 789–848.
——, "Reputational Bonding of Job Performance: The Wage Consequences of Being Fired," unpublished paper, July 1993.
Riley, J. G., "Testing the Educational Screening Hypothesis," *Journal of Political Economy,* LXXXIV (1979), S227–252.
Spence, A. M., "Job Market Signaling," *Quarterly Journal of Economics,* LXXXVII (1973), 355–379.
Spletzer, J., and M. Lowenstein, "Belated Training: The Relationship between Training, Tenure and Wages," unpublished paper, Bureau of Labor Statistics, 1996.
Taber, C. R., "The Rising College Premium in the Eighties: Return to College or Return to Ability," unpublished paper, Northwestern University, April 1996.
Waldman, M., "Job Assignments, Signalling, and Efficiency," *Rand Journal of Economics,* XV (1984), 255–267.
Weiss, A., "Human Capital and Sorting Models," *Journal of Economic Perspectives,* IX (1995), 133–154.