Post-Schooling Training Investment and Employer Learning

Barış Kaymak*
Université de Montréal

Abstract

Wage growth among young workers relates positively to ability and negatively to education, conditional on ability. This was interpreted as evidence for statistical discrimination with employer learning. I show that this pattern is also consistent with a version of the Ben-Porath model of skill formation where (i) workers differ in their learning ability and (ii) job training is a substitute for formal schooling. Data on job training from the NLSY confirm both modeling extensions for young workers. Nonetheless, the substitutability of training for schooling fades too quickly to explain the slow wage growth by education among workers with comparable ability.

J.E.L. Codes: J24, J31, J71
Keywords: On-the-job training, human capital, employer learning, statistical discrimination

*Department of Economics, Université de Montréal, C.P. 6128 succursale Centre-ville, Montréal, QC H3C 3J7. E-mail: baris.kaymak@umontreal.ca
Workers with higher cognitive ability, as measured by aptitude tests, enjoy faster wage growth throughout their careers. On the other hand, among workers with similar abilities, those with higher educational attainment have slower wage growth. This pattern is particularly strong for young workers and lasts for about 10-15 years.

These empirical regularities are considered as evidence for statistical discrimination based on educational attainment. When a prospective worker’s productivity is not directly observed, employers rely on indicators of quality, such as education, to set a wage rate. Eventually, they learn to distinguish between workers with similar educational backgrounds based on their job performance, which brings the wage rate closer to productivity. The statistical discrimination model thus predicts that correlates of productivity that are unobservable to employers do not explain wages of young workers, but become important determinants with experience. Furthermore, if these correlates are related to education, then educational attainment is expected to be less and less important for experienced workers (Farber and Gibbons, 1996; Altonji and Pierret, 2001).

In this paper, I develop and test a human capital model that is also consistent with the empirical regularities above. It is natural for workers with better learning skills to have a comparative advantage in acquiring new skills, not only through formal education, but also through job training and formation. Consequently, these workers have higher educational attainment on average, and they enjoy faster earnings growth throughout their career. It is conceivable, however, that, conditional on ability, workers with less schooling invest more in their skills to compensate for their lack of formal education, and thereby enjoy faster earnings growth. In this case, earnings of workers with similar ability, but different schooling levels, converge to each other, and education becomes a poor determinant of earnings, especially for experienced workers.

1 This implicitly assumes a competitive labor market where potential employers have the same information as the current employer. Several papers in the learning literature have explored deviations from this assumption (see Gibbons and Katz (1991) and Schönberg (2007) among others.)

2 Several papers rely on these predictions to estimate patterns of employer learning, e.g. Lange (2007) estimates the speed of employer learning to bound the signaling role of education, Mansour (2012) measures the variation in employer learning by occupation, and Arcidiacono, Bayer, and Hizmo (2010) by education.

3 The human capital model developed here is treated as a distinct alternative to the employer-learning model, thereby abstracting from potential interactions between the two models. Gibbons and Waldman (1999, 2006) study in this vein the task assignment problem under uncertainty and link changes in marginal
I formalize this argument using a variation of the Ben-Porath (1967) model of life-cycle skill formation and derive conditions under which the model delivers the observed empirical patterns of wage growth by ability and education. These conditions are critical for delivering salient features of the wage data, such as the concavity of the wage experience profile. They are, therefore, met in reasonable formulations of the Ben-Porath model.

The key features that allow the model to match the observed wage patterns is the substitutability of job training for education and the complementarity between training time and ability in acquisition of new skills. Several testable implications emerge as a result. The model implies that workers with higher ability invest more time in job training, and that training activity is decreasing in education for workers of comparable ability. Furthermore, the return to training investment is similar to the return to education, especially for younger workers. Later on, potential complementarity between training time and human capital may lead to larger marginal return to training.

I test these implications using data on training programs from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY). Results from the Armed Forces Qualification Test (AFQT) is used as a proxy for ability following the literature. The data confirm a strong positive link between ability and training investment. Conditional on ability, workers with higher educational attainment spend less time on training activities in total. Furthermore, the (annualized) return to training is estimated to be 7.8 percent for young workers, which is similar to the return to a year of education at 8.3 percent. These findings provide a direct support for the human capital model.

Further investigation of training activity by experience nonetheless reveals some conflicting evidence. It appears that training and education are substitutes around the time of labor market entry, but a positive correlation between education and training emerges with experience. Although workers with higher educational attainment invest less in training early on, total training investments are similar after 15 years of experience for workers with similar abilities. By contrast, the gap in training investment by ability widens with experience as predicted by the human capital model. This suggests that the negative association between education and wage growth for workers with similar abilities remains a qualitative diagnostic for the statistical discrimination model.
The implications of the statistical discrimination model for wage growth were also used to test whether employers discriminate workers based on race (Altonji and Pierret, 2001). Yet, it is possible that employers are fully informed, and race is simply correlated with investment in human capital and the ability to acquire new skills. The data on training activity in the NLSY, however, do not show any systematic differences by race. This undermines the human capital story for understanding the racial differences in earnings growth, and reinforces the statistical discrimination theory as a more viable interpretation of racial differences in wage growth for young workers.

This paper is closest, in spirit, to Kahn and Lange (2010) who test the employer learning model using direct productivity measures from a single, large firm (the Baker-Gibbs-Holmstrom data). They find that the variation in individual productivity is the predominant factor in pay dispersion, suggesting a reinterpretation of the employer learning literature. The findings here confirm this conclusion using a more specific model of productivity changes, one that depends on job training, but in a more general sample.

Next section presents the model and outlines the key conditions that allow the model to generate the observed differences in wage growth by ability and education. Section 2 describes the data and Section 3 reviews the wage regressions that provide the base for the recent employer learning literature. Section 4 evaluates the implications of the human capital model for training and wages. Section 5 concludes.

1 A Model of Skill Formation with Heterogenous Agents

This section develops a variation of the Ben-Porath model of life-cycle skill formation. To study the differences in earnings growth by education and ability, the model is extended by allowing workers to differ in their ability to acquire new skills, and in their preferences for formal education. Workers with higher abilities learn more skills for a given amount of time. Their aptitude for learning facilitates acquisition of skills both during school and later at work. Workers also have varying preferences for education. This variation may originate from differences in social background, parents’ attitudes towards education, or,
simply, personal tastes for school as opposed to work.\footnote{A similar heterogeneity structure is used by Willis and Rosen (1979) who allow workers to differ in their productivity and in their cost of education as captured by the discount rate. This type of variation would not work here since workers with higher discount rates not only spend less time at school but also invest less in training in the Ben-Porath model.}

Workers live a finite life of $T$ periods. $K(t)$ denotes the human capital stock of a worker at age $t \in [0, T]$. The market for human capital is competitive. A unit of human capital is valued at rate $R$. Each worker is endowed with a unit of work effort at any time. Workers can accumulate human capital by investing a fraction $l$ of their effort to training. A worker’s current earnings are:

$$y(t) = RK(t)(1 - l(t)).$$

(1)

The human capital stock of a worker evolves according to the following production function.

$$\dot{K}(t) = z_i g(l(t), K(t)),$$

(2)

where $z_i$ is the individual specific skill parameter that captures differences in ability to accumulate additional skills. The production function is increasing in both arguments and displays diminishing returns to scale. It is assumed that the depreciation rate of human capital is zero.\footnote{This assumption does not alter our theoretical results and is consistent with the estimates reported in Browning, Hansen, and Heckman (1999).}

To capture the differences in preferences for schooling, let the flow variable $\psi_i$ denote the monetary equivalent of the net utility gain associated with schooling. $\psi_i$ and $z_i$ are independent.\footnote{This assumption is innocuous. The results presented here would hold as long as the correlation between the two terms is less than perfect.} Assuming that the utility function is linear in earnings, $\psi_i$ can also be interpreted as an idiosyncratic cost of schooling. The technology for human capital production is the same at school and at work. If the worker chooses to attend school, however, he has to devote all his time to human capital accumulation: $l(t) = 1$. Worker’s problem is to choose the optimal investment level to maximize his lifetime earnings. The individual
control problem can be formalized as follows.

\[
\max_{l(t) \in [0,1]} \int_0^{T(s)} e^{-rt} R k(t)(1 - l(t)) + \psi_l(t) l(t) I_{l(t) \geq 1} \left[ 1 - l(t) \right] + \lambda(t) z_l g(l(t), K(t)) dt 
\]

subject to

\[ \dot{K}(t) = z_l g(l(t), K(t)). \]

\( I_{l(t) \geq 1} \) is an indicator function which equals 1 if worker is engaged in full-time human capital production. \( T(s) \) is the length of a worker’s career which may depend on his education. It is assumed, in general, that \( T(s) \leq T + s \). The Hamiltonian for the control problem above is:

\[ H = e^{-rt} R k(t)(1 - l(t)) + \psi_l(t) l(t) I_{l(t) \geq 1} + \lambda(t) z_l g(l(t), K(t)) \]

where \( \lambda(t) \) is the shadow value of investment in human capital, and \( K(0) \) is given and common to all workers. The interior solution is fully characterized by the following equations:

\[
\begin{align*}
-\frac{\partial H}{\partial K} &= -e^{-rt} R(1 - l(t)) - \lambda(t) z_l g_k(l(t), K(t)) = \dot{\lambda}(t) \quad (3) \\
\frac{\partial H}{\partial l} &= -e^{-rt} K(t) R + \lambda(t) z_l g_l(l(t), K(t)) = 0 \quad (4) \\
\lambda(T(s)) K(T(s)) &= 0 \quad (5)
\end{align*}
\]

Equation (3) implies that the shadow value of investment is decreasing in time as the worker gets closer to the end of his career. This generates a decreasing investment profile and an increasing wage profile over a worker’s career.

Equation (4) is the interior optimality condition for the time invested in training. As \( \lambda(t) \) decreases, the marginal return to training investment decreases. Meanwhile, capital stock accumulates raising the opportunity cost of training time. If the complementarity between capital stock and time investment, \( g_{lK} \), is not too large, then the model generates a monotonically decreasing investment profile and an increasing earnings profile over worker’s life.
Workers with higher ability have higher marginal return to investment conditional on the stock of human capital, and they spend more time on training activities. This generates lower earnings at the beginning of their career and yields faster wage growth. If these workers also attain higher education levels, as discussed next, then education and wage growth are positively related.

1.1 Schooling Period

Workers who attend school spend their entire time endowment on human capital accumulation at school. To analyze the schooling choice, consider the first order condition when \( l(t) \) approaches 1 from the right, \( \frac{\partial H^+}{\partial l} \). An agent chooses to specialize in schooling if the marginal return to investment is strictly greater than the marginal cost at \( l(t) = 1 \).

\[
\frac{\partial H^+}{\partial l} = -e^{-r(t)}K(t)R + \lambda(t)z_i g(t)(1, K(t)) + \psi_i > 0
\]  

(6)

This occurs if the initial level of human capital is low and the marginal return to training is high. As the worker invests in his capital, \( K(t) \) increases, raising the time cost of investment. At the same time the marginal return to additional human capital declines as the worker’s horizon shortens, given, once again, that \( g_{ik} \) is not too high. Therefore the specialization period only happens at the beginning of a worker’s life. The worker leaves school when the inequality above is reversed.

Equation (6) yields a testable implication of the model. At the time a worker leaves school, the marginal return to a year of investment in training equals the marginal return to the last year of schooling on average. This is a direct consequence of the substitutability of training for schooling. Furthermore, if existing capital and time are complements in production of new human capital, \( g_{ik} > 0 \), then the marginal return to training increases as workers build human capital. A second implication, therefore, is that the return to training increases with experience. I test both these implications below.

Workers with higher utility attachments to school, \( \psi_i \), or with higher ability, \( z_i \), are more likely to stay longer at school. These are the driving forces of the heterogeneity in the model. If ability is not directly observable to the econometrician, then higher wage growth caused by higher ability is attributed to education. However, conditional on abil-
ity, a higher education level results in a higher stock of human capital at graduation, which reduces the optimal training investment over the life-cycle, due to diminishing returns. I next focus on the heterogeneity in wage growth and formalize the main result.

1.2 Ability, Education and Wage Growth

Since ability raises the speed of human capital accumulation in this model, those with higher ability experience faster wage growth on average. To see this, consider the average wage growth over a worker’s career:

\[
\frac{w(T(s) + s) - w(s)}{T(s) - s} = \frac{RK(T(s) + s) - RK(s)(1 - l(s))}{T(s) - s} = z \int_0^{T(s)} g(l(t), K(t)) dt + l(s)K(s)
\]

where the terminal condition \(l(T) = 0\) is imposed. Workers with higher ability not only produce more human capital for a given level of investment, \(l(t)\), but they also invest more intensively in training according to equation (4). This implies that the term above is unambiguously increasing in \(z\).

Perhaps less evident is the negative effect of educational attainment on the rate of wage growth conditional on ability. Denoting the observed earnings for a worker with \(s\) years of education and \(x\) years of experience by \(w(x, s)\), the growth rate of earnings at experience level \(x\) is

\[
\frac{w_x(x, s)}{w(x, s)} = \frac{K_x(x, s)(1 - l(x, s)) - l_x(x, s)K(x, s)}{K(x, s)(1 - l(x, s))} = \frac{zg(l(x, s), K(x, s))}{K(x, s)} - \frac{l_x(x, s)}{1 - l(x, s)}.
\]

The dependence on \(z\) is suppressed for simplicity. The change in wage growth in response to education given ability generally depends on the properties of the human capital production function \(g(l, K)\). In what follows I formally establish this result for two broadly used classes of functions \(g(l, K)\).
1.2.1 Case I: $g(l, K) = f(l)K^\alpha$

Consider first the extreme case when $\alpha = 1$, i.e. production of new capital is a linear function of the level of current human capital. Suppose, for the moment, that a worker’s career is independent of his education: $T(s) = T + s$. The Hamiltonian conditions become,

\begin{align*}
-\frac{\partial H}{\partial K} &= -e^{-rt}R(1 - l(t)) - \lambda(t)z_if(l(t)) = \dot{\lambda}(t) \quad (8) \\
\frac{\partial H}{\partial l} &= -e^{-rt}R + \lambda(t)z_if'(l(t)) = 0 \quad (9) \\
\lambda(T)K(T) &= 0 \quad (10)
\end{align*}

This system defines a path for $\lambda(t)$ and $l(t)$ that is independent of capital stock. In general the shadow value implicitly depends on the educational choice if the work horizon is affected by years spent at school. Otherwise, the solution defined by the set of equations above is identical for all workers given $z_i$. The wage growth defined in equation (7) can be rewritten as,

$$
\frac{w_x(x,s)}{w(x,s)} = zf(l(x)) + \frac{l_x(x)}{1 - l(x)}.
$$

Since the time investment in training is independent of education, the only variation in wage growth comes from $z$. Conditional on $z$, log-wage and log-capital stock profiles over experience are parallel for different educational choices. This critical case displays a relative neutrality, in the sense that the marginal return to and the marginal cost of training time relative to potential earnings is independent of the human capital stock.

When the marginal return to training time increases less than proportional to potential earnings, the workers with higher potential earnings upon graduation find it optimal to spend less time on training activities. The following proposition establishes this result (the formal proof is relegated to the appendix).

**Proposition 1** Suppose that $g(l(t), K(t)) = f(l)K^\alpha$ and that the length of a worker’s career is fixed independent of the schooling level, $T(s) = T + s$. Then the rate of wage growth over experience is decreasing in education, conditional on $z$, if $\alpha \leq 1$, with equality if and only if $\alpha = 1$. 

9
Diminishing return to scale generates lower earnings growth through three distinct channels. First, with $\alpha < 1$, the rate of growth of potential earnings is decreasing in the level of capital stock for a given rate of investment. Second, the workers with higher capital stock invest less time in human capital accumulation which further slows down the growth of human capital. This results in reduced growth of potential earnings for workers with the same ability level, but higher educational attainment. Third, since actual earnings equal potential earnings net of investment expenditures, the rate of decline in investment expenditures relative to earnings also matters for the differences in the growth rates of earnings. Workers with higher capital stock not only spend less time training, but also choose to decrease their training time more slowly with experience. As less effort is released from the training activities, and allocated to work, earnings grow more slowly. Thus, the differences in the growth rates of earnings are more pronounced than those of potential earnings.

It was assumed that the length of a worker’s career is fixed regardless of the time spent at school. If a worker’s career is limited by age, e.g. due to mandatory retirement, the workers with more schooling bear an additional cost to investment in education. After schooling period is over, they face a shorter work horizon and invest even less in training, which amplifies the result above.

1.2.2 Case II: $g(l, K) = g(lK)$

A widely used functional restriction on the technology of human capital production assumes that current human capital raises the return to market time and training time equally. This is achieved when the training time and the current capital stock enter the production function multiplicatively: $g(l(t), K(t)) = h(l(t)K(t))$, where $h' > 0$, and $h'' < 0$. This is neither a special nor a more general case of the functional form assumed before. This production function is said to be “neutral” in the sense that the life-cycle profiles of the level of potential earnings for workers with different educational attainments are parallel. We refer to this case as absolute neutrality.

Under absolute neutrality, an explicit solution for $\lambda(t)$ exists, which simplifies the expressions for optimal investment and wage growth. Nevertheless, the generality of $h(.)$
prevents an equally general result in this case. Lower earnings growth for workers with more schooling is achieved only under a third order condition. However, it is consequently demonstrated that this condition is necessary to produce log-earnings profiles that are concave in experience, a salient feature of earnings consistently observed in the data.

Before we continue, let $Q(t) = l(t)K(t)$ denote the total outlays on training. Combining the Hamiltonian conditions with the transversality condition yields the following particular solution: $\lambda(t) = Re^{-rt}(1 - e^{-r(T-t)})/r$. Substituting this solution in the first order condition for training gives:

$$z_i h'(Q(t))(1 - e^{-r(T(s)-t)}) = r.$$  \hspace{1cm} (11)

Under absolute neutrality, the optimal investment expenditure depends only on time (or rather on the work horizon), and, hence, are independent of the level of current human capital. If we also suppose for a moment that $T(s) = T + s$ is fixed, then the workers with different education but same ability have the same optimal training expenditures. This implies that the potential earnings profiles are parallel. In this case, the rate of growth of earnings can be written as follows.

$$\frac{w_x(x,s)}{w(x,s)} = \frac{zg(Q(x)) - Q_x(x)}{K(x,s) - Q(x)}.$$ \hspace{1cm} (12)

The numerator of this expression is identical for all workers with the same ability, at any experience level, since they have the same training expenditure. The only difference is the the stock of human capital which is increasing in education. Since the workers with higher education start their careers with higher levels of human capital, their earnings are always equally higher. This immediately implies that the changes in earnings over experience are the same for all workers, but the growth rate is lower for workers with more education. This is merely a scale effect.

Figure 1 depicts the training intensity for two workers with the same learning ability but different schooling levels. The optimal path for the worker with higher education lies below the other worker, because (i) he has less time to reap the benefits of his invest-
ment, (ii) he has a higher stock of human capital. The total investment expenditure is the same for both workers as it is determined only by age. This generates parallel potential earnings profiles as a function of age (Figure 2a). However, the worker with higher education has a shorter career, therefore he faces a lower marginal rate of return to investment conditional on experience. Figure 2b shows potential and observed earnings as a function of experience. The shifted profiles display a lower level of investment expenditures, \( Q_s(s, x) < 0 \), and a lower growth in potential earnings at any experience level, \( K_{xs} < 0 \).

In order to link the argument to observed earnings, we also need to consider the rate of decline in the investment expenditures. If the training expenditure is convex in age, indicating that the investment decreases faster for younger workers, then the observed earnings growth is also lower for the educated, conditional on ability\(^7\). The convexity of the training expenditures in turn depends on a third order condition on the production function. The following proposition formalizes the result.

**Proposition 2** Suppose that \( g(l(t), K(t)) = h(l(t)K(t)) \). Then the rate of wage growth over experience is decreasing in education, conditional on \( z \), if

\[
\frac{h''' h'}{(h'')}^2 = k \geq 1,
\]

where \( k \) is the hyperbolic absolute risk aversion (HARA) in the context of utility functions. The exponential function, \( h(x) = -e^{\theta x}/\theta \), satisfies this condition with equality. The Cobb-Douglas function, which is a widely used parameterization of the human capital production function in the literature, also satisfies this condition.\(^8\) Nevertheless, this condition is not necessary. In fact, when this condition is satisfied with an equality, the proposition still holds due to the scale effect discussed above. The condition may seem arbitrary in the context of production functions, but it is critical for generating earnings profiles that are increasing and concave in experience. The following proposition states the equivalence of the two empirical regularities.

\(^7\)Note that the denominator in equation (7) is still increasing in \( s \). If, at any time, the worker with less education were to catch up, then the two workers would have the same wage profile from then on.

\(^8\)The exponential function displays a constant absolute risk aversion (CARA) of 1 in utility theory. CRRA has \( k > 1 \), and quadratic has \( k = 0 \).
**Proposition 3** The rate of wage growth over experience is decreasing in education, conditional on $z$, if wage profiles are log-concave in experience.

This is a direct result of the absolute neutrality assumption. Recall from equation (12) that wage growth is summarized as a function of the investment level, $Q$. Under absolute neutrality, the optimal investment choice depends only on the time left to retirement, $T - t$, which is equivalently $T - s - x$. Then the rate of change in investment with respect to age, experience and education are equivalent. This allows us to write the proposition in terms of changes with respect to time. See appendix for the derivation.

### 1.3 Wage Dispersion over Experience

The employer learning model also predicts that the variance of earnings, conditional on education, grows over time. Workers with the same education level start out with similar wages, but as employers learn to distinguish among these workers, wages become more disperse.\(^9\) On the other hand, the variance of earnings, conditional on ability, diminish over time. In fact, if education is a pure indicator, without any value added to productivity, the variance converges to zero. Both of these predictions are consistent with the human capital model presented above.

Take two workers with the same level of education, but with different ability levels. We have established above that the worker with higher ability has a higher wage growth over experience. This is due to two reasons. First, the worker with higher ability is better at accumulating human capital, leading faster wage growth for similar investments. Second, since the worker with lower ability is relatively over-invested in human capital through education, he chooses to invest less in training, further slowing down his wage growth. This widens the wage gap for people with similar education, leading to higher variance over time conditional on education.\(^10\)

---

\(^9\)Mansour (2012) uses this prediction to support the employer learning literature, and applies it to gauge employer learning in various occupations.

\(^10\)If the high ability worker invests sufficiently more in training, he may start out at a lower wage. In this case the variance may decrease at the beginning for a brief period. Heckman, Lochner, and Todd (2006) provide some evidence for this prediction.
If we compare workers with the same ability level but different educational attainments, the wage gap will be largest at the beginning. As the worker with lower education invests more in training, he compensates for his lack of education and the wage gap narrows (See Figure - 2b). The variance conditional on ability, therefore, diminishes with experience.

The key component in the model is that the level of investment in human capital is decreasing in education among workers of comparable ability levels. The model also predicts that the difference in investment levels is particularly large during the early career, and shrinks over experience. Next I examine the empirical validity of these predictions.

2 Data

The data is taken from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY). NLSY contains detailed information on workers in transition from school to labor market when most training activity and employer learning occurs. The subjects were surveyed on several economic variables including their education, family background and labor market outcomes. NLSY also reports the training activities of workers on and off the job throughout the survey, and it contains the scores of the Armed Forces Qualification Test (AFQT), a variable that is widely used as a proxy for cognitive ability in the literature.

In regard to training activities, the respondents were asked whether they had participated in any vocational or technical training programs. Then, detailed information on the duration of each training spell and the type of the training activity were collected. In particular, the respondents indicated one of the following categories: (1) Business College, (2) Nurses Program, (3) Apprenticeship, (4) Vocational Technical Institute, (5) Barber-Beauty School, (6) Flight School, (7) Correspondence, (8) Company Training and (9) other. Beginning in 1988, category (8) was split into three as (8) Formal Company Training, (9) Seminars or training programs at work not run by employer and (10) Seminars or training programs outside of work. I group the reported categories in two different ways and analyze them separately. First, I test the implications of the model using all training activities. Second, I present the results separately for on-site company training (category (8))
before 1988 and categories (8) and (9) after 1988.). These categories are often considered as on-the-job training spells in the literature (e.g. Parent (1999)).

To test the implications of the model for training, first, an annual variable for training investment is calculated going back exactly one year from the survey date. This results in loss of some training spells that occur between interviews that are more than a year apart. Consequently, the observed incidence of training is low. Only 6.9% of the workers report investment in training during the year prior to the survey. This figure is 2.6% if we restrict attention to company training only. This is not surprising since the data reports only formal training spells. Several activities on-the-job could qualify as training, though not necessarily reported as one.\footnote{For instance, if a worker spends a few hours reading the manual of a new machine that he needs to operate, or spends an hour listening to a co-worker’s advice on how to perform his task better, he would be considered investing in his human capital.} Therefore the training observed in the data is likely a lower bound for the actual investment in human capital. A second feature in the data is that most training occurs at the onset of a worker’s career. This is evident in Figure 3, which shows the annual hours of total training and on-the-job training by experience.

The low incidence of training observed in the data results in heavy abundance of zeros, making it harder to draw robust inference. To get around this problem, a cumulative variable is created in total hours of training investment. This reduces the incidence of zeros by cumulating training spells over a worker’s career, and by drawing in additional training spells between interviews that are further apart. For the sample used in the analysis the incidence of zeros for total investment is 52%. This percentage declines from 76% among workers with less than one year experience to 35% among those with fifteen years of experience.

Like training, most of the employer learning takes place when workers have just entered the labor market. Lange (2007) estimates that 90% of the learning is completed by 15 years. The sample is, therefore, restricted to workers who are within 15 years of their entry into the labor market.\footnote{Since the NLSY contains relatively younger workers, the sample size also decreases quickly with experience.}
3 Ability, Education and Wage Growth

Table 1 shows the wage regressions that constitute the motivation for the employer learning literature. The first column shows the results of a regression of log wages on education, AFQT score and an interaction of education with experience. The control variables are indicators for years of experience, survey years and race. As expected, wages vary positively with both ability and education. Return to a year of education is 8.0% and the market value of one standard deviation higher AFQT score is 9.3%.

The interaction term between education and experience is statistically zero. The second column adds the interaction of AFQT with experience. Results show that wages grow 6.8% faster for every 10 years for one standard deviation higher AFQT score above the mean. More interestingly, the coefficient on education interacted with experience is now significantly lower. A year of education translates into 1.6% lower wage growth per decade. These findings are consistent with Altonji and Pierret (2001) and Lange (2007). The effects are particularly stark early on. Column 3 estimates the same specification for workers with 10 or less years of experience. The coefficient on education interacted with experience is in fact -2.5% for the first 10 years.

Altonji and Pierret (2001) use differences in wage growth by race to test whether employers engage in statistical discrimination based on race. Column 4 includes an interaction term between experience and an indicator for race. When ability is excluded from the regression, black workers seem to have similar starting salaries, conditional on education, but they have 1.08% lower wage growth per year for the first fifteen years. Once ability is controlled for, however, the starting salaries are about 7% lower for blacks, and they have 0.51% lower wage growth per year.

Interpreted within the confines of a statistical discrimination model, the findings above suggest that employers initially use educational attainment to distinguish between workers of different ability. As workers’ true abilities are revealed by their performance on the job, education becomes less important relative to the AFQT score. The findings on the growth patterns by race indicate that employers also use race, or another variable that is correlated with race, to distinguish workers’ abilities. Since the average racial wage gap is about 11%, this constitutes about two thirds of the racial wage gap. The remaining
component of the wage gap emerges as wages further reflect ability over experience.

4 Testing the Human Capital Model with Job Training

The capacity of the human capital model to explain the patterns of wage growth discussed above depends on the value of job training in the market and on the differences in training investment by ability and education. Thus I begin the analysis by estimating the return to job training.

4.1 Return to Job Training for Young Workers

Table 2 reports the results from the regression of log-wages on education, ability and training. The training variables included in the regression are cumulative training investment and hours of training activity during the year before the survey. The former is akin to years of schooling in that it captures the total time input to human capital production. The second variable reflects the opportunity cost of training activity in terms of foregone earnings during training. All training variables are annualized based on 52 40-hour weeks. The control variables are indicators for potential experience, race and survey year.

The estimated return to a year of general job training is 7.9%. The return to a formal company training is 8.0%. Both of these are similar to the return to a year of education estimated at 8.3%. Workers that are currently engaged in general training activities are subject to a 13.7% reduction in wages. This effect is consistent with the model where time is the main input, albeit quantitatively small for a year of full-time training. In fact, workers that have recently received company training appear to have 17.4% higher wages. One possibility is that the receipt of company training is correlated with components of worker productivity other than those included in the regression. To get around this, columns (3) and (4) include fixed worker effects in the regression. Since education and ability are also fixed worker traits, they are not included in these estimations. The effect of current training on wages decreases slightly to -16.9% for general training and drastically to -3.7% for company training, although the latter is not statistically significant. While these es-
imates suggest that workers do not bear a large part of the training costs, as modeled above, it is also possible that the cost of training is internalized through an employment margin where the worker does not receive any pay for a period of time, or through a wage margin, but over a longer term. Another possibility is that training investment is more prone to measurement error relative to years of schooling, and, therefore, subject to a more severe attenuation bias.

When fixed effects are included, the return to a year of training is 9.3% for general training and 9.6% for company training, slightly higher than the return to education. When there are complementarities between capital stock and time investment, the model predicts that the marginal return to training is similar to that of education for young workers, and increases with experience. To test this, the last two columns repeat the fixed-effects specification for workers with 5 years of experience or less. The return to general training is 7.8%, which is lower than the 9.3% estimated in the whole sample although the difference is not statistically significant. While this seems to somewhat support the model’s prediction, the return to company training goes in the opposite direction. The estimated return to company training is 12.6% for young workers, higher than the 9.6% estimated for the sample.

Overall the estimates of the return to training suggest that training investment is as important as schooling, which supports the notion that training and education are substitutes. The results on the secondary predictions of the model are less clear. Results based on general training activities lend somewhat stronger support to the model than on-site company training.

4.2 Ability, Education and Training Intensity

The results in the previous section attribute a crucial role for training investment in wages. Then, could the differences in wage growth by education and ability be explained, at least in part, by differences in training? The results presented in Section 3 are consistent with the human capital model if workers with ability are more likely to receive training and workers with higher education and blacks are less likely to receive training, conditional on ability. In this section I directly test for these results.
Table 3 presents the marginal effects from a tobit regression of training investment on educational attainment, AFQT and race. The control variables are indicators for survey years, potential experience and race. The results show that training investment is strongly related to ability. One standard deviation higher AFQT score is associated with 132 hours of additional training activity in total, and 113 hours of additional company training. Conditional on ability, however, training activity is negatively related to educational attainment. Workers with an additional year of education have 30 hours less training activity in total, and receive 27 hours less company training.\(^{13}\)

Given the estimated returns to training investment in table 2, the estimated differences in training activity are not likely to generate large differences in wage growth.\(^{14}\) Nonetheless, the patterns in Table 3 are qualitatively consistent with the human capital model. If the informal training activity that goes unmeasured in the data displays similar qualitative features, perhaps the differences in wage growth by education and ability could very well be explained by a training model.

By contrast, there is no apparent relationship between race and training, once education and ability are controlled for. The results indicate that Blacks have slightly less training activity, but the difference is not statistically significant. These findings lend support to the validity of the statistical discrimination hypothesis by ruling out human capital formation by training as a viable explanation for racial differences in earnings growth.

Recall from Table 1 that the differences in wage growth by education and ability are particularly strong early in the career. To generate this the differences in total training investment by education and ability should widen with experience. To test this implication, the last two columns of Table 3 displays the marginal effects of ability and education on total training investment by experience. The results show that workers with higher ability invest more in training both initially and later on. A worker with one standard deviation higher AFQT score engages in 73 hours of additional training initially, and continues to

\(^{13}\)Tobit seems appropriate given the number of zeros in the data. When a least squares regression is estimated, the corresponding coefficients of ability are 155 hours for total training and 144 for company training, and those of education are -53 for total training and -50 for company training.

\(^{14}\)132 hours of additional training for a one standard deviation increase in the AFQT score brings about 0.59% increase in wages (evaluated at 9.3% return to job training), while the average return to AFQT in the data is 9.3%.
invest 70 hours more per decade. A similar result is obtained for company training where ability is associated with 54 hours of more company training initially, and 64 hours of additional investment per decade.

The results show an interesting pattern for the relationship between education and training investment. Conditional on ability, a worker with one more year of education invests 68 hours less in training at the beginning of his career, but invests 49 hours more per decade during the first 15 years of his career. This stands in contrast to the human capital model. While education and training are substitutes in formation of human capital at the onset of a worker’s career, they are rather complements in general. Those with higher education are likely to experience a higher wage growth conditional on ability as they accumulate longer training hours.

The results suggest that when training activity is not accounted for, wage regressions are likely to exaggerate the role of employer learning for differences in wage growth by ability. Furthermore, one is likely to find positive or no association between wage growth and education, conditional on ability.

Comparable estimates are hard to come by as the earlier literature on the determinants of training as data on the intensity of training and ability are rarely available together. Blundell, Dearden, and Meghir (1996) and Lynch and Black (1998), among others, finds a positive relation between training and education. This is not conditional on ability. Altonji and Spletzer (1991) include SAT scores of high schools seniors in their regressions, and find that the receipt of training depends positively on both ability and education for workers with more than 10 years of experience in the NLS72 survey. This is consistent with the findings here. Two other studies obtained mixed results. Lynch (1992) reports a higher incidence of training for high-school graduates in the NLSY in 1983. For workers with post-secondary education, off-the job training activity is lower while on-the-job training is similar. Veum (1995) studies the young workers in the NLSY during 1986 - 1990 and finds that the incidence of training is increasing in education, conditional on the AFQT score, with the exception of apprenticeships.
5 Conclusion

A model of statistical discrimination, where employers use education and racial background as indicators of otherwise unobserved productivity, is a coherent interpretation of observed differences in wage growth among young workers. It provides a rational interpretation for the racial wage gap, and explains why education has a diminishing role for wages of more experienced workers with similar abilities. This paper develops and tests a model of human capital as an alternative to the employer learning model.

The findings based on the data on job training from the NLSY lend partial support to the human capital model. Increased intensity of training activity could potentially explain the faster wage growth among young workers with higher cognitive ability. On the other hand, the substitution of training for education appears to be short-lived in the data. Therefore, the negative relationship between wage growth and educational attainment among young workers with comparable abilities remains a challenge for the human capital model. Furthermore, the absence of a meaningful relationship between racial background and job training casts doubt on the human capital theory as a suitable interpretation of the widening racial wage gap by experience.
References


A Technical Appendix

Claim 1 The rate of wage growth by experience is increasing in ability $z$ conditional on educational attainment.

Proof The proof is done in two steps. First, it is shown that the rate of wage growth is strictly increasing in the role of existing capital stock on human capital accumulation.
Second, it is shown that when \( g_k = 0 \), wage growth is increasing in ability conditional on education. The derivative of equation (7) is

\[
\frac{d(w_x/w)}{dz} = (zg_K \frac{dK}{dz} + zg_l \frac{dl}{dz} + g)K - zg_K \frac{dK}{dz} - \frac{dl_x}{dz} \frac{1}{1 - l_x},
\]

which is increasing in \( g_k \) since \( \frac{dK}{dz} > 0 \) given \( s \).

Next, suppose that \( g_k = 0 \), i.e. \( \dot{K} = zf(l) \). In this case, total capital stock is linear in \( z \) at any experience level. To see this, note that \( K(t) = K(0) + \int_0^t zf(l(t))dt \). Equation (7) becomes

\[
\frac{w_x}{w} = \frac{zf(l(t))}{K(0) + z \int_0^t f(l(t))dt} - \frac{l_x(t)}{1 - l(t)}.
\]

Note that, conditional on the path of \( l(t) \), wage growth is increasing in \( z \). Next, I show that when \( g_k = 0 \), investment intensity, \( l(t) \), is increasing in ability. This is evident from the optimality condition (4), which becomes

\[
e^{-rt}R[K(0) + zf(1)s + \int_0^x f(l(t))dt] = \lambda(t)zi f'(l)
\]

under the assumption that for \( g \). The term on the left is the marginal cost, and the term on the right is the marginal benefit of time investment in training. When \( K(0) = 0 \), \( z \) increases the marginal cost and benefit of training proportionally, leading to the same training paths for two workers with different abilities. When \( K(0) > 0 \), marginal cost responds less than proportionally to ability, encouraging high ability to worker to invest more in training.

Claim 2 Suppose that \( g(l(t), K(t)) = f(l)K^\alpha \) and that the length of a worker’s career is fixed independent of the schooling level. Then the rate of wage growth over experience is decreasing in education conditional on \( z \) if \( \alpha \leq 1 \), with equality if and only if \( \alpha = 1 \).

Proof With this formulation, we have a finite time control problem where workers differ only in their capital stock upon graduation from school. The Hamiltonian conditions
become\textsuperscript{15}

\begin{align*}
\frac{\partial H}{\partial K} &= -e^{-rt}R(1-l) - \lambda z_i f(l) K^{\alpha-1} = \dot{\lambda}(t) \quad (13) \\
\frac{\partial H}{\partial l} &= -e^{-rt}KR + \lambda z_i f'(l) K^\alpha = 0 \quad (14) \\
\lambda(T+s)K(T+s) &= 0 \quad (15)
\end{align*}

\textbf{Case 1: }\alpha = 1

Notice that if \(\alpha = 1\), this system identifies a time path for \(l(t)\) and \(\lambda(t)\) that is independent of the existing capital stock. The wage growth over experience is given by the following formula.

\[
\frac{\dot{w}}{w} = \frac{\dot{K}(1-l) - \dot{l}K}{K(1-l)} = z_i f(l) - \dot{l} / (1-l)
\]

where the second equality follows from the hypothesis of the proposition. Since the time investment, \(l\) is independent of the initial capital stock, wage growth conditional on \(z_i\) is identical for all education groups.

\textbf{Case 2: }\alpha < 1

The wage growth over experience in this case is

\[
\frac{\dot{w}}{w} = z_i f(l) K^{\alpha-1} - \dot{l} / (1-l)
\]

The first term is decreasing in current capital stock. The second term is positive and is also decreasing in capital stock. To see this rewrite the first order condition,

\[
f'(l) = \frac{e^{-rt}RK^{1-\alpha}}{\lambda z_i}
\]

Notice that the concavity of \(f(l)\) implies that the fraction of time allocated to training is decreasing in the current capital stock for a given lambda. Taking logs and differentiating with respect to time we get;

\[
\frac{f''(l)\dot{l}}{f'(l)} = -r + (1 - \alpha)z_i f(l) K^{\alpha-1} - \dot{\lambda} / \lambda
\]

\textsuperscript{15}Time arguments are suppressed for convenience
The second term on the right hand side of this equation is decreasing in \( K \). To investigate the third component, combine the first two Hamiltonians.

\[
\dot{\lambda} = -e^{-rt}R((1 - l) + \alpha f(l) / f'(l))
\]

The time path for \( \lambda \) depends only on \( l \). The response in \( \dot{\lambda} \) for a given change in \( l \) is

\[
\frac{\partial \dot{\lambda}}{\partial l} = -e^{-rt}R(-1 + \alpha(1 - f'' f / (f')^2))
\]

A rise in time investment slows down the decline in \( \lambda \) if \( \alpha \) is low enough. If this is the case, than the right hand side of equation (17) is definitely decreasing in \( K \) (by the chain rule \( \frac{\partial \dot{\lambda}}{\partial K} = \frac{\partial \dot{\lambda}}{\partial l} \times \frac{\partial l}{\partial K} < 0 \)). Since \( f'' < 0 \), equation (17) implies \( d\dot{l}/dK > 0 \) and hence the wage growth in equation (16) is decreasing in capital stock.

If \( \alpha \) is closer to 1, then \( \frac{\partial \dot{l}}{\partial K} < 0 \) implying that the sign of \( d\dot{l}/dK \) is indeterminate. But note that this derivative is monotonic in \( \alpha \), since \( \frac{\partial^2 \dot{\lambda}}{\partial \alpha \partial l} < 0 \), and that at best when \( \alpha = 1, \frac{d\dot{l}}{dK} = 0 \). Therefore for any \( \alpha < 1, \dot{l} \) is increasing in current capital and \( l \) is decreasing. This ensures that a worker with higher capital upon graduation (due to more schooling) will invest less today and his investment profile will decrease less than a worker with less capital (\( -\frac{\partial \dot{l}}{\partial K} < 0 \)). Since the terminal point for both workers is the same, \( l(T + s)\lambda(T + s) = 0 \), the investment profiles won’t cross. Hence, wage growth is decreasing in capital stock at all levels of experience. ■

Claim 3 \( \frac{\partial^2 \ln w}{\partial s^2 x} < \frac{\partial^2 \ln w}{\partial x^2} = \ddot{\ln w} \)

Proof Recall that the wage growth is

\[
\frac{w_x}{w} = \frac{zg(Q(x + s) - Q_s(x + s)) - Q_x(x + s)}{K(x, s) - Q(x + s)}
\]

Since \( t = x + s \), we have \( Q_x(x + s) = Q_s(x + s) = \dot{Q}(t) \), and hence \( w_x = \dot{w} \), and \( w_{xs} = w_{xx} = \ddot{w} \). Consequently, the numerator changes at the same rate with respect to education, experience and age. We are done if the denominator rises faster with education than it does with time/experience. Since education and experience enter symmetrically
in the investment, this is true if $K_s(x, s) > K_x(x, s)$. The equation for the capital stock at experience level $x$ is,

$$K(x, s) = K(s) + \int_s^{s+x} zg(Q(t))dt = K(0) + \int_0^s zg(K(t))dt + \int_s^{s+x} zg(Q(t))dt$$

Taking the derivative with respect to $x$ and $s$ separately and taking the difference,

$$K_s - K_x = zg(K(s)) + zg(Q(x + s)) - zg(Q(s)) - zg(Q(x + s)) = z(g(K(s)) - g(Q(s)) = z(g(K(s)) - g(l(s)K(s))$$

$\geq 0$

where the last inequality follows from the fact that $l(s) \leq 1$ at the time of graduation (which in turn comes from $\psi \geq 0$).

**Claim 4** The rate of wage growth is decreasing in education if $rzg'''(.) \geq [zg''(.)]^2(1 - e^{-2r(T-t)})$.

**Proof** From the previous claim we have, $w_{xs} = w_{xx} = \ddot{\omega}$. Since the wage itself (the denominator) is increasing in education, it would suffice to find a condition that guarantees $\ddot{\omega} < 0$.

$$\ddot{\omega} = zg'(Q)\dot{Q} - \ddot{Q}$$

Applying the implicit function theorem to equation (11), we get

$$\dot{Q} = -\frac{r^2e^{-r(T-t)}}{(1 - e^{-r(T-t)})^2zg''(Q)}.$$  

Taking the time derivative one more time and simplifying terms,

$$\ddot{Q} = \frac{r^3e^{-r(T-t)}(1 - e^{-2r(T-t)})zg''(Q) - r^4e^{-2r(T-t)})g'''(Q)/g''(Q)}{(1 - e^{-r(T-t)})^4[zg''(Q)]^2}$$

Plugging these in the equation for $\ddot{\omega}$ yields,

$$\ddot{\omega} = r^2e^{-r(T-t)} \left[ z^2g'g''(1 - e^{-r(T-t)})^2 - r(1 - e^{-2r(T-t)})gz'' + r^2e^{-r(T-t)}g'''g'' \right]$$
Replacing $zg'(Q)$ with $r/(1 - e^{-r(T-t)})$ from equation (11), this term is negative if and only if

$$rzg'''(.) \geq [zg''(.)]^2(1 - e^{-2r(T-t)})$$

### B Data

The data is taken from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY). The results here use the male workers from the nationally representative sample. Wages are defined as hourly earnings. All values are converted to 2004 dollars using the CPI. The Armed Forces Qualification Test Score is standardized by each age group as the respondents were of different age when they took the test.

**Training Variables:** At every survey respondents were asked if they had participated in a training program since the previous interview. Detailed information, then, were collected on the duration, intensity and the type of these training spells. In particular, the respondents reported the beginning and end dates of each training spell (in month and year), total number of weeks and the average number of hours per week in training. This enables the construction of the total time investment in training in hours since the last interview. The earlier surveys, 1979 to 1986, did not provide details for training spells that lasted less than a month. Workers who reportedly participated in a training program that lasted less than a month were assigned two weeks of training. A quarter to a third of the training spells during these years were less than a month. Training data was not available in 1987 although several workers reported on training spells during that time in later surveys. These were duly added to worker’s training history. Until 1988, up to three training spells were recorded. Later this limit was raised to four. The respondents were however asked if they had a fourth (fifth after 1986) training program to report. Based on this question, it is possible to calculate the number of workers for which this limit was binding. The limit was binding for a total of only 80 observations (about 0.2% of the sample) in all years.
### Table 1: Education, Ability and Wage Growth

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>0.080</td>
<td>0.094</td>
<td>0.096</td>
<td>0.080</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>AFQT</td>
<td>0.093</td>
<td>0.041</td>
<td>0.029</td>
<td>0.093</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.009)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.112</td>
<td>-0.112</td>
<td>-0.101</td>
<td>-0.025</td>
<td>-0.071</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td>(0.029)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Education×Exp/10</td>
<td>0.002</td>
<td>-0.016</td>
<td>-0.025</td>
<td>0.001</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.010)</td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>AFQT×Exp/10</td>
<td>0.068</td>
<td>0.099</td>
<td>0.062</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black×Exp/10</td>
<td></td>
<td></td>
<td>-0.108</td>
<td>-0.051</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.027)</td>
<td>(0.029)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>26,792</td>
<td>26,792</td>
<td>18,584</td>
<td>26,792</td>
<td>26,792</td>
</tr>
</tbody>
</table>

Note — All specifications control for indicators for potential experience, survey year and race. Standard errors are clustered by respondent and are reported in italics. NLSY 1979 men with 15 years of experience or less. Column (3) uses observations with up to 10 years of potential experience.
Table 2: Return to Job Training

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: log-hourly wages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Education</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
</tr>
<tr>
<td>AFQT</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
</tr>
<tr>
<td>Cumulative Training/(52x40)</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Current Training/(52x40)</td>
<td>-0.137</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative OJT Training/(52x40)</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Current OJT Training/(52x40)</td>
<td>0.174</td>
</tr>
<tr>
<td>Fixed Worker Effects</td>
<td>No</td>
</tr>
<tr>
<td>N</td>
<td>26,792</td>
</tr>
</tbody>
</table>

Note — All specifications control for indicators potential experience, survey year and race. Robust standard errors are clustered by respondent and are reported in italics. NLSY 1979 Men with 15 years of experience or less. Columns (5) and (6) use observations with less than 5 years of experience.
### Table 3: The Determinants of Training

<table>
<thead>
<tr>
<th>Training Type:</th>
<th>All</th>
<th>OJT</th>
<th>All</th>
<th>OJT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>-29.8</td>
<td>-27.3</td>
<td>-67.6</td>
<td>-53.7</td>
</tr>
<tr>
<td></td>
<td>(8.7)</td>
<td>(7.5)</td>
<td>(12.6)</td>
<td>(11.0)</td>
</tr>
<tr>
<td>AFQT</td>
<td>131.6</td>
<td>112.5</td>
<td>72.5</td>
<td>54.5</td>
</tr>
<tr>
<td></td>
<td>(19.5)</td>
<td>(16.8)</td>
<td>(24.6)</td>
<td>(21.4)</td>
</tr>
<tr>
<td>Black</td>
<td>-63.9</td>
<td>-49.0</td>
<td>-67.0</td>
<td>-53.5</td>
</tr>
<tr>
<td></td>
<td>(40.6)</td>
<td>(35.2)</td>
<td>(64.4)</td>
<td>(56.2)</td>
</tr>
<tr>
<td>Education $\times$ Exp/10</td>
<td>48.8</td>
<td>37.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(11.5)</td>
<td>(9.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFQT $\times$ Exp/10</td>
<td>69.6</td>
<td>63.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(19.6)</td>
<td>(17.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black $\times$ Exp/10</td>
<td>3.6</td>
<td>1.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(51.0)</td>
<td>(45.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>26,792</td>
<td>26,792</td>
<td>26,792</td>
<td>26,792</td>
</tr>
</tbody>
</table>

Note — Results from Tobit regressions. All specifications control for indicators potential experience, survey year and race. Robust standard errors are clustered by respondent and are reported in italics. NLSY 1979 Men with 15 years of experience or less.
Figure 1: Training Intensity with Absolute Neutrality

Figure 2: Human Capital, Earnings and Experience under Absolute Neutrality
(a) Annual Training Hours

(b) Annual On-the-job Training

Figure 3: Job Training by Experience – NLSY 1979 Men