NBER WORKING PAPERS SERIES

OMITTED-ABILITY BIAS AND THE INCREASE IN THE RETURN TO SCHOOLING

McKinley L. Blackburn

David Neumark

Working Paper No. 3693

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1991

The authors wish to thank John Bishop, David Card, Bill Carter, John Chilton, Elchanan Cohn, Andrew Foster, Jacob Mincer, Paul Taubman, and seminar participants at the University of Pennsylvania, the University of Maryland, and Columbia University, for helpful comments. Sadiq Currimbhoy provided excellent research assistance. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3693
May 1991

OMITTED-ABILITY BIAS AND THE INCREASE IN THE RETURN TO SCHOOLING

ABSTRACT

Over the 1980s there were sharp increases in the return to schooling estimated with conventional wage regressions. We use both a signaling model and a human capital model to explore how the relationship between ability and schooling could have changed over this period in ways that would have increased the schooling coefficient in these regressions. Our empirical results reject the hypothesis that an increase in the upward bias of the schooling coefficient, due to a change in the relationship between ability and schooling, underlies the observed increase in the return to education over the 1980s. We also find that the increase in the return to education has occurred largely for workers with relatively high levels of academic ability.

McKinley L. Blackburn
Department of Economics
University of South Carolina
Columbia, SC 29208

David Neumark
Department of Economics
University of Pennsylvania
Philadelphia, PA 19104
and NBER

I. Introduction

In the 1980s, the United States experienced considerable changes in the structure of wages paid to different demographic and educational groups. The most notable of these changes is a large increase in wage differences of white males at different educational levels, as the wages of more-educated workers increased relative to their less-educated counterparts.[1] For instance, among white males between the ages of 25 and 34 in 1979, college graduates earned roughly 15 percent more than workers who had only completed high school; among 25-34 year-olds in 1987, college graduates earned 33 percent more than high-school graduates.[2] Blackburn, Bloom, and Freeman (1990b) show that wage differences associated with education also increased for white females, while Katz and Revenga (1989) also point to increases in the 1980s in wage differences associated with the level of labor market experience of white males.

Several studies have attempted to explain why we observe an increase in education-related wage differentials in the 1980s. Perhaps the most commonly-offered explanations have been associated with changes in the relative demand for workers at different educational levels. For example, it has been suggested that changes in international trade patterns have shifted relative labor demand curves in favor of the more-educated. An associated explanation singles out shifts in the industrial structure of the economy towards service-oriented production as the important factor. However, the available evidence suggests that these changes have at most played a minor role in the changes in earnings differentials. There is more evidence that

---

[1] See, for example, Katz and Revenga (1989), and Blackburn, Bloom, and Freeman (1990a).

[2] These statistics are taken from Table 1 of Blackburn, Bloom, and Freeman (1990a).

changes in the supply of workers at different educational levels have contributed to the changes in earnings differentials, especially for younger white males, though the magnitude of this effect is sensitive to assumptions about the substitutability between more- and less-educated workers.[3] An alternative explanation (discussed in Blackburn, Bloom and Freeman, 1990a) is that there are changes in the average level of productive ability of workers in different educational classes. One reason for such a change could be changes over time in the quality of primary and secondary education.[4]

There are several theories that suggest that there might be a relationship between a worker's inherent ability (*i.e.*, ability not affected by acquisition of schooling) and his level of schooling. In the following section, we discuss two such theories -- a signaling model, and a human capital model -- and explore what might change over time so as to affect the schooling-ability relationship.[5] Since most empirical studies of the increase in the return to schooling do not attempt to control for the effect of unobserved ability on wages, any such changes in the schooling-ability relationship could have led to changes over time in the observed return to schooling.[6] In fact, it could be that the "true" return to schooling (*i.e.*,

---

[3] Stapleton and Young (1988) develop a model in which changes in sizes of entering cohorts affect the returns to higher education, because substitutability between young and old workers is higher for low-educated than high-educated workers.

[4] Blackburn, Bloom, and Freeman (1990a) rule out changes in the quality of education as a complete explanation by noting that education-related earnings differentials increased over the 1980s for older cohorts whose educational quality could not have changed over the period. However, the fact that differentials increased more rapidly for younger workers suggests that this may still be a partial explanation of increased differentials.

[5] Bishop (1989) suggests that changes in the return to schooling (ignoring ability) may be due to varying selectivity of colleges and graduate schools, or changes in the rewards to "credentials" *per se*.

[6] Taubman and Wales (1972) examine changes in the relationship between ability and schooling over a much longer period, utilizing estimates from a wide

the increase in earnings from one more year of schooling for a worker of a given quality) has not changed over time, and that the observed increase in earnings differentials is attributable to changes in the correlation between schooling and ability.[7]

We test the explanations of the apparent increased return to schooling generated by these models. First, we test the proposition -- generated by the signaling model -- that changes in the distribution of the workforce across education classes underlie the increases in the estimated return to schooling. We use CPS estimates of changes in returns to schooling, and changes in the distribution of workers across education levels, coupled with parametric assumptions about the distribution of unobserved ability, to explore the plausibility of the signaling explanation.

Second, we consider the proposition -- generated by the human capital model -- that, because of changes in the relationship between ability and schooling, the omission of ability as an independent variable in wage equations has led to an "observed" rather than a true increase in the return to schooling in the 1980s. Our empirical testing uses a sample of young white males from the National Longitudinal Survey Youth Cohort. The models we estimate take advantage of scores on several tests measuring academic (or cognitive) and mechanical ability that are available for each individual surveyed in the Youth Cohort data. These test scores are used as (potentially error-prone) measures of ability in wage equations.

Our empirical findings provide little or no support for the idea that changes in the relationship between ability and schooling in the 1980s are

---

variety of data sets. They do not, however, relate these changes to variation in the estimated return to schooling over time.

[7] Alternatively, an increase in the return to ability could lead to a spurious increase in the estimated return to schooling.

responsible for the increase in education-related earnings differentials. However, we do find evidence that the increase in earnings differentials has occurred primarily for workers with higher academic ability.

II. Theoretical Discussion of the Ability-Schooling Relationship

In this section, we consider the effect that the relationship between omitted ability and schooling has on empirical estimates of the relationship between education and earnings, and how changes in the ability-schooling relationship may change the estimated relationship between education and earnings. Our model for earnings is of the form

$$(1) \qquad w = \beta_1 S + \beta_2 A + \epsilon \quad ,$$

where w is the log of the wage, S is an education variable, A is an ability variable, and ϵ is an error term distributed independently of S and A.[8] Since A is not observed in the data sets used in recent studies of changes in the return to schooling, these studies have used the simple-regression coefficient b_{wS} as an estimate for β_1 (*i.e.*, they estimate [1] omitting ability from the regression).[9] The remainder of this section develops two models that show that a changing relationship between ability and schooling could, in principle, underlie the estimated increase in b_{wS} that has been found by other researchers.

First, a signaling model is developed, in which β_2 is zero because ability is unobserved by employers. In this model, b_{wS} reflects the combined effect of the contribution of schooling to a worker's marginal product, and

---

[8]The wage equation can be thought of as the partial relation of w with S and A, the correlation of other variables with w, S, and A having been removed.

[9]$b_{wS} = Cov(w,S)/s_S^2$, where Cov(w,S) is the sample covariance between log wages and schooling, and s_S is the sample standard deviation of schooling.

the information that schooling contains for the employer's expectation of the worker's ability. We show that changes in the distribution of workers across education classes can alter this second component of b_{wS}.

Second, we develop a human capital model in which both schooling and ability appear in the wage equation (*i.e.*, both β_1 and β_2 in equation (1) are non-zero). In this case,

$$(2) \quad E(b_{wS}) = \beta_1 + \beta_2 b_{AS} = \beta_1 + \beta_2 r_{AS}\left[\frac{s_A}{s_S}\right] ,$$

where r_{AS} is the (sample) correlation coefficient between ability and schooling, and s_A and s_S are the (sample) standard deviations of ability and schooling. If $\beta_2 > 0$ and $r_{AS} > 0$, b_{wS} is an upward-biased estimate of β_1. As we discuss below, changes in the distribution of ability and/or opportunity can lead to $E(b_{wS})$ changing while β_1 remains constant.

A. Signaling Model

In the signaling model of Spence (1973), high-ability workers obtain more schooling than low-ability workers because schooling provides a signal to employers that they have high levels of ability. Arrow (1973) has also presented a model where education serves a screening function in sorting out high-ability from low-ability workers. We present a simplified version of Arrow's model to study the potential importance to earnings differentials of sorting of workers into educational classes on the basis of their productive ability.

We make four simplifying assumptions about the screening process associated with education. First, the schooling decision is dichotomous; S is a dummy variable equal to one if a worker is in the "high" education class, and zero if he is in the "low" education class.[10] Second, the single equilibrium involves education sorting workers perfectly on the basis of ability, such that if $A_i > A_c$ then $S_i = 1$, and if $A_i \leq A_c$ then $S_i = 0$. Third, ability is completely unobservable while schooling is observable, so that β_2 in equation (1) is zero. And, fourth, employers pay log wages equal to $E(MP_i | S_i)$, where MP_i is the log of the marginal product of worker i.[11]

If MP_i is a linear function of schooling and ability,

$$(3) \quad MP_i = \gamma_1 S_i + \gamma_2 A_i \quad ,$$

then,

$$(4) \quad E(MP_i | S_i) = \gamma_1 S_i + \gamma_2 E(A_i | S_i = 0) \cdot (1 - S_i) + \gamma_2 E(A_i | S_i = 1) \cdot S_i$$

$$= \gamma_1 S_i + \gamma_2 E(A_i | A_i \leq A_c) \cdot (1 - S_i) + \gamma_2 E(A_i | A_i > A_c) \cdot S_i \quad .$$

Denoting the conditional means of A_i as

$$E(A_i | A_i \leq A_c) = \frac{\int_{-\infty}^{A_c} af(a)da}{F(A_c)} = \bar{A}_L \quad ,$$

$$E(A_i | A_i > A_c) = \frac{\int_{A_c}^{\infty} af(a)da}{1 - F(A_c)} = \bar{A}_H \quad ,$$

where $f(\cdot)$ and $F(\cdot)$ are the density and distribution functions for ability,

---

[10] For example, S might be a dummy variable for whether or not the worker has a high school education or better.

[11] The assumption that employers equate log wages with the $E(MP_i | S_i)$ is made to ease the computational burden of the simulations that follow. It does imply that firms are risk-averse in their assessment of a worker's ability.

equation (4) can be expressed

$$(5) \quad E(MP_i | S_i) = \gamma_1 S_i + \gamma_2 \bar{A}_L + \gamma_2 (\bar{A}_H - \bar{A}_L) \cdot S_i \quad .$$

If equation (5), plus a stochastic error term, determines the wage, then the expected value of the coefficient estimate obtained from the regression of w on S is

$$(6) \quad E(b_{wS}) = \gamma_1 + \gamma_2 (\bar{A}_H - \bar{A}_L) \quad .$$

It follows that the usual estimate b_{wS} provides an upward-biased estimate of the contribution of schooling to the actual marginal product.[12]

The estimated partial effect of schooling on wages can change without changes in the direct effect of schooling on productivity (*i.e.*, γ_1). In fact, increases in the educational attainment of a population can cause $E(b_{wS})$ to change, since a fall in A_c will affect $(\bar{A}_H - \bar{A}_L)$. Is it possible that such a scenario underlies the increase in the return to schooling in the 1980s? In general, it is not possible to determine how $(\bar{A}_H - \bar{A}_L)$ will vary with declines in A_c.[13] For instance, if the distribution of ability is uniform with finite

---

[12]In terms of the population counterparts to the sample statistics that contribute to the bias in equation (2), we have $\sigma_S^2 = p(1-p)$ and $\rho_{AS} = \left[p(1-p)\right]^{1/2} (\bar{A}_H - \bar{A}_L) / \sigma_A$, where p is the percentage of the population in the high-education category. Note that both σ_S^2 and ρ_{AS} change when A_c changes, while σ_A^2 does not change.

[13]From our expressions for the conditional means of A, we can obtain

$$\frac{\partial E(\bar{A}_H - \bar{A}_L)}{\partial A_c} = \frac{f(A_c)}{F(A_c)[1-F(A_c)]} \left[F(A_c)(\bar{A}_H - \bar{A}_L) - (A_c - \bar{A}_L)\right] \quad ,$$

the sign of which is indeterminate.

bounds, then it can be shown that $(\bar{A}_H - \bar{A}_L)$ does not depend on A_c. However, other distributions for ability do generate a difference between the conditional means that varies as the percentage of workers in the high-education class changes. In particular, we consider two single-peaked, symmetric densities for ability -- the standard normal distribution, and a triangular distribution with density

$$f(A) = \begin{cases} 1 - |A| & \text{if } -1 \le A \le 1 \\ 0 & \text{otherwise} \end{cases}$$

With both distributions, the difference between the conditional means of ability is at a minimum when $A_c = 0$, *i.e.*, $p = P(A \ge A_c) = 0.5$.[14] If $p < .5$, so that more-educated workers are a minority, then the difference falls as more-educated workers become more numerous; but once more-educated workers become a majority, the difference begins to increase again.[15] The changes in the difference are not small; *e.g.*, going from $p = .5$ to $p = .66$ would increase the difference between the conditional means by 13 percent if ability followed the standard normal distribution.

However, studies of the education-earnings relationship generally study more than a simple two-way breakdown of the education distribution. While the

---

[14] Numerical simulations revealed that if A follows an exponential distribution, the differential will follow a U-shaped pattern with respect to A_c. However, the minimum differential is reached at the mean of A, which is greater than the median.

[15] Let $\lambda(A_c) = E(A|A>A_c)$. For any symmetric distribution centered at zero, $E(A|A \le A_c) = -E(A|A>-A_c) = -\lambda(-A_c)$, so that $E(b_{wS}) = \gamma_1 + \gamma_2[\lambda(A_c) + \lambda(-A_c)]$. For the standard normal distribution, $\lambda'(A_c) > 0$ and $\lambda''(A_c) > 0$ (see Heckman and Honore, 1990); in this case the differential will increase with increases in A_c when $A_c > 0$ [since $\lambda'(A_c) > \lambda'(-A_c)$], and will decrease with increases in A_c when $A_c < 0$. We were unable to sign $\lambda''(A_c)$ for the triangular distribution; however, numerical calculations showed that the differential was also U-shaped, with a minimum at $A_c = 0$.

restriction to equal logarithmic changes in earnings per year of schooling is the most common empirical restriction, some studies have also used dummy variables for various educational classes. For example, Blackburn, Bloom, and Freeman (1990a) compare the earnings of workers who had less than a high-school education to those with a high-school education, and to those with a college education. Their findings (for white males aged 25-64 in the U.S. in 1973, 1979, and 1987) are reproduced in panel A of Table 1. Their results show only small changes in these differentials from 1973 to 1979, but much larger changes from 1979 to 1987.

Can changes in the percentage of workers in these various classes explain the observed movements in earnings differentials? Changes in educational attainment were considerable for white males over this period. For example, the percentage of working males 25-64 with less than a high-school education fell from .26 to .13 from 1973 to 1987, while the percent with a college degree increased from .21 to .32. Assuming that education perfectly sorts workers by ability and that the distribution of ability remains unchanged, and choosing a parametric distribution for ability, it is possible to predict how the observed changes in the percentage of workers in the various classes would affect the earnings differentials between these classes. These results are reported in the bottom two panels of Table 1.[16] For both the standard normal

[16]These simulations are based on an extreme version of the signaling model in which γ_1 in equation (3) is set to zero. The simulated earnings differentials are determined only up to a constant factor of proportionality. The coefficient on ability, γ_2, was standardized so that the college graduate to less than high school earnings differential in 1973 was the same in the actual and simulated numbers. Using this coefficient, the actual and simulated earnings differences in 1973 are very close for the college to high school graduate comparison. This experiment may overstate the effect of changes in the ability differentials between different educational classes, because the entire earnings differential is attributed to ability.

and triangular distributions, the results suggest that only a small part of the change in differentials over the 1979-1987 period might be explained by changes in educational attainment.[17]

Of course, the screening model formulation permits other reasons for the bias associated with using b_{wS} to change over time, including a change in the sorting process of education (*e.g.*, if education becomes a better filter over time), increases in the dispersion of ability, or an increase in the return to ability.

B. Human Capital Model

In the human-capital model of the relationship between schooling and earnings offered by Mincer (1974, Ch. 1), all workers have identical opportunities for, and receive identical rewards from, human-capital investments. In contrast, a model suggested by Becker (1975) allows ability to affect the rate of return to human capital investment, as he assumes that workers with higher levels of innate ability will also receive higher returns to their human capital.[18] Following Becker, we assume that for each individual the marginal benefit and marginal cost of education take the form

$$MB(A_i, S_i) = \exp(kA_i)S_i^b$$

---

[17] Jacob Mincer points out that the screening model suggests that the earnings of college graduates relative to all other workers should be falling over the 1973-1987 period, since college graduates were a growing minority (among 25-64 year-olds) throughout this period. While this differential did fall from 1973 to 1979, it increased from 1979 to 1987.

The simulations were also performed for changes in educational attainment for 25-34 year-old white males. Since there was almost no change in educational attainment for this group over the years 1979-1987, these results provide even less support for the importance of signaling explanations for the change in earnings differentials.

[18] The model is discussed in his Woytinsky Lecture of 1967, which is reprinted in Becker (1975).

$$MC(P_i, S_i) = \exp(-P_i)S_i^d \quad ,$$

where P_i represents the ith individual's opportunities for investing in human capital.[19] Assuming that b<d, the optimal level of investment in schooling is equal to

$$(7) \quad S_i^* = \exp\left[(kA_i+P_i)/(d-b)\right] \quad .$$

so that workers with higher ability (or higher opportunity) will invest more in education. The wage is given by $W_i = \int_0^{S_i^*} MB(A_i, x_i)dx_i$, which leads to the log-linear wage equation

$$(8) \quad w_i = \log(W_i) = -\log(1+b) + (1+b)\log(S_i^*) + kA_i \quad .$$

This form for the wage equation shows that if ability is held constant, then the empirical relationship between w_i and S_i^* reflects the (logarithmic) slope of the marginal benefit function for schooling.[20] However, if ability is

---

[19] A similar framework is used in Behrman, Pollak, and Taubman (1990). We assume that benefits and costs are expressed in constant per-period flows corresponding to the period for which the wage or income variable is measured.

[20] It is not possible, in this framework, to derive a marginal benefit function for schooling that leads to the commonly-used semi-log specification for the wage equation. To see this, note that the expression for W_i would have to be

$$W_i = \exp(f + gS_i^* + hA_i) \quad ,$$

which implies the marginal benefit function

$$MB = dW_i/dS_i^* = g \cdot \exp(f + gS_i^* + hA_i) \quad .$$

But this marginal benefit function, when integrated from 0 to S_i^*, does not yield the required form for W_i. At the least, this suggests that one should try more than one specification in estimating the effects of schooling on wages. Because we do not use log schooling, we do not impose the restriction in equation (8) that the negative of the log of the intercept is equal to the schooling coefficient.

omitted, a positive correlation between S_i^* and A_i would lead to the estimated return to schooling tending to overstate the marginal benefit of schooling, as those workers with higher levels of observed schooling will also be those with higher marginal benefits of education.

From equation (7), the covariance of log schooling and ability is

$$\sigma_{\log(S^*),A} = \frac{1}{d-b}\left[k\sigma_A^2 + \sigma_{P,A}\right] \quad , \tag{9}$$

where $\sigma_{P,A}$ is the covariance of ability and opportunity. If ability and opportunity are uncorrelated, or positively correlated, the covariance between log schooling and ability is positive, and the omission of ability from wage equations should lead to upward-biased estimates of the return to schooling. However, if A and P are negatively correlated, the sign of the covariance is not clear. Since opportunity represents anything that shifts the marginal cost curve for schooling, and since high-ability individuals are likely to face high foregone-earnings costs at the margin, P and A may be negatively correlated. In particular, if we allow wages to be a direct determinant of the marginal cost of education, we have

$$MC = \exp(C_i)W_i = \exp(C_i)\int_0^{S_i} MB(A_i,x_i)dx_i = \frac{1}{(1+b)}\exp(kA_i+C_i)S_i^{(1+b)} \quad . \tag{10}$$

In terms of the earlier expression for marginal cost, $d=(1+b)$ and $P_i=-kA_i-C_i+\log(1+b)$. With this restriction, $\sigma_{\log(S^*),A} = -\sigma_{C,A}$, so the sign is still indeterminate. If more able individuals have lower costs of education, net of opportunity-wage costs (as seems most plausible), the covariance between ability and log schooling will be positive, and the OLS schooling coefficient estimate will be an upward-biased estimate of the marginal benefit of education.[21]

The human-capital model suggests an interesting result for how changes in the distribution of opportunity will affect the bias in schooling coefficients. Assuming ability and opportunity are uncorrelated, an increase in the variation of marginal costs across individuals (*i.e.*, an increase in σ_P^2) should increase the variance of log schooling, but have no effect on the correlation between ability and log schooling.[22] From equation (2), it follows that the bias in the schooling coefficient (using the logarithm of schooling rather than the level of schooling as the independent variable) will fall as a result of the increase in the variation of opportunity. The intuition is that with higher variation in opportunity a smaller part of observed differences in schooling are due to differences in ability. Increases in the variance of ability will increase both the variance of schooling and the covariance between ability and schooling, but the net effect is that the bias in the schooling coefficient will increase.[23] Not surprisingly, increases in k -- the

---

[21]Strictly speaking, this discussion implies that if we controlled for ability, the schooling coefficient would be unbiased. However, if the ability controls are imperfect, or if there is measurement error in the schooling variable, then the schooling coefficient may remain biased once ability is included. Griliches and Mason (1972) argue that measurement error in the schooling variable should lead to a downward bias in the schooling coefficient once ability is included as a regressor.

[22]The variance of schooling can, from (7), be expressed as

$$\sigma^2_{\log(S^*)} = \frac{1}{(d-b)^2}\left[k^2\sigma_A^2 + \sigma_P^2\right]$$

if ability and opportunity are independent.

[23]The partial regression coefficient (b_{AS}) from the regression of A on log S satisfies

$$E(b_{AS}) = \frac{\sigma_{\log(S^*),A}}{\sigma^2_{\log(S^*)}} = \frac{(d-b)k\sigma_A^2}{k^2\sigma_A^2 + \sigma_P^2}$$

return to ability (β_2) in this model -- will increase the bias in the usual schooling coefficient estimate (though it will also increase the variance of schooling, and the covariance between schooling and ability).[24] Finally, increases in the "true" marginal benefit of schooling (*i.e.*, b) will increase the usual schooling coefficient estimate. However, the increase in the estimate will be less than the increase in b, because the bias in the schooling coefficient falls as b increases.[25]

The conclusions reached in the previous paragraph follow under the assumption that ability and opportunity are uncorrelated. However, if we assume that wages directly affect MC, as in equation (10), and we assume $\sigma_{C,A}<0$, the results of the previous paragraph are essentially unchanged. The one exception is that changes in b will no longer affect the bias in the schooling coefficient. A summary of the effects of changes in the human-capital parameters on the bias in b_{wS} is provided in Table 2.

The human capital model offers several ways in which the observed wage/schooling relationship may change without there being changes in the "true" effect of schooling on wages. Of course, the extent to which the joint

---

which increases with increases in σ_A^2.

[24]For example, Juhn, Murphy, and Pierce (1990) have suggested that recent increases in the residual variance in wage equations for males reflect increases in the returns to ability. If so, one consequence could be that observed increases in the return to schooling overstate true increases in β_1.

[25]The bias falls because $E(b_{AS})$ falls as b increases. However,

$$\frac{\partial E(b_{wS})}{\partial b} = \frac{\sigma_P^2}{k^2\sigma_A^2 + \sigma_P^2} > 0 \quad ,$$

so that the usual coefficient estimate will increase when b increases.

distribution of schooling and ability may have changed over the 1980s is an empirical question. In the following section, we estimate models that attempt to control for the effects of ability on wages in order to ascertain the importance of omitted ability bias for recent increases in education-related earnings differentials.

III. **Empirical Analysis**

A. Data

The data we use come from the National Longitudinal Survey Youth Cohort. This cohort was first surveyed in 1979, when the respondents were between the ages of 14 and 22. They have been reinterviewed each year since 1979; we use data through the 1987 interview. The information extracted for each year includes wages on the current job, schooling status, labor market activity over the previous year, and industry, occupation, and union coverage on the current job. The 1979 interview also collected several variables associated with the family background of the respondent, which we use in our empirical analysis.

Most importantly, the data set includes scores of each respondent on the Armed Services Vocational Aptitude Battery (ASVAB) tests. Ten test scores are available, for a variety of cognitive and mechanical aptitudes. The test areas are: general science; arithmetic reasoning; mathematics knowledge; word knowledge; paragraph comprehension; mechanical comprehension; numerical operations; electronic information; auto and shop information; and coding speed. The ASVAB tests were administered to all survey respondents between the 1979 and 1980 surveys, with a 94 percent completion rate. The availability of these test scores, along with the time period over which the data were collected, make the NLS Youth Cohort a useful data set for studying changes in education-related earnings differentials for young white males in the early 1980s, and whether shifts in the ability-schooling relationship underlie these changes.[26]

There is more than one manner in which the NLSY data could be used to measure earnings differentials for young workers. One possibility is to estimate wage equations for each year from 1979 to 1987, using any respondent in the year who was working and was not in school.[27] One problem with this type of analysis is that the sample in the later years will be increasingly made up of workers more established in their labor market positions; signaling and learning models suggest that schooling may become less important for wages, and ability more important, as workers accumulate experience,[28] so that this analysis could confuse age (actually, experience) effects in the return to schooling with the desired period effects. To minimize this problem, we instead construct our sample so that we use only one wage for each respondent; the wage we choose is the first wage available after the respondent has completed his schooling (*i.e.*, the respondent does not return to school by the 1987 wave of the survey). With this sample, we hope to capture the schooling effects among workers competing for their first jobs, for it is among this group that the effects of relative demand and supply shifts should be most important. However, the restriction to using only one wage per respondent makes it impractical to try to carry out an analysis of bias in schooling

---

[26]The NLS Youth Cohort has also not suffered from sample attrition to the same degree as earlier longitudinal labor-market surveys; by 1987, roughly 90 percent of the original cohort was still responding to interview requests.

[27]Bishop (1989) uses a restricted version of this setup, in which coefficients in wage equations are allowed to vary along a linear trend over the sample period. Also, he does not restrict the sample to individuals who are out of school.

[28]See, *e.g.*, Harris and Holmstrom (1982) and Farber and Gibbons (1990). Indeed, Farber and Gibbons derive a further restriction that returns to ability will increase with experience, while returns to schooling will remain constant.

coefficients in regressions estimated separately for each year from 1979 to 1987, so we allow our schooling coefficients to vary along a linear time trend over the 1979-1987 period.[29]

Sample means and standard deviations for many of the variables used in our wage-equation estimation are presented in the first column of Table 3. The average age of our sample of workers on their first post-schooling job is fairly young, though the amount of labor-market experience (*i.e.*, hours worked in year-equivalent units) shows that on average our respondents had worked over 2 years before they enter the sample.[30] The educational-attainment statistics for our sample show slightly lower average education levels than other estimates for this cohort, largely due to the fact that some of the eventually more-educated members of this cohort are still in school in 1987.

In column (2) of Table 3, we report coefficient estimates from an individual-level regression of some of the variables on a constant and a time trend. These estimates show how the composition of the sample changes as we move through the 1979-1987 period. The wage variable we use is a measure of hourly earnings (in current dollars) on the primary job held at the time of the interview; the trend coefficient shows that this wage has increased by almost 8 percent per year over the 1979-1987 period. While part of this increase is due to inflation, the increase also reflects the fact that the individuals in the later years have a higher average level of education, have more experience at the time of the first post-schooling job observation, and are older at the time of the first job. Wages in the later years may also be

---

[29]Our sample size for white males is 2451. We exclude the self-employed, farm laborers, and respondents reporting a wage lower than one-half of the federal minimum wage prevailing in the year from which the observation is drawn.

[30]Of course, much of this experience may have been obtained in jobs held while still in school.

higher because returns to education (and experience) increased over the period.

Table 3 also reports sample statistics for averages of three subsets of the ASVAB test scores. Since the individuals in the cohort were of different ages when the tests were administered, age effects were removed from the scores by regressing each of the individual (normalized) test scores on a set of individual-year age dummies. The residuals from selected tests were then averaged to form the three composite test scores identified in Table 3. Following Bishop (1989), we dropped the coding speed test, and classified the remaining tests as either academic, technical, or computational; details of this classification are provided in the footnotes to the table. As our wage-equation estimates suggest that the technical and computational composites have very similar effects on wages, we also present sample statistics for the sum of these two composites (*i.e.*, the non-academic test). The trend coefficients for the test scores show that all three composites tend to be higher for those individuals whose first jobs were in the later years, with the increase over time largest for the academic test and smallest for the technical test.

B. Wage Equation Estimates

Using our hourly wage variable, we initially estimate equations of the form

$$(11)\quad \log(w_i) = \beta_1 S_i + \beta_2 (T_i S_i) + \beta_3 X_i + \beta_4 (T_i X_i) + \beta_5 Y_i + \epsilon_i \ ,$$

where w is the wage, S is years of schooling, T is the value of the time trend for the year in which the observation is taken, X is a vector of other factors that affect the wage, Y is a set of year dummies, and ϵ is an error term. The trend has a value of zero for the first year (1979), and increases by one for each following year. Including year dummies effectively controls for variation in wages due to inflation, productivity growth, or other cyclical factors. The other variables included in X are experience, age, a union membership dummy, a marriage dummy, and an urban dummy; the estimated trends in the coefficients for the marriage dummy and the urban dummy were essentially zero, so in our reported estimates we constrained these trends to be zero.

Ordinary least-squares estimates of β_1 and β_2 in equation (11) are reported in column (1) of Table 4. The results replicate the increased return to schooling in this period found in other data sets.[31] The estimates suggest that the linear return to schooling was .032 in 1979, and that this coefficient has increased by .0034 in each following year; by 1987, the estimate for the return to schooling is .059. This estimate for the increase in the return to schooling is somewhat larger than estimates suggested by previous studies, though other studies have used samples of workers that are older and more established than the workers in our sample.[32]

To further explore the robustness of the increased return to schooling in the NLSY, in the remaining columns of Table 4 we include other

---

[31]We also estimated separate wage equations for each year, using our sample of first wages. The schooling coefficient estimates (and standard errors) are:

1979	.021 (.011)
1980	.048 (.011)
1981	.039 (.011)
1982	.051 (.012)
1983	.041 (.012)
1984	.039 (.013)
1985	.062 (.015)
1986	.050 (.015)
1987	.050 (.021)

[32]While we do not focus on these coefficients in this paper, the OLS estimates show the coefficients for age and union membership to have declined over time, while the coefficient for experience increased.

education-related variables as regressors to pick up nonlinear effects, while continuing to interact the years-of-education variable with a time trend. The linear specification could lead to incorrect inferences concerning the sign or magnitude of the increase in the return to schooling if the true relationship between schooling and log wages is nonlinear.[33] This is potentially a serious problem in analyzing our sample, since the individuals whose first jobs are from the earlier years tend to be less educated than the individuals from the later years. For example, if the return to schooling were higher for college than high school, our finding that the return to schooling is higher in the later years studied could entirely be due to the fact that individuals from the later years have more years of college education.

In column (2) of Table 4, we add years of college as an additional regressor; the estimates suggest the return to college years is higher than to pre-college years, but including this variable only marginally reduces the years-of-education coefficient trend. Including years of high school along with years of college (column 3) reduces the education-coefficient trend by considerably more (though it provides an unlikely negative coefficient for high-school years). In column (4) we include a college graduate dummy, and in column (5) we include both a high-school graduate and a college graduate dummy; estimates of both specifications continue to provide evidence of an increasing return to schooling. In sum, nonlinear effects of schooling on log wages, combined with the nature of our sample, may explain some of the large increase in the return to schooling suggested in column (1), but even after

---

[33] One possible nonlinear relationship between wages and schooling is the log-linear model for wages suggested in section II. Empirical results using the log of schooling in place of the level of schooling in our wage-equation estimations uniformly provided poorer fits, so we continued to use the level of schooling. Conclusions about the trend in the education coefficient, however, were very similar for both schooling specifications.

controlling for these effects we continue to see a reasonably large rise in the schooling coefficient.

As discussed in Section II, all of the schooling coefficient estimates in Table 4, and in particular the years-of-education coefficient trend, potentially suffer from biases resulting from the error term being partly composed of individual abilities not captured in **X**, and from changes in the correlations between these abilities and schooling. In Table 5, we attempt to provide some idea of the importance of omitted-ability bias by including our test score measures as proxies for this omitted ability. In column (1) we include the individual's academic, technical, and computational test scores as independent variables. As the coefficient estimates for the technical and computational tests are very close, and the coefficient estimate for the academic test is negative and statistically insignificant, we estimated a specification that excludes the academic test, and includes the sum of the technical and computational tests; these results are in column (2). Both regressions provide highly significant coefficient estimates for the non-academic test scores, and inclusion of the test scores reduces the estimates for the schooling coefficient at any point in time (e.g., in column (1) the coefficient estimate for 1979 is .013, and for 1987 it is .051). However, the magnitude of the estimated increase in the schooling coefficient does not decline after including the test scores, but rather slightly increases.[34]

Simply including the test scores as regressors may not be the best way to use the information in these variables to control for "ability." It seems

---

[34]This result is due to the fact that, in our data set, the partial regression coefficient in the auxiliary regression of ability on schooling displays a (statistically significant) negative trend for each of the individual and composite test scores that we use.

reasonable to expect that the productive ability that employers value is at least partly reflected in our test scores, but that several other factors also affect the outcome of the tests (*e.g.*, test-taking ability, sleep the previous night, *etc.*). We might write this process as

$$TS_i = \gamma A_i + \upsilon_i \quad ,$$

where TS is the test score, A is ability rewarded in the labor market, and υ is other factors that affect the test score. If we assume that A and υ are uncorrelated, we have the classical errors-in-variables setup, suggesting that Table 5's OLS estimates of the test score coefficients, and coefficients for variables correlated with the test scores, are inconsistent. As a remedy, we assume that ability is correlated with the family background of the individual through the equation

$$A_i = \lambda F_i + \eta_i \quad ,$$

where F is a vector of family-background variables.[35] Instrumental-variable estimation of the wage equation when the test score is included as a regressor, using F as an instrument for TS, should eliminate the inconsistency in the wage-equation estimates resulting from measurement error in the test scores.[36]

---

[35] We assume that F and υ are uncorrelated. Prior research suggests that it is reasonable to exclude family-background variables as regressors in a wage equation (see the discussion in Blackburn and Neumark, forthcoming). Note that equations (7) and (8) imply that variables that determine P_i are obvious instruments for schooling, but not necessarily for test scores. While family background may seem most directly related to P_i, family-background variables may also be highly correlated with TS_i (and especially if A_i is directly determined by these variables).

[36] This method for controlling for unobserved ability was originally suggested by Griliches and Mason (1972); see also the surveys in Griliches (1977, 1979).

Schooling and test-score coefficient estimates from an estimation in which the test scores are treated as error-ridden are presented in columns (3) and (4) of Table 5. The family-background variables used as instruments are listed in the appendix table, along with the coefficient estimates in the first-stage regressions for the test scores.[37] The wage-equation estimates show the estimate of the "ability" effect to be larger than without instrumenting, as would usually be expected if the OLS coefficients suffered from measurement-error bias. The schooling coefficients also decline, for any given year, and appear to be essentially zero in the earliest years. But the increase in the return to schooling is slightly larger as a result of instrumenting for the test scores, again suggesting that omitted ability plays no role in explaining increases in the return to schooling. We also performed specification tests (suggested by Hausman, 1978) for the presence of measurement-error bias in the ability coefficient; the probability values for the null hypothesis of no measurement-error bias are also reported in Table 5. The specification tests actually suggest that instrumenting is not necessary for the test scores.

Our instrumental-variable estimates may also be inconsistent if the level of schooling is not exogenous with respect to the post-schooling wage. Models in which schooling decisions depend upon the wage (such as the human-capital model of section II, with a fixed effect in the wage-equation error not captured by the test scores), or measurement error in the schooling variable

---

[37]For observations in which a family-background variable is missing, we set the variable to zero; we also include dummy variables for each family-background variable being missing. This is essentially a first-order regression method for handling missing regressors. This method is likely to provide inconsistent coefficient estimates (see Kmenta, 1986), so the coefficient estimates we report in the appendix table are likely biased estimates of λ. However, the inconsistency in these estimates should not affect the consistency of our estimates of the wage equation.

(Griliches and Mason, 1972), suggest that schooling should be treated as potentially correlated with the wage-equation error. Given the young age of our sample, endogeneity is a potentially serious problem, since the wages we observe are likely highly correlated with those relevant to their schooling decisions. To explore this possibility, we use our family-background variables to instrument for both the non-academic test score and education; these results are reported in column (5) of Table 5.[38] This technique does affect the point estimates for the schooling and ability coefficients, but it also leads to a considerable increase in the standard errors associated with these coefficients; the increase in the return to schooling implied by the point estimates is even larger. If we instrument for schooling but not the test scores (column 6), we find similar results to those in column (5). Hausman tests provide no support for the joint hypothesis of endogeneity (or measurement error) in schooling and measurement error in ability, or for the simple hypothesis that schooling is endogenous or measured with error.

We were concerned that our results may be partially driven by the failure to adequately control for interactive effects among the determinants of the wage. For instance, there may be an interaction between education and ability in wage equations, *e.g.*, education may have a larger impact on the wages of more able workers.[39] This may be particularly important given that, in our sample, individuals observed in the later years have both higher test scores

---

[38] Rather than instrumenting for the education/trend interaction, we used the first-stage predicted value for education and interacted it with the trend variable in estimating the wage equation, since there is no reason to expect the trend variable to be correlated with the error term.

[39] An interaction between ability and log schooling would arise in the human-capital model if

$$MB(A_i,S_i) = \exp(kA_i)S_i^{b+hA_i}(1+b+hA_i) \quad .$$

and more schooling.

The first column of Table 6 presents estimates for the specification in the second column of Table 5 with an interaction between schooling and non-academic ability as a regressor.[40] The education/ability coefficient estimate is statistically significant, and the education/trend coefficient estimate is reduced considerably. This result suggests that our finding of an increase in the return to schooling may be partly due to a combination of an ability/education interactive effect and the fact that average levels of ability are increasing over time in our sample. However, the estimates in columns (2) and (3) suggest that this interactive effect is much less important if we allow the education/non-academic-test coefficient to vary over time, and also add interactions of the non-academic test score with age, experience and a trend. For example, in column (3) the education coefficient estimate still displays an upward (although statistically insignificant) trend, though there is also some (slight) evidence that the education/non-academic-test coefficient is increasing over time.

It is perhaps not surprising that there appears to be little evidence of an interactive effect of education and non-academic ability, since there is no clear reason why non-academic ability would be expected to increase the beneficial effects of education. In columns (4)-(6) of Table 6, we repeat the estimations of columns (1)-(3) using the academic test score in place of the non-academic. Including an interaction between the academic test and

---

[40] We estimated these specifications by OLS using the test scores as ability measures. This method of estimation is supported by the insignificant Hausman test statistics of Table 5 when using the test scores. Using different data, Blackburn and Neumark (forthcoming) found that instrumenting was necessary when using an IQ test score; the difference in findings may be due to the ASVAB test scores being less error-prone than the IQ test scores in the other data.

education provides a significant coefficient estimate for the interaction, and leaves the estimated increase in the schooling coefficient at essentially zero. This interactive effect of academic ability and schooling appears to primarily be present in the later years of our sample, as the test-score/education trend coefficient is significant in columns (5) and (6). If we include all interactions for both the academic and non-academic test scores in the same equation (column 7), the academic-test/education trend coefficient estimate is still large, but becomes statistically insignificant because of a much higher standard error. In column (8), we exclude all variables with clearly insignificant coefficient estimates, leaving the academic-test/education trend coefficient estimate virtually unchanged but with a much smaller standard error. While any hypothesis testing associated with column (8) does suffer from pretest bias, the t-statistic of 2.9 is rather large. In addition, in both columns (7) and (8), an estimated increase in the coefficient on education is no longer present.

To summarize the findings from Tables 5 and 6, the increase in the return to education over the 1980s persists when account is taken of the potential relationships between ability, schooling and wages. However, our investigation leads to a refinement of the finding: the increase in the return to education occurred for workers with relatively high academic ability. Existing estimates of the increase in the return to schooling, from data sets without ability measures, overstate the relative wage gains that education would have imparted to a randomly chosen (or marginal) worker.

Several checks of the robustness of our findings to changes in the sample used in estimating the wage equation were performed, and the results are reported in Table 7. The first two columns report the OLS estimate of the education coefficient and its trend in specifications without ability controls. The next two columns report these coefficients' estimates from specifications that include the test-score interactions included in column (8) of Table 6. We also report the academic-test/education trend coefficient estimate, and the estimate for the coefficient on the non-academic test score.

We tried two alternative restrictions in selecting the sample. First, we excluded any individual with labor market experience greater than three years, to avoid using individuals who may be firmly ensconced in the labor force and therefore not competing for a new job, and to reduce the potential confounding influence of learning. Second, we excluded observations whose first job observations were in 1982 or 1983, to enhance the comparison between the low schooling-return and high schooling-return periods, and to eliminate severe recession years. Both redefinitions of the sample lead to similar conclusions as with the full sample, with the sample with the maximum-experience restriction providing a larger estimated increase in the academic-test/education interaction effect than was provided by the full sample. The fourth row of Table 7 presents results from estimations that include a college graduation dummy; again, conclusions are essentially unchanged.

We also estimated wage equations that include twelve industry and eleven occupation dummies as regressors; these results are reported in the bottom three rows of Table 7. These estimates suggest that occupational shifts (but not industry shifts) can explain a considerable portion of the schooling return increase if we omit ability from the specifications.[41] Occupation (and to a lesser extent industry) shifts also appear to account for at least part

[41]This findings differs from that of Blackburn, Bloom, and Freeman (1990a), who found that changing industrial composition of employment explains up to 25 percent of the increase in education-related earnings differentials, but that occupational changes played no role in the schooling-return increase.

of the increase in the interactive effect of academic ability and education, suggesting that at least part of the increased importance of education to wages for high academic-ability males is due to demand shifts towards occupations (and industries) that tend to employ high-education/high-ability individuals.

IV. Summary

Much attention has been paid to explaining recent increases in the return to schooling among males in the U.S. Estimates of these increases are generally obtained from wage regressions that are potentially biased by the presence of "unobserved" ability in the wage-equation error. We have shown that both a signaling model and a human capital model suggest that changes in the relationship between ability and schooling could underlie the increases in the schooling return. We offer evidence on the plausibility of these explanations, with a particular focus on using test scores as a proxy for ability in wage regressions. Our results provide little or no support in favor of the hypothesis that the increases in the return to schooling reflect an increased upward bias in the schooling coefficient estimate due to a change in the ability-schooling relationship. But our results do provide an interesting refinement of the stylized fact that education returns have been increasing in the 1980's -- the increase in the return to education has occurred largely for workers with higher levels of "academic" ability.

What can explain an increase in the return to education for high-ability workers only? Supply-side explanations can be constructed. For example, if it were the case that education and ability were becoming *less* correlated over time, then there would be relatively fewer of those workers with both high levels of education and ability; also, if the average level of ability were to fall, this could create a growing scarcity of high-education, high-ability workers. But at present it is difficult to assess the existence or importance of such supply-side changes.[42] While occupation shifts appear to be of some importance, what is causing these shifts is still an open question. Skill-biased technical change is one possibility, though the evidence in favor of this argument is still limited.[43]

---

[42] In fact, Bishop (1989) refers to results that suggest that average test scores of young individuals were increasing over the 1980s.

[43] Davis and Haltiwanger (1990) offer evidence on changes in wage dispersion across manufacturing plants that, they argue, supports the skill-biased technical change hypothesis.

REFERENCES

Arrow, Kenneth. 1973. "Higher Education as a Filter." *Journal of Public Economics*. 2: 193-216.

Becker, Gary. 1975. *Human Capital*. Second Edition. Chicago: University of Chicago Press.

Behrman, Jere, Robert Pollak and Paul Taubman. 1990. "The Impact of Ability and Parental Income on Wages and Schooling." Unpublished paper, University of Pennsylvania.

Bishop, John. 1989. "Achievement, Test Scores, and Relative Wages." Unpublished paper, Cornell University.

Blackburn, McKinley, David Bloom, and Richard Freeman. 1990a. "The Declining Economic Position of Less-Skilled American Men." In *A Future of Lousy Jobs? The Changing Structure of U.S. Wages*. Edited by Gary Burtless. Washington, D.C.: The Brookings Institution.

________. 1990b. "An Era of Falling Earnings and Rising Inequality?" *The Brookings Review*. Winter 1990-1991: 38-43.

Blackburn, McKinley and David Neumark. "Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials." Forthcoming in *Quarterly Journal of Economics*.

Davis, Steve J. and John Haltiwanger. 1990. "Wage Dispersion Between and Within U.S. Manufacturing Plants, 1963-1986." Unpublished paper, University of Chicago.

Farber, Henry, and Robert Gibbons. 1990. "Learning and Wage Determination." Unpublished paper, Massachusetts Institute of Technology.

Griliches, Zvi. 1977. "Estimating the Returns of Schooling: Some Econometric Problems." *Econometrica*. 45: 1-22.

________. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy*. 87: S37-64.

Griliches, Zvi, and William Mason. 1972. "Education, Income and Ability." *Journal of Political Economy*. 80: S74-103.

Harris, Milton and Bengt Holmstrom. 1982. "A Theory of Wage Dynamics." *Review of Economic Studies*. 44(3): 315-333.

Hausman, Jerry A. 1978. "Specification Tests in Econometrics." *Econometrica*. 46(6): 1251-1271.

Heckman, James and Bo Honoré. 1990. "The Empirical Content of the Roy Model." *Econometrica*. 58 (5): 1121-1149.

Juhn, Chinhui, Kevin Murphy, and Brooks Pierce. 1990. "Wage Inequality and the Rise in Returns to Skill." Unpublished paper, University of Chicago.

Katz, Lawrence, and Ana Revenga. 1989. "Changes in the Structure of Wages: The United States vs Japan." *Journal of the Japanese and International Economies*. 3: 522-553.

Kmenta, Jan. 1986. *Elements of Econometrics*. Second Edition. New York: MacMillan.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings*. New York: Columbia University Press.

________. 1991. "Human Capital, Technology, and the Wage Structure." Unpublished paper, Columbia University.

Spence, Michael. 1973. *Market Signalling: Information Transfer in Hiring and Related Processes*. Cambridge, Mass.: Harvard University Press.

Stapleton, David C. and Douglas J. Young. 1988. "Educational Attainment and Cohort Size." *Journal of Labor Economics*. 6(3): 330-361.

Taubman, Paul and Terence Wales. 1972. "Mental Ability and Higher Educational Attainment in the 20th Century." National Bureau of Economic Research Occasional Paper 118.

Table 1
Actual and Predicted Changes in Logarithmic Differentials
for White Males in 1973, 1979, and 1987

Comparison	Differential 1973	1979	1987	Δ in Differential '73 to '79	'79 to '87
A: Actual*					
CG to LTHS	.47	.46	.63	-.01	.17
HS to LTHS	.19	.20	.25	.01	.05
CG to HS	.28	.26	.38	-.02	.12
B: Standard Normal					
CG to LTHS	.47	.48	.50	.01	.02
HS to LTHS	.20	.20	.20	0	0
CG to HS	.27	.28	.30	.01	.02
C: Triangular					
CG to LTHS	.47	.47	.49	0	.02
HS to LTHS	.20	.19	.19	-.01	0
CG to HS	.27	.28	.30	.01	.02

*These numbers are taken from Table 1 in Blackburn, Bloom, and Freeman (1990a), and refer to workers aged 25-64. CG refers to college graduates, HS to high school graduates, and LTHS to workers with less than a high school degree.

Table 2
Summary of Results from the Human-Capital Model

$$MB = \exp(kA_i)S_i^b$$

$$MC = \exp(C_i)W_i$$

$$\sigma_{C,A} < 0$$

Parameter	Sign of partial derivative of bias with respect to parameter
k	+
b	0
σ_A^2	+
σ_C^2	-
$\lvert\sigma_{C,A}\rvert$	+

Table 3
Descriptive Statistics for First Post-Schooling Labor Market Observation

	Mean (Standard deviation)	Trend[1]
Log wage	1.616 (.451)	.077 (.003)
Years of education	12.729 (2.449)	.558 (.016)
High school graduate (12 years)	.423	...
College graduate (16 years)	.156	...
Experience	2.180 (1.754)	.395 (.012)
Age	20.975 (2.740)	.685 (.017)
Married, spouse present	.166	...
Urban	.710	...
Union	.176	...
Number of observations	2451	...
1979	463	...
1980	315	...
1981	309	...
1982	341	...
1983	287	...
1984	228	...
1985	185	...
1986	168	...
1987	155	...
Test scores (age-neutral):[2]		
Academic test[3]	.012 (.857) [-.61,.69]	.140 (.006)
Technical test[4]	-.010 (.840) [-.54,.61]	.065 (.007)
Computational test	.010 (.976) [-.65,.77]	.119 (.008)
Non-academic test (technical + computational)	.000 (1.541) [-.91,1.16]	.184 (.012)

1. Coefficient from regression on intercept and time trend. The time trend is defined as zero in 1979. Standard errors reported in parentheses.
2. Residuals from regressions of normalized test scores on individual year age dummy variables. Lower and upper quartiles are reported in square brackets.
3. Average of residuals for tests of arithmetic, mathematics, word knowledge, paragraph comprehension, and general science.
4. Average of residuals for tests of auto and shop knowledge, electronics, and mechanical knowledge.

Table 4
OLS Log Wage Equation Estimates[1]

	(1)	(2)	(3)	(4)	(5)
Years of education	.032 (.007)	.029 (.008)	.054 (.014)	.028 (.008)	.030 (.008)
Years of education × trend	.0034 (.0017)	.0030 (.0017)	.0018 (.0018)	.0027 (.0017)	.0023 (.0018)
Years of high school	...	...	−.013 (.006)	...	...
Years of college	...	.002 (.002)	.001 (.002)	...	...
High school graduate	...	...	...	...	−.011 (.017)
College graduate	...	...	....	.080 (.024)	.075 (.025)
$\bar{R}^2$	.393	.393	.394	.396	.396

1. Specifications also include experience, age and union status (each interacted with a time trend defined as zero in 1979), dummy variables for urban residence and married, spouse present, and an intercept and single year dummy variables. Standard errors are reported in parentheses.

Table 5
OLS and IV Log Wage Equation Estimates,
Including Test Scores with Constant Coefficients[1]

	OLS		IV for Test Scores		IV for Test Scores and Schooling	IV for Schooling
	(1)	(2)	(3)	(4)	(5)	(6)
Years of education	.013 (.008)	.012 (.008)	-.000 (.013)	-.001 (.013)	.031 (.035)	.020 (.016)
Years of education × trend	.0048 (.0017)	.0048 (.0017)	.0062 (.0023)	.0057 (.0018)	.0064 (.0034)	.0095 (.0031)
Academic test	-.010 (.017)	...	-.057 (.152)	...	...	...
Technical test	.044 (.013)	...	...	...	...	...
Computational test	.041 (.010)	...	...	...	...	...
Non-academic test	...	.038 (.006)	.094 (.081)	.064 (.020)	.025 (.043)	.028 (.009)
$\bar{R}^2$	.404	.409	...	...	...	...
Measurement error/ endogeneity tests: p-value[2]	...	...	.331	.170	.418	.136

1. Specifications also include experience, age and union status (each interacted with a time trend defined as zero in 1979), dummy variables for urban residence and married, spouse present, and an intercept and single year dummy variables. Standard errors are reported in parentheses. Instrumental variables are listed in Table A1.
2. P-value from F-test of significance of coefficients of residuals from first-stage instrumental variables regressions, in log wage equation estimated with OLS.

Table 6
OLS Log Wage Equation Estimates,
Alternative Trend and Interactive Specifications[1]

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years of education	.026 (.009)	.025 (.009)	.019 (.009)	.031 (.009)	.029 (.009)	.025 (.009)	.025 (.009)	.023 (.009)
Years of education × trend	.0015 (.0018)	.0015 (.0018)	.0022 (.0019)	.0001 (.0019)	-.0003 (.0020)	-.0001 (.0021)	-.0003 (.0021)	.0004 (.0020)
Non-academic test	.038 (.006)	.036 (.006)	.052 (.009)	...	...	...	.067 (.016)	.061 (.012)
Non-academic test × years of education	.009 (.002)	.006 (.003)	.002 (.004)	...	...	...	.005 (.007)	...
Non-academic test × years of education × trend	...	.0008 (.0008)	.0010 (.0008)	...	...	...	-.0005 (.0015)	...
Non-academic test × trend	...	...	-.0041 (.0028)	...	...	...	-.0057 (.0050)	-.0037 (.0027)
Non-academic test × experience	...	...	-.005 (.004)	...	...	...	-.002 (.006)	...
Non-academic test × age	...	...	.010 (.003)	...	...	...	.006 (.005)	.006 (.003)
Academic test	...	...	...	.048 (.011)	.043 (.012)	.058 (.020)	-.043 (.032)	-.021 (.017)
Academic test × years of education	...	...	...	.015 (.004)	.005 (.007)	-.003 (.007)	-.004 (.014)	...
Academic test × years of education × trend	...	...	...	...	.0029 (.0014)	.0032 (.0015)	.0034 (.0028)	.0029 (.0010)
Academic test × trend	...	...	...	...	...	-.0022 (.0054)	.0068 (.0094)	...
Academic test × experience	...	...	...	...	...	-.012 (.007)	-.007 (.011)	...
Academic test × age	...	...	...	...	...	.019 (.005)	.009 (.009)	.009 (.007)
$\bar{R}^2$	.409	.409	.411	.402	.403	.405	.413	.414

1. Specifications also include experience, age and union status (each interacted with a time trend defined as zero in 1979), dummy variables for urban residence and married, spouse present, and an intercept and single year dummy variables. Standard errors are reported in parentheses.

Table 7
Schooling and Schooling × Ability Coefficients and Trend Interactions for Alternative Samples and Specifications[1]

Specification:	Years of education coefficients				Years of education × academic test × trend coefficient	Non-academic test score coefficient
	Table 4 (1)		Table 6 (8)		Table 6 (8)	Table 6 (8)
	Linear	Trend	Linear	Trend	Trend	Linear
Results from earlier tables	.032 (.007)	.0034 (.0017)	.023 (.009)	.0004 (.0020)	.0029 (.0010)	.061 (.012)
Maximum experience of three years (N=1888)	.034 (.008)	.0041 (.0020)	.027 (.009)	.0008 (.0024)	.0045 (.0012)	.064 (.012)
1982 and 1983 excluded (N=1823)	.030 (.008)	.0036 (.0017)	.018 (.009)	.0009 (.0021)	.0026 (.0011)	.058 (.013)
College graduation dummy variable included	.028 (.008)	.0027 (.0017)	.018 (.009)	-.0001 (.0020)	.0030 (.0010)	.064 (.012)
One-digit industry dummy variables included (N=2297)	.038 (.007)	.0033 (.0016)	.028 (.009)	.0009 (.0019)	.0023 (.0010)	.052 (.012)
One-digit occupation dummy variables included (N=2297)	.029 (.008)	.0013 (.0017)	.021 (.009)	-.0014 (.0018)	.0017 (.0010)	.045 (.012)
One-digit industry and occupation dummy variables included (N=2297)	.032 (.007)	.0016 (.0016)	.022 (.008)	.0003 (.0019)	.0014 (.0010)	.047 (.011)

1. See footnotes in corresponding Tables 3-6.

Appendix Table

Coefficients of Instrumental Variables in First-Stage Regressions[1]

	(1) Non-Academic Test[2]	(2) Academic Test[2]	(3) Schooling
Magazines in home (age 14)	.231 (.060)	.149 (.030)	.484 (.073)
Newspapers in home (age 14)	.311 (.077)	.108 (.038)	.152 (.095)
Library card in home (age 14)	.067 (.061)	.076 (.030)	.145 (.075)
Father's education (1979)	.029 (.010)	.015 (.005)	.080 (.012)
Mother's education (1979)	.072 (.013)	.036 (.006)	.094 (.016)
Number of siblings (1979)	-.047 (.019)	-.037 (.010)	-.151 (.022)
Number of older siblings (1979)	.015 (.024)	.015 (.012)	.086 (.029)
Highest grade of oldest sibling (1979)	.029 (.014)	.012 (.007)	.105 (.017)
Foreign language spoken in home (age 14)	.011 (.082)	-.052 (.041)	.177 (.100)
Father and mother in home (age 14)	-.062 (.082)	-.029 (.040)	.387 (.100)
No adult male in home (age 14)	-.024 (.106)	-.021 (.052)	.187 (.130)
$\bar{R}^2$	.368	.502	.625

1. Coefficients are reported for first-stage regressions using linear schooling. Specifications also include intercepts, single year dummy variables, all other variables included in specifications of wage equations in Table 2, and dummy variables for each of these set equal to one when data were missing on the instruments (in which case the variables were set equal to zero). (We distinguish between the highest grade of oldest sibling missing in the usual sense, and missing because the respondent is the oldest sibling.) Standard errors are reported in parentheses.
2. Specification includes schooling and its trend interaction.