NBER Working Paper Series

EDUCATION: CONSUMPTION OR PRODUCTION

by

Edward Lazear*

Working Paper No. 104

CENTER FOR ECONOMIC ANALYSIS OF HUMAN BEHAVIOR
AND SOCIAL INSTITUTIONS

National Bureau of Economic Research, Inc.
204 Junipero Serra Boulevard, Stanford, CA 94305

*University of Chicago and National Bureau of Economic Research

September 1975

Preliminary; not for quotation.

Financial support from the National Institute of Mental Health and the Rockefeller Foundation is gratefully acknowledged.

Human capitalists often find themselves confronted with the charge that the human capital framework, although consistent with empirical results, does not appropriately capture the true direction of causality. In particular, Spence (1972) has argued that education may simply provide a signal to employers of the applicant's ability and may not actually increase the individual's productivity.[1] Although the signalling hypothesis has implications with respect to the social returns to education which differ from those of the traditional human capital story, the implications for the individual are essentially the same. Regardless of whether schooling is a screen or a truly productive asset, it is still traditional for the individual investor to acquire formal schooling up to the point where the present value of additional education is equal to the cost of its acquisition. There is, however, another motivation for the connection between schooling and wealth which has implications at the level of the individual different from those traditionally ascribed to human capital. Specifically, it can be claimed that education is simply a normal consumption good and that like all other normal goods, an increase in wealth will produce an increase in the amount of schooling purchased. Increased incomes are associated with higher schooling attainment as the simple result of an income effect. Bowles (1972), for example, finds that 52% of the variation in levels of schooling can be explained by family background variables. He argues that the exclusion of these variables from earnings functions causes an upward bias in the estimated returns to schooling.[2] If this is so, schooling increases an individual's wealth only by the consumption value of the good, since it is a non-saleable

---

The author wishes to thank Gary Becker, Charlie Brown, William Brock, Gary Chamberlain, Victor Fuchs, Zvi Griliches, James Heckman, Lee Lillard, Robert Lucas, José Scheinkman, Theodore Schultz and Kenneth Wolpin for useful suggestions and David Feldman, Kyle Johnson, and Maureen Lahiff for computer assistance.

asset. This paper will attempt to determine empirically the amount by which an increase in wealth is caused by schooling as distinguished from the amount by which the demand for schooling increases as the result of an increase in wealth.

The problem then is to estimate the amount of education that would be acquired in the absence of its consumptive benefits.[3] There are two conceptual ways to do this; one can measure the difference between the total amount of education acquired and that amount which would be acquired in the absence of the wealth-increasing properties of education. The residual can then be called the investment component. This approach, for one thing, requires that all uses of education as a consumption good be known. A considerable amount of research has been done on the use of education in non-market activities. Michael (1973), for example, claims that more educated individuals consume other commodities as if they have higher wealth. Libebowitz (1974) finds that an increase in a mother's level of education is associated with a higher measured IQ for her child. In the areas of the family, Michael (1973a) has found that more educated individuals are better contraceptors and are thus less likely to experience an unwanted birth. Benham (1974) finds that an increase in a woman's education is associated with higher earnings on the part of her spouse, given his education. Education thus can have an impact on an individual's health as well. Williams (1974) finds that a child's birthweight (which is generally positively related to its chance of survival) is also positively associated with parental education, given family income. Additional examples can be given, but the point is clear - education affects many aspects of an individual's behavior in addition to its impact on his earnings capacity. Yet it would be extremely difficult to attempt to estimate the total impact that education has on one's ability to consume. Nor, as it turns out, would this assist us in disentangling consumption from investment components of education.

There have also been attempts to establish the existence of an education effect on worker productivity. Welch (1970, 1973) and Griliches (1964, 1968), among others, have argued that education plays an important role in the formation of and adaptation to technological change. An alternative approach is to treat the problem as one of finding the difference between the wealth-maximizing and utility-maximizing levels of education. In the model developed below, this is found to be a more conceptually appropriate and empirically useful formulation of the problem.

A Model

Let us start by defining the opportunity set of values of goods consumed (including leisure), X^c, and education "consumed", E^c, by the individual. An individual starts out with an endowment $(X,E) = (X^e,0)$. He may convert resources into education according to the following relationship, $E = f(X^p)$ where X^p is the amount of resources used in the production of schooling (measured in terms of grades completed). $f(X^p)$ is assumed to be monotonic, continuous, and differentiable throughout. If X is the numeraire commodity, then the individual may "sell" a unit of education for the present price, P_E goods, where "present" is defined as the date of birth. The individual sells his education by working at jobs which pay higher wages to more educated individuals. It is assumed that the selling price of all units of schooling are the same.[4] The individual then maximizes his wealth by producing education up to the point, E^*, where

$\frac{1}{f'(X^P)} = P_E$ (since $\frac{1}{f'(X^P)}$ is the marginal cost of producing a unit of education). This can be seen on Figure 1.

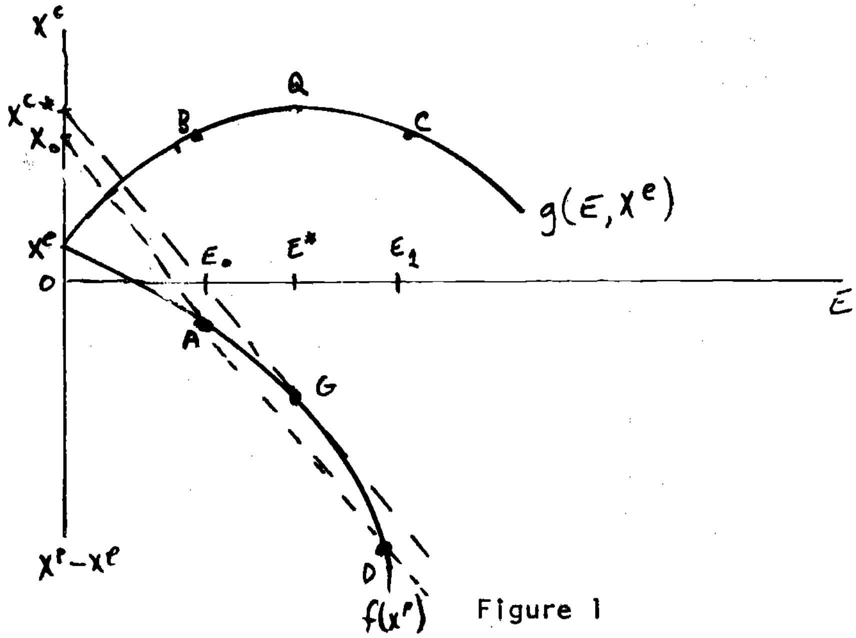

Figure 1

Education is plotted along the horizontal axis and goods are plotted on the vertical axis. If the individual starts with an endowment at X^e, he produces education by moving along the curve, $f(X^P)$. The individual's ex post measured wealth position is given by the intersection of the line through G (with slope = $-P_E$) with the vertical axis. Thus, if the individual maximizes his measured wealth, he can consume X^{c*} goods. However, he also consumes E^* of education because the "sale" of education does not preclude the simultaneous use of education in consumption. The amount of education produced and sold equals the amount consumed so that $E^c = E$.[5] If the individual were to produce E_0 education, his opportunity to purchase goods would be given by X_0 and his consumption of education would equal E_0 so that he would end up at point B. It is clear that Q is superior to B and all other points along X^e Q as long as both X and

E are goods. However, if the individual were to produce E_1 units of education, he would be able to consume at C. If E is a good, no unambiguous statement can be made with respect to the superiority of any point between Q and C to C since this essentially depends upon the individual's willingness to trade consumption of goods for consumption of education.[6]

One may describe the opportunity set through X^eBQC by

$$(1) \qquad X^c = h(X^p, X^e) = X^e + [f(X^p)]P_E - X^p$$

Since $E = f(X^p)$ is monotonic and continuous, there exists an inverse function $X^p = f^{-1}(E)$ so that (1) can be rewritten as

$$(2) \qquad X^c = g(E, X^e) = X^e + EP_E - f^{-1}(E) \ .$$

If E enters the utility function at all, the consumptive optimum can occur at Q only if there is a kink in $g(E,X^e)$ at point Q. But since $f(X^p)$ is differentiable and non-zero, the inverse of $f(X^p)$ is also differentiable. Thus, $g(E,X^e)$ is differentiable throughout so that no kink can occur at Q. This simply states that as long as education enters the utility function (either positively or negatively), there exists some price at which the individual is willing to trade some goods for more, or in the case where E enters negatively, less education.[7] Thus, the separation principle is invalidated because one of the factors of wealth production, namely education, enters the utility function directly. The individual necessarily acquires an amount of education different than the wealth-maximizing level.

The problem facing the consumer, then, is to maximize utility given by $U = U(X^c,E)$ subject to the constraint that $X^c = g(E)$. Therefore,

(3) $\mathcal{L} = U(X^c,E) - \lambda[X^c - g(E)]$

so that the first order conditions for an optimum are given by

(4) (a) $U_{X^c} - \lambda = 0$

(b) $U_E + \lambda g' = 0$

(c) $X^c - g(E) = 0$

or,

(5) (a) $U_E + U_X g' = 0$

(b) $X^c - g(E) = 0$

The importance of the conclusion that the optimal level of schooling acquisition differs from the wealth maximizing level depends, of course, on the magnitude of the effect. If U_E is close to zero for all levels of E, then most of education's effect on behavior works through its ability to alter wages. On the other hand, it may be the case that P_E is close to zero. If so, education is acquired primarily because it has a positive (direct) effect on the individual's utility. In order to ascertain the relevant magnitudes, it is necessary to parameterize the functions. Thus, let

(6) $U(X^c,E) = X^c e^{\Theta E}$

and let

(7) $E = f(X^p) = (AX^p)^\gamma$

where A is a scalar which converts units.

The choice of the functional forms is not arbitrary. The exponential form of (7) is natural since it yields $f(0) = 0$, $f' > 0$ and $f'' < 0$ for $0 < \gamma < 1$. The form of the utility function is less obvious. The main reason for its choice is that it has the very unusual property that $\frac{d^2X}{dE^2} = \Theta^2 X \geq 0$ for all values of Θ. This implies that an interior solution is always obtained. Since we observe that individuals consume positive levels of both X^c and E, this property is indeed desirable (although institutional constraints may account for the failure to observe corners to some extent).[8]

From equation (2),

$$(8) \qquad g(E,X^c) = X^e + P_E E - \frac{1}{A} E^{1/\gamma} .$$

The wealth maximizing level of education, E^*, is found when $g'(E,X^e) = 0$, or when

$$(9) \qquad g' = P_E - \frac{1}{\gamma A} E^{\frac{1-\gamma}{\gamma}} = 0$$

so that

$$(10) \qquad E^* = (A\gamma P_E)^{\frac{\gamma}{1-\gamma}}$$

The wealth-maximizing level of education, E^*, is independent of initial wealth because the way by which an individual will maximize his wealth does not vary with initial endowment in the absence of capital constraints. From (5a) the optimization condition is

$$(11) \qquad \Theta X^c = \frac{1}{\gamma A} E^{\frac{1-\gamma}{\gamma}} - P_E$$

$$= \frac{X^p}{\gamma E} - P_E$$

and upon substituting 5(b) into (11) one obtains

$$\Theta(X^e + P_E E - X^p) = \frac{X^p}{\gamma E} - P_E \tag{12}$$

or

$$E = \frac{1}{P_E}(X^p - X^e) + (\frac{1}{\Theta})(\frac{1}{P_E})(\frac{X^p}{\gamma E}) - \frac{1}{\Theta} \tag{13}$$

Equation (13) identifies P_E, and could be estimated by OLS except that E appears on both sides of the equation. This by itself is easily treated, but before pursuing that, a few points should be made.

First, the model gives a framework in which observed levels of schooling can differ across individuals even though their ability levels do not.[9] This is an important aspect of the model; it does not suffer from the drawback that all individuals are assumed to be alike and yet obtain different levels of education. The reason that individuals differ in their levels of schooling, at least up to this point, is that their endowed wealth differs and their post-schooling wealth therefore differs.

Additional variation in schooling may result when it is recognized that endowments may affect the cost of obtaining schooling. In particular, any aspect of an individual's endowment which affects his wage rate will certainly alter the cost of spending time at school. As found below (see Table 1), "endowed" characteristics which are important in the determination of wage rates are endowed IQ, mother's level of education and a white/non-white dummy. Thus rewrite (7) as

$$E = (AX^p)^{\gamma} K^{\delta} M^{\eta} e^{\lambda D} \tag{14}$$

where K is a measure of endowed IQ (defined below)

M is the highest grade of schooling completed by the mother

D is a dummy equal to one for whites.[10]

Equation (14), it should be remembered, is essentially an identity which relates the number of years of education to cost of acquiring them. Thus, δ, for example, measures the effect of IQ on the costs of producing education as it works by changing the cost of time. For a given dollar expenditure, higher ability individuals produce fewer _years_ of education since time is more valuable to them. Any correct measure of X^P will reflect this variation.

If endowment variables affect the costs of schooling by making individuals more productive workers, it is also reasonable to expect that their ability to learn productive skills will be affected by these same variables. More able individuals, for example, are expected to acquire more units of human capital per year spent at school than less able ones. Therefore, let the present price of a unit of education vary with the endowment variables so that $1/P_E$ in (13) becomes

$$\frac{1}{P_E} \equiv \beta_1 + \beta_1 K + \beta_2 M + \beta_3 D$$

where the β_i are unknown parameters.[11]

So far, it has been assumed that all expenditure on education is financed by the individual. However, some parents clearly contribute toward financing their child's education. That part of parental aid which is contingent upon the child's attending school reduces the cost of schooling to the child by that same amount.[12] Although it is only the child's expenditure on schooling X^P, which directly enters the calculation of the opportunity set [equations (1) and (2)], it is total expenditure by parent and child which is appropriate for the education production function. The X^P term of (14) should then be replaced by $\left(\frac{X^P}{1-F}\right)$ where F is the proportion of total schooling expenditure financed by parental subsidy (i.e., total expenditure times (1-F) equals X^P, the expenditure of the child).[13]

Incorporating the changes, (14) becomes

$$(15) \quad E = \left(\frac{AX^P}{1-F}\right)^{\gamma} K^{\delta} M^{\eta} e^{\lambda D}$$

so that (8) is rewritten as

$$(16) \quad g(E,X^e,K,M,D) = X^e + [\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D]^{-1} E - \frac{1-F}{A} E^{\frac{1}{\gamma}} K^{\frac{-\delta}{\gamma}} M^{\frac{-\eta}{\gamma}} e^{\frac{-\lambda D}{\gamma}} .$$

Equation (9) then becomes

$$(17) \quad \frac{\partial g}{\partial E} = (\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)^{-1} - \frac{1-F}{\gamma A} E^{\frac{1}{\gamma}} K^{\frac{-\delta}{\gamma}} M^{\frac{-\eta}{\gamma}} e^{\frac{-\lambda D}{\gamma}}$$

so that (10) becomes

$$(18) \quad E^* = [(\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)^{-1} (\frac{A\gamma}{1-F} K^{\frac{\delta}{\gamma}} M^{\frac{\eta}{\gamma}} e^{\frac{\lambda D}{\gamma}})]^{\frac{\gamma}{1-\gamma}}$$

Finally, the identifying equation (13), can be rewritten as

$$(19) \quad E = \beta_0(X^P-X^e) + \beta_1 K(X^P-X^e) + \beta_2 M(X^P-X^e) + \beta_3 D(X^P-X^e)$$
$$+ \frac{\beta_0}{\theta}\left(\frac{X^P}{\gamma E}\right) + \frac{\beta_1 K}{\theta}\left(\frac{X^P}{\gamma E}\right) + \frac{\beta_2 M}{\theta}\left(\frac{X^P}{\gamma E}\right) + \frac{\beta_3 D}{\theta}\left(\frac{X^P}{\gamma E}\right) - \frac{1}{\theta} .$$

The term γ in equation (19) is a technological and not a behavioral parameter. It is determined exogenously from information on the costs of attending school and should not be thought of as estimable from (19). Theoretically, one could solve algebraically for γ, δ, η and λ simply by knowing how the endowment variables affect wages and therefore the foregone earnings component of school costs. However, the computer can be used more easily to find an appropriate solution for γ, δ, η, and λ. Rewrite (15) as

$$(20) \quad \ell n\ E = \gamma\ \ell n\ A + \gamma\ \ell n\left\{\frac{X^P}{1-F}\right\} + \delta\ \ell n\ K + \eta\ \ell n\ M + \lambda D\ .$$

Once values of E, K, M and D are selected, X^P is determined by the technological relationship between schooling costs and foregone earnings (described in more detail below). Any five arbitrarily chosen, linearly independent, vectors of E, K, M, and D will therefore allow one to solve for γ, A, δ, η, and λ. The extent to which the solutions will vary with different arbitrary sets of vectors depends upon how good an approximation the exponential production function is to the true production function. One may go further and select any T number of arbitrary vectors and solve for the parameters by least-squares. As R^2 approaches 1, the exponential approximation becomes exact and invariant to subsets of different arbitrary vectors. This is the procedure followed below. R^2 is quite high and it was found that the estimates of A, γ, δ, η, and λ were insensitive to different sets of arbitrary vectors.

Finally, β and $\frac{1}{\theta}$ can be estimated from (19) directly, and although overidentified, maximum-likelihood techniques are easily applied to obtain unique estimates.

Again it is clear that P_E can conceivably be close or equal to zero. That is, the wealth-increasing component of education can be small or non-existent. In that case, education should be regarded simply as a consumption good, the correlation between education and income being purely the result of an income effect. The framework therefore does what it sets out to do; namely, it isolates the wealth-increasing effect of education. In addition, since the parameters can be identified, one can solve for E^*. Thus $(E - E^*)$, the amount of education acquired in excess of the wealth-maximizing level, is also identified.[14]

Estimation Procedure

Before actually estimating (19), a good deal of data manipulation must be done to resolve definitional questions. The first part of this section is devoted to this methodologically cumbersome task. The data used come from the National Longitudinal Survey (1966-69) on young men, 14-24 years of age in 1966. When estimating (19), a subsample was selected such that no individual was currently attending school, nor had he attended during the previous two years.[15] E is therefore defined as the highest grade of schooling completed.[16]

Direct information on IQ is not available for each individual. However, all persons in the survey were given a test which measures their "knowledge of the world of work," (KWW). Scores range from zero to 56 with a mean of 32.6 and a standard deviation of 9.2. Griliches (1974) reports that this variable is highly correlated with IQ and in many cases, performs better.[17] However, it is still true that measured KWW may not be an accurate representation of endowed ability. Since it is the endowed values which are appropriate in a theoretical sense, $\widehat{KWW}$ rather than KWW is used as a measure of K where $\widehat{KWW}$ is constructed as follows:

One would like to know what the IQ level (even measured IQ level) would have been in the absence of schooling. That more highly schooled individuals tend to have higher IQ's is, in part, attributable to the fact

IQ tests often measure what is taught in school.[18] One may estimate the effect that schooling has on measured KWW by specifying the following relationship:[19]

$$\text{(21)} \quad KWW = \alpha_0 + \alpha_1 S_{66} + \alpha_2 (S_{66})^2 + \alpha_3 D + \alpha_4 M + \alpha_5 F + \alpha_6 FI$$

where KWW is the 1966 test score

S_{66} is the highest level of schooling completed in 1966

D is a dummy equal to 1 for whites

M is the highest grade of school completed by the mother

F is the highest grade of school completed by the father

FI is the median income of the father's occupation according to the 1960 Census.

The immediate problem is one of potential simultaneity bias. That is, not only may there be an effect of schooling on measured IQ, but there may also be an effect of true IQ on the optimal level of schooling. To the extent that measured and true IQ are correlated, simultaneity bias exists. One way to avoid this problem is to estimate (21) for groups for which there is no relationship between currently observed schooling levels and optimal schooling. For example, if all individuals in the sample attended school through grade 12, but some are still in the process of doing that, we can measure the effect of schooling on measured IQ correctly for these levels of schooling. Extrapolation then gives unbiased estimates of the effect of schooling on measured IQ. This is what was done. Since in 1966, individuals in our sample were 14-24 years old, a substantial portion were still a long way from the school-leaving margin. Thus a sample was selected with the following characteristics: all individuals in the sample were in

1966 currently attending schools in a grade of 12 or less. By 1969, all had graduated from high school. There is therefore no relationship between the observed level of schooling in 1966 and the optimal level since all in the sample had completed at least through 12th grade. The effect of schooling on measured IQ is therefore given correctly by α_1 and α_2 of equation (21) in Table 1. One should note in passing that mothers' education and both fathers' education and occupational status are positively associated with measured IQ. It comes as no surprise that measured IQ depends on race as well.

The value of endowed measured IQ - what measured IQ would have been in the absence of schooling - can now be estimated. Let KWW be defined as (21) with S_{66} set equal to zero so that[20]

$$\hat{KWW} \equiv \alpha_0 + \alpha_3 D + \alpha_4 M + \alpha_5 F + \alpha_6 FI$$

Although it is true that D, M, F and FI affect measured IQ, we do not want to hold these constant when defining the endowed ability level since they are all part of the endowment themselves. (We are not claiming to find "true" IQ, but simply endowed measured IQ where endowed means "in the absence of schooling.")

The most difficult variable to obtain information on is X^e, the endowed wealth of the individual. From Figure 1, the endowed wealth of the individual is seen to be the level that his wealth would have been in the absence of any schooling. The major component of X^e is thus the present value of the individual's earnings stream with zero years of education. The value of this stream is not equal for all individuals; higher IQ persons earn higher wages even with zero years of schooling;

individuals who grow up in the homes of highly educated parents may obtain human capital at home even without any formal schooling; whites earn more than blacks. To ascertain the level of wages at each age with zero years of schooling, the following relationship can be estimated:[21]

$$(22) \quad W_{69} = a_0 + a_1 S_{69} + a_2 \text{ Age} + a_3 \hat{KWW} + a_4 M + a_5 D + a_6 IS$$

where W_{69} is the wage rate in cents-per-hour during 1969 and IS is a dummy equal to 1 for individuals who are currently attending school.

Cross-sectional estimation of (22) done in the usual way suffers from a fundamental flaw in the context of this analysis. By calling a_1 the effect of schooling on wage rates, one assumes the answer to the question posed; i.e., the total correlation between schooling and wage rates would be called production. Longitudinal data allows one to avoid this problem by taking differences so that (22) implies

$$(22') \quad W_{69} - W_{66} = a_1(S_{69} - S_{66}) + a_2(3) + a_6(IS_{69} - IS_{66})$$

One can argue that the a_1 obtained from (22') does not contain the effects of income on schooling as the result of consumption. If all wage changes are anticipated, there is no reason why more rapid wage growth during any particular period in one's life should affect optimal schooling consumption.[22] Even if some of the wage changes are unanticipated, the consumption bias on a_1 should be very small for the following reasons: First, even if this period's wage change reflects a true change in permanent income, only a small part of the increased consumption of schooling should take place during this period since individuals tend to spread consumption evenly over their lifetimes. Thus, $S_{69} - S_{66}$ induced by a given change in permanent income is likely to be small. Second, even if some of the wage change is unanticipated, it is unlikely that a large part of it is unexpected. As the unanticipated component falls, the amount of a_1 that reflects the consumption effect falls. Finally, even if all of the wage change were unanticipated, the elasticity of wealth with respect to the wage change is likely to be small for two reasons. First, to the extent that the change is regarded as a transitory phenomenon, the effect on total wealth will be trivial. Second, even the part that is not regarded as transitory can only affect the calculation of future wages which will lessen the elasticity of wealth with respect to current wage changes as well.

Taking the estimates from (22') to be unbiased, one can then estimate the constrained version of (22) to obtain unbiased estimates of a_0, a_3, a_4 and a_5. The constrained equation is:[23]

$$(22'')\quad W_{69} - \hat{a}_1 S_{69} - \hat{a}_2 \text{ Age} - \hat{a}_6 \text{ IS} = a_0 + a_3 \widehat{KWW} + a_4 M + a_5 D$$

For estimation, one would like to have information on individuals throughout their lifetimes. The Census tapes, for example, provide earnings data on individuals of all ages. Unfortunately, they do not provide data on the crucial endowment variables. Specifically, IQ data are missing from almost all data sets which contain variation in other variables (schooling levels and age, especially). We are therefore forced to rely on the NLS and to extrapolate to older age levels.[24]

Define W_t, the hourly wage in cents at age t, as the predicted value from (22) when S_{69} is set equal to zero, age is set equal to t, and IS is equal to zero. (When $W_t < 0$, W_t is set equal to zero). The present value in dollars of the earnings stream with schooling equal to zero is then

(23) $$X^e = \sum_{t=o}^{65} (.01W_t) \left(\frac{1}{1+.10}\right)^t (2000)\ R$$

where $R = 1.013^{\mu - Age_i}$

μ = mean age of the sample

Age_i = the individual's actual age.

(Based on a 2000 hour work year). R is a correction for vintage effects. Since (26) is estimated from a cross-section, W_t overstates the obtainable wage at age t for individuals who are older than the sample mean age and understates if for those who are younger. The annual vintage effect is estimated elsewhere[25] to be 1.3% per year.[26] Estimates of the wage and IQ equations are contained in Table 1.[27]

It is now necessary to measure X^P, the value of resources used in the production of schooling. Let the cost to the individual of attending the jth grade of schooling be approximated by

(24) $\Delta X^P_j = 1.5C_j - T_j$ where

C_j is the indirect (foregone wage) cost associated with the jth year of schooling and T_j is the financial aid obtained from the parent. Assume $T_j/1.5C_j \equiv F$ to be constant for all j so that (24) can be written as[28]

(25) $\Delta X^P_j = 1.5C_j\ (1-F).$

Then

(26) $$X^P_j = 1.5(1-F) \sum_{j=1}^{J} C_j \left(\frac{1}{1+.10}\right)^{j+5}$$

where X^P_j is the total cost to the individual of acquiring schooling through the jth grade.

TABLE 1

REGRESSION RESULTS

Variable	(21) Dependent Variable $=KWW$	(22') Dependent Variable $=W_{69} - W_{66}$	(22'') Dependent Variable $=W_{69} - \hat{a}_1 S_{69} - \hat{a}_2 Age - \hat{a}_6 IS_{69}$
$\hat{KWW}$			7.1341 (3.056)
$(S_{69} - S_{66})$		2.6741 (2.279)	
$(IS_{69} - IS_{66})$		-44.497 (5.661)	
M	.1279 (.0829)		3.3875 (1.490)
D	3.566 (0.487)		6.7820 (14.65)
F	.2999 (.0678)		
FI	.00013 (.00010)		
S_{66}	-.9318 (1.31)		
S_{66}^2	.1482 (.0702)		
Constant	18.71	86.33 (3.94)	-560.6 (59.8)
R^2	.2069	.0412	.0486
N	1245	1722	1722
Sample Criterion	In school in 1966 with grade level < 12, completed high school by 1969, complete information	Complete Information	Complete Information

Note: Standard errors in parentheses. Those in (22'') are uncorrected for variance introduced by the constraint.

The indirect cost of schooling, C_j, consists of hours spent in school times the price of time. Hours spent in school is approximated by the following linear function, consistent with 25 hours per week at grade 1 to 58 hours per week at grade 18.[29]

$$(27) \quad H(j) = 830 + 70j \ .$$

The price of time can be estimated by examining the wage rates of individuals who are not currently attending school. In order to find the price of time during, say, the 11th grade, one can look at the mean wage of individuals who are not currently enrolled in school and whose highest level of education is grade 10. It is possible, however, that individuals in these two groups are not alike with respect to other variables which influence the wage rate. For example, those who remain in school may have higher values of $\hat{KWW}$ than those out of school. Since $\hat{KWW}$ enters (22), correction for this must be made in calculating the price of time. Define W^*_{ji} as the corrected wage rate for the ith individual during the jth year of school. As above, W^*_{ji} is the predicted wage rate for individual i from (22) where S_{69} is set equal to j-1, age is equal to j+5 and $\hat{KWW}$, M and D assume their actual values. (IS is equal to zero since the wage rate understates the price of time for those in school by minus the coefficient of IS). So

$$(28) \quad W^*_j = [-560.6 + 2.67(j-1) + 28.7(j+5) + 7.134\ \hat{KWW} + 3.388M + 6.782D]R^{\mu-Age}$$

By using this corrected wage (which varies over years of schooling) one can obtain an estimate of the cost to each individual of his actual level of education. The cost of J years of schooling to individual i is given by

$$(29) \quad X^P_j = 1.5(1-F) \sum_{j=1}^{J} H(j)(.01W^*_j) \left(\frac{1}{1+.10}\right)^{j+5} .$$

Finally, one needs a measure of F. Since by assumption, $T_j/1.5C_j$ is constant over all j, we can take T_j to be the current year's financial aid from the parent where j refers to the 1969 schooling level. A problem immediately arises. In order to have a well-defined F, the individual must currently be attending school. However, to estimate (19), the sample is to be restricted to those not attending school in 1969 (or during the previous two years) so that F is not obtainable for this sample.

As above, the solution is to estimate T_j by regressing it on exogenous endowment variables M, D, F, FI, NFAM and $(NFAM)^2$ where NFAM is the number of members in the individual's family other than himself. ($\hat{KWW}$ is excluded since it is a linear combination of other included variables). The sample for estimation purposes is restricted to those individuals for whom non-zero values of T_j were obtained.[30] The results are contained in Table 2. The only important variable in the determination of T_j was the number of other family members. Most significantly, there seems to be no relationship between parental wealth and the level of transfer for schooling.[31] This suggests that parental-wealth related schooling cost differences across individuals may well be small. (This does not state, of course, that all cost differences are small. Schools are more likely to give scholarships, for example, to more able individuals).

Estimation of (20) is now a simple matter. As described above, arbitrary vectors of E, M, K and D can be used to generate the solutions for A, γ, δ, η, λ. Observations from the NLS sample were used as the arbitrary vectors. The results are contained in Table 2 and yield the following estimates: γ = .2061; A = 963,400 α = -.5659; η = -.0478; and λ = -.0019. The selection of different subsets of vectors yielded very similar coefficients.

We are now ready to estimate equation (19). There are two difficulties encountered. First, E appears on both sides of the equation so that the independent variables are not uncorrelated with the error. The solution to this is to use a two-stage estimation procedure. The first stage consists of obtaining predicted values for the right-hand variables by regressing them on a set of instruments. Then E is regressed on the predicted values to obtain unbiased estimates of the coefficients in (19). Since everything in the model depends fundamentally on a few exogenous endowment variables--M, F, FI, NFAM, D--we know that there is some underlying polynomial expression which expresses E as a function of these variables and approximates the analytical solution to (19). We therefore use as instruments the terms of a second degree polynomial formed from the endowment variables. Since these are uncorrelated with the error term, tye produce consistent estimates.

The second difficulty involves the overidentification of the parameters of (19) caused by the non-linear structure. The two-stage estimation process above gives consistent estimates of $1/\Theta$ and β_1 from the intercept and first four coefficients, respectively. However, a method analogous to the Durbin two-stage technique of time series analysis produces estimates which incorporate the information of the non-linear constraint. Using the constant from the initial regression as a consistent estimate of $-\frac{1}{\Theta}$, rewrite (19) as

TABLE 2

REGRESSION RESULTS

Variable	(20) Dependent Variable $=\ell n\ E$	Dependent Variable $-T_j$
Constant	2.839	1064 (311)
$\ell n\ (\frac{X^P}{1-F})$*	0.2061 (0.0017)	
$\ell n\ (K)$	-0.5659 (0.0105)	
$\ell n\ (M)$	-0.0477 (0.0038)	
D	-0.0019 (0.0031)	51.58 (180.6)
M		21.44 (25.3)
F		-14.89 (23.42)
FI		.02942 (1.018)
NFAM		-158.11 (72.4)
$(NFAM)^2$		10.876 (7.308)
R^2	.8402	.0398
N	3366	258
Sample Criterion	Complete information	Complete information, in school and $T_j > 0$

*The actual value of $[1.5 \sum_{j=1}^{J} C_j (\frac{1}{1.10})^{j+5}]$ from (29) is used here since the 1-F terms cancel.

$$(30) \quad E + \hat{\frac{1}{\theta}} = \beta_o(X^p - X^e + \frac{X^p}{\hat{\theta}\gamma E}) + \beta_1(X^p - X^e + \frac{X^p}{\hat{\theta}\gamma E})K + \beta_2(X^p - X^e + \frac{X^p}{\hat{\theta}\gamma E})M + \beta_3(X^p - X^e + \frac{X^p}{\hat{\theta}\gamma E})D.$$

The same 2SLS procedure used to estimate (19) can now be used to obtain consistent estimates of β_1 whose efficiency exceed that of the initial β_1. The results are presented in Table 3.

The estimates of $1/\theta$ and the β_1 are presented below:

$1/\theta$	-22.68
β_0	.000804
β_1	-.0000083
β_2	-.0000237
β_3	-.0000747

Additional Results and Implications

The most important finding is that $\theta < 0$; education is a bad! This does not imply, of course, that education provides no positive utility at any level, but simply states that the representative individual pushes consumption of education to the point where he would no longer accept it at zero cost. I.e., if someone were to offer to cover the cost of schooling (both tuition and the foregone earnings component) on the condition that

TABLE 3

2SLS REGRESSION RESULTS

Independent Variable	(19) Dependent Variable $= E$	(30) Dependent Variable $= E + \frac{\hat{1}}{\theta}$
$(X^P - X^e)$	.00122 (.00023)	
$(X^P - X^e)K$	-.0000087 (.0000038)	
$(X^P - X^e)M$	-.0000133 (.0000027)	
$(X^P - X^e)D$	-.0000339 (.0000078)	
$X^P/\theta\gamma E$	.00401 (.00058)	
$(X^P/\theta\gamma E)K$	-.0000406 (.0000043)	
$(X^P/\theta\gamma E)D$	-.000489 (.000125)	
$(X^P/\theta\gamma E)M$	.000108 (.000012)	
$(X^P - X^e + X^P/\theta\gamma E)$		.000804 (.000017)
$(X^P - X^e + X^P/\theta\gamma E)K$		-.0000083 (.0000004)
$(X^P - X^e + X^P/\theta\gamma E)M$		-.0000237 (.0000010)
$(X^P - X^e + X^P/\theta\gamma E)D$		-.0000747 (.0000077)
Constant	22.68 (2.04)	
SEE	8505	4.999
N	1636	1636
Sample Criterion	Out of school for past 2 years, complete information $E \neq 0$	

future wages would remain unaltered by schooling, the offer would be declined. Optimal schooling falls short of the wealth-maximizing level.

It is also found that $\frac{\partial P_E}{\partial K}$ and $\frac{\partial P_E}{\partial M}$ are positive and substantial and that whites receive a higher return to education than do blacks. Since

$P_E = (\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)^{-1}$,

$$\frac{\partial P_E}{\partial K} \cdot \frac{K}{P_E} = -\beta_1(\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)^{-1} K = .82$$

or $\frac{\partial P_E}{\partial K}$ = \$102 at the point of means. Similarly,

$$\frac{\partial P_E}{\partial M} \cdot \frac{M}{P_E} = -\beta_2(\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)^{-1} M = .80$$

or $\frac{\partial P_E}{\partial M}$ = \$307.

Both changes in measured IQ and changes in parental education have large and similar effects on the returns to education. More can be said on this. From Table 1, one can calculate the wage elasticity (given education) with respect to M and K. At the mans, $\frac{\partial W}{\partial K} \cdot \frac{K}{W} = .67$ while $\frac{\partial W}{\partial M} \cdot \frac{M}{W} = .11$. Although the effect of IQ on wages is similar to that of IQ on education, parental education seems to be much more important in influencing the returns to schooling than in affecting wage rates. This seems reasonable. Ability (especially as KWW measures it) is likely to be a more generally applicable asset than is parental education. The latter might be expected to have specific components and one is not surprised that additional parental education is most useful in teaching children how to benefit from schooling.

The fact that $\Theta < 0$, i.e., that education is considered a bad, implies that the wealth-maximizing level of education, E^*, should exceed the optimal level of education. From (18), one can solve for E^* directly. At the means, a white individual's wealth-maximizing level of education is 16.0 years. The mean level of attained education among whites is 11.9 years. Individuals tend to quit school after high school even though the wealth-maximizing stopping point is college graduation because they regard school attendance as an unpleasant activity. The calculation for blacks reveals that the mean level of P_E for blacks is only 70% of the mean P_E for whites. This is consistent with other results on black/white returns to schooling during the late 60's. The mean black's wealth-maximizing level of education is 15.5 years; he acquires 10.3 years of education so that both the actual and wealth-maximizing levels of education are lower for blacks than whites. Although the costs of attending school are lower for blacks (foregone earnings are lower) this is more than offset by the smaller returns to schooling that accrue to blacks.

One can perform the same analysis with respect to differences in the $\hat{KWW}$ and M variables. Differentiating (18) with respect to K one obtains:

$$(31)\quad \frac{\partial E^*}{\partial K} = \frac{\gamma}{1-\gamma}\left[(\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)^{-1}\left(\frac{A\gamma}{1-F} K^{\delta/\gamma} M^{\eta/\gamma} e^{\lambda D/\gamma}\right)\right]^{\frac{\gamma}{1-\gamma} - 1}$$

$$\cdot \left[\frac{\delta}{\gamma} K^{\frac{\delta-\gamma}{\gamma}} \frac{A\gamma}{1-F} M^{\eta/\gamma} e^{\lambda D/\gamma}(\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)^{-1}\right.$$

$$\left. - \frac{A\gamma}{1-F} K^{\delta/\gamma} M^{\eta/\gamma} e^{\lambda D/\gamma}(\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)^{-2}\beta_1\right]$$

Evaluation at the means yields $\frac{\partial E^*}{\partial K} = -.284$ so $\frac{\partial E^*}{\partial K} \cdot \frac{K}{E^*} = -.50$. Similarly,

$\frac{\partial E^*}{\partial M} = .242$ and $\frac{\partial E^*}{\partial M} \cdot \frac{M}{E^*} = .14$. This result is consistent with the above discussion. Since parental education affects the returns to schooling by a much greater amount than it affects wage rates (and the cost of schooling), it makes sense that individuals with more educated parents should have higher wealth-maximizing levels of schooling. However, since the effect of higher ability on returns to schooling is similar to that of parental education, while ability plays a much more important role in raising wages, the negative $\frac{\partial E^*}{\partial K}$ can be reconciled with the positive $\frac{\partial E^*}{\partial M}$.

This does not i ply that one should observe a negative correlation between schooling and IQ. There are two reasons: First, observed IQ is directly and positively affected by schooling so that it may be the case that the higher observed IQ individuals actually have lower endowed IQ's. But this is not necessary. Even if observed IQ were a perfect measure of endowed IQ, the attained level of education may be higher for higher IQ individuals even though the wealth-maximizing level is not. In terms of Figure 1, this could happen if the situation were the one pictured in Figure 2. Here individual 2 is the high ability worker. Although $E^*_1 > E^*_2$, $E^c_2 > E^c_1$ because the high ability opportunity set is higher and steeper than the low ability opportunity set.

We therefore want to consider the realtionship between the optimal E and K and M. The difficulties involved in solving (19) explicitly for E have already been mentioned. However, the implicit function theorem allows us to describe its behavior without obtaining an explicit solution. Write (19) as

$$(32) \quad Q(E, K, M, D, X^P, X^e) = (\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)(X^P - X^e + \frac{X^P}{\theta\gamma E}) - \frac{1}{\theta} - E = 0$$

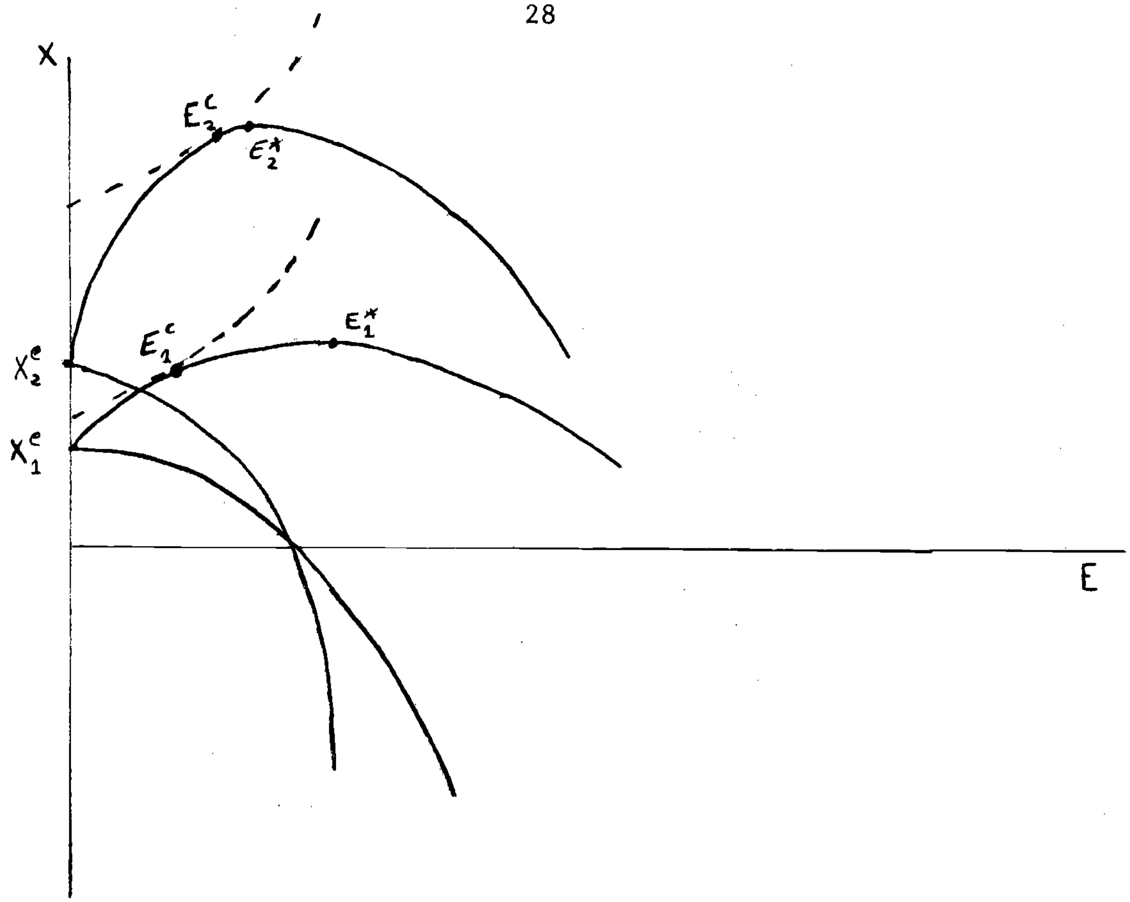

Figure 2

It is then true that

$$-\frac{dE}{dK} = \frac{\frac{\partial Q}{\partial K}}{\frac{\partial Q}{\partial E}} .$$

The terms on the right hand side are easily evaluated.

$$\frac{\partial Q}{\partial K} = (\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)\left(\frac{\partial X^p}{\partial K} - \frac{\partial X^e}{\partial K} + \frac{\partial X^p}{\partial K} \cdot \frac{1}{\Theta\gamma E}\right) + \beta_1 (X^p - X^e) + \frac{\beta_1 X^p}{\Theta\gamma E} .$$

The inverse of (15) implies that $\frac{\partial X^p}{\partial K} = -\frac{\delta X^p}{\gamma K} = 305$ at the point of means.

From (23), we obtain at the mean

$$\frac{\partial X^e}{\partial K} = 20 \sum_{0}^{65} \frac{W_t}{K} \left(\frac{1}{1.10}\right)^t = (20)(10.47)a_3 = 1493.$$

Similarly,

$$\frac{\partial Q}{\partial E} = (\beta_0 + \beta_1 K + \beta_2 M + \beta_3 D)\left(\frac{\partial X^P}{\partial E}\right)\left(\frac{1}{\theta\gamma E} + 1\right) - \frac{(\beta_0+\beta_1 K+\beta_2 M+\beta_3 D) X^P}{\theta\gamma E^2} - 1 .$$

From the inverse of (15), $\frac{\partial X^P}{\partial E} = \frac{X^P}{\gamma E} = 1255$. Plugging these values in yields

$$\frac{dE}{dK} = -.23$$

so that

$$\frac{dE}{dK} \cdot \frac{K}{E} = -.54$$

The optimal level of schooling and endowed ability are negatively related at the means. Yet, one observes a positive relationship between measured ability and attained schooling. This suggests that the observation is to a large extent the result of measurement error which alters ex post measured ability, but not endowed measured ability. Table 4 reveals that the average level of $\widehat{KWW}$ is very stable across schooling levels. This simple correlation of zero between $\widehat{KWW}$ and E lends some support to the reasonableness of the result.[32] Also, it should be remembered that $\frac{dE}{dK}$ holds other things as parental education and wealth constant. Since endowed ability and other background variables are likely to be positively related, a negative partial but zero simple correlation between endowed ability and schooling is indeed reasonable.[33]

In the same way, $\frac{dE}{dM}$ can be calculated. At the point of means, $\frac{dE}{dM} = .20$ and $\frac{dE}{dM} \cdot \frac{M}{E} = .16$. The optimal level of education and parental education vary directly. Holding ability and other background variables constant, more educated parents can be expected to have more educated children. Neither the simple nor partial relationship is surprising.

We observe that blacks on average opt for about 1.6 fewer years of schooling than do whites. It is also the case that the average value of M is 8.0 for blacks and 10.1 for whites and that the average $\hat{KWW}$ is similarly 26.6 and 28.4 respectively. From the above calculations, one can approximate the amount of additional education that blacks would obtain if they had the whites' level of M and K. Thus

$$(33) \quad \hat{E}_B - E_B = \frac{dE}{dM}(2.1) + \frac{dE}{dK}(1.8) = .42 - .41 = .01 \quad .$$

The effect is trivial. The effect of blacks lower $\hat{KWW}$ almost exactly offsets that of their lower value of M. Part of the residual is the result of differences in mean $\hat{F}$ which is .004 for blacks and .015 for whites. Most of the difference, however, results from much lower returns to schooling for blacks than for whites.

The question was orginally posed in terms of wealth and education, while discussion up to this point has been cast in terms of endowment variables and education. The reason is straightforward. The way by which an individual's wealth varies in this model is through interpersonal differences in ability or background which alter the expected lifetime earnings stream. This implies that it is misleading to talk about a pure wealth effect on education since wealth cannot change without also changing the cost of education, i.e., without shifting the education production function. There exists no conceptual experiment analogous to a pure increase in wealth when discussing education. When wealth changes, so must prices.

The original question can now be answered. "Is the correlation between education and wealth the result of consumption or production?" Part of the answer has already been presented. Individuals do not appear to attend school for consumption purposes. School attendance appears to be justified by its ability to affect wealth. The fact that 97% of all individuals in the sampel had a negative E-E* attests to this.

Table 4 reveals an important fact. The wealth-maximizing and observed level of schooling are positively correlated and $|E-E^*|$ falls as E increases. For individuals with $E \geq 18$, the mean level of E exceeds E*. These individuals push education beyond the wealth-maximizing level. This suggests that at least for individuals who achieve high levels of education, (e.g., M.A.'s and Ph.D.'s), schooling is considered a good. Firm conclusions on this cannot be reached, however, without appropriately segregating the sample.[34]

Additional information is obtained by considering the regressions in Table 5. The results of (34) show that although there is a strong positive association between E and E*, only 6-1/2% of the total variation in E is explained by variation in E*. There is a great deal of residual variation left to explain. Part of this variance can be explained by consumption while part is simply "randon error."

Regressions (35) and (36) verify the consistency of the last few pages worth of calculations. From (35), one sees that the partial relationship between K and M and E* are negative and positive, respectively. From (36), we find that the partial relationships between K and M and E are

TABLE 4

Highest Grade of Schooling Completed	$\hat{KWW}$ (1)	M (2)	E-E* (3)	σ(E*) = σ(E-E*) (4)	E* (5)
6	28.9	5.37	-8.0	.75	14.0
7	28.4	6.91	-8.1	1.34	15.0
8	27.0	7.32	-8.3	1.6	16.3
9	27.5	7.99	-7.3	1.6	16.3
10	28.5	8.87	-6.0	2.0	16.0
11	27.6	9.41	-5.6	1.3	16.6
12	27.9	9.76	-4.7	1.7	16.7
13	28.1	11.1	-4.1	1.2	17.1
14	29.1	11.6	-2.7	2.5	16.7
15	27.2	11.9	-2.7	1.3	17.7
16	28.4	12.0	-1.5	1.4	17.5
17	27.7	11.6	-0.2	0.9	17.7
≥ 18	28.4	12.5	+0.3	1.3	17.7
Total 11.7	27.9	9.56	-4.9	σ(E*) = 1.7	16.6
				σ(E-E*) = 2.5	
Whites 12.0	28.5	10.2	-4.8	σ(E*) = 1.7	16.8
				σ(E-E*) = 2.6	
Non-Whites 10.8	27.0	8.3	-5.2	σ(E*) = 1.7	16.0
				σ(E-E*) = 2.2	

*There were no observations in the subsample for which $E < 6$ and very few for low levels of E so these entries should be regarded with skepticism.

TABLE 5

REGRESSION RESULTS

Variable	(34) Dependent Variable = E	(35) Dependent Variable = E*	(36) Dependent Variable = E
E*	.334 (.031)		
$\hat{KWW}$		-.304 (.006)	-.054 (.012)
M		.237 (.008)	.302 (.018)
D		.860 (.006)	.737 (.125)
Constant	6.17 (0.54)	22.2 (0.2)	9.75
R^2	.065	.692	.200
SEE	2.20	.969	2.04
N	1636	1636	1636

also negative and positive, respectively. Blacks have lower levels of E and E* than do whites. All signs correspond to those predicted in the analysis of previous pages. These findings again support the claim that variance in observed E moves in the direction predicted by the utility maximizing framework set out in this paper. Furthermore, since K, M and D by themselves can, even in this simple equation, explain 20% of the variation in E, differences in wealth-maximizing levels of education across individuals (as formulated in this model) tell only part of the story. This does not say that the investment motive is unimportant. The findings of this paper suggest that it is supreme. Virtually all education is wealth-increasing. We simply find that variations in E* across individuals do not go far to explain variations in E.

The robustness of the obtained results was tested by altering various assumptions made throughout the model. The first variation allowed $\widehat{KWW}$ to depend on age as well as schooling. The results were virtually identical to those previously obtained: $1/\Theta = -22.51$, $\beta_0 = .000778$, $\beta_1 = -.0000080$, $\beta_2 = -.0000231$, $\beta_3 = -.0000743$, $\gamma = .1931$. This resulted in $P_E = \$3816$ and E* of 15.7 at the mean.

Second, the discount rate was changed to .05 from .10. γ became .2001 and $1/\Theta = -25.12$. All other coefficients changed almost proportionately to 1/3 their original values. This resulted in $P_E = \$10{,}526$ (higher because of the lower discount rate), but E* increased only to 17.3 years from 16.0. All other qualitative conclusions were the same. Thus, even with such drastic alteration of the discount rate, the findings remain intact.

Third, the hours-in-school equation was assumed to be 2/3 of its original magnitude, thereby lowering the cost of a year of schooling.

Again, all coefficients in (30) change by a scalar, P_E fell to \$2136, but $1/\Theta$ is still negative at -28.62 and E* remains at 16.1. Again, all previous statements hold.

Finally, the assumption on direct costs of schooling was changed. Instead of direct costs being one-half of foregone earnings, they were assumed to be zero through grade 12. This did not affect the results at all. All coefficients were very similar, P_E = \$2891 and E* was 16.0 years again for the mean individual.

One final check on the reasonableness of the results can be made by examining the predicted wealth from (2). Evaluating all terms at the means, one obtains that P_E = \$3571 and X^c = \$57,305. At the assumed 10% rate of interest this yields a permanent annual income (consumption) figure of \$5730. This seems in line with casual notions of average permanent income. In addition since $\bar{X}^e$ = \$18,557, schooling triples the wealth of the representative individual.

As a final point, it is now useful to consider the welfare effects of government subsidization of education. One of the primary motivations of government subsidization seems to be the redistribution of opportunity if not wealth. If it were the case that education did not enter the utility function at all, then individuals would necessarily go to their wealth-maximizing levels. Redistributive arguments for giving education to the poor instead of assets paying the market rate of interest would have to rest on differences in the cost of obtaining funds across individuals (ignoring externalities).[35] Since education enters the utility function negatively, however, the individual stops short of the wealth-maximizing level of education. This means that education yields a higher pecuniary return at the margin than do other assets (the difference being the value of the marginal disutility). By subsidizing education, the government could induce individuals to move to the wealth-maximizing level of education. If the only objective were to raise pecuniary wealth of particular individuals, this method would yield a higher return than the same amount of transfer expenditure in non-human capital. It is still true, however, that the utility level of the recipients would be lower when the transfer is in education (a non-saleable asset) rather than in bonds.

The results of the above analysis allow us to calculate the amount of subsidy necessary to move the mean individual to the wealth-maximizing level of education. Since $\mathcal{E}(E^*-E) = 4.9$ years, the government should set as an approximation $\frac{\partial E}{\partial X^P} dX^P = 4.9$ (since $\partial E/\partial X^P$ varies with X^P, this is only an approximation). dX^P is, like parental transfer, an education-specific subsidy. From (32),

$$\frac{\partial E}{\partial X^P} = -\frac{\frac{1}{P_E} + \frac{1}{\Theta P_E \gamma E}}{\frac{\partial Q}{\partial E}} = .0007$$

so that $dX^P = \$6872$ in period zero dollars. Since this occurs at the average school quitting age of about 18, in age 18 current dollars the expenditure is $[e^{(.1)(18)}][6872] = \$41{,}575$, a substantial subsidy. The exact expenditures varies across individuals. Table 4 reveals that $|E-E^*|$ is greater for non-whites than it is whites. The wealth-maximizing scheme would therefore require that a larger subsidy go to non-white individuals. Similary, since $\frac{\partial E^*}{\partial K} = -.28$ and $\frac{\partial E}{\partial K} = -.23$, the difference between E and E^* decreases with increased ability. The required subsidy would therefore be larger for low ability workers. It should be remembered, of course, that all would be better off with a non-wealth-maximizing transfer of money of the corresponding amount than with an educational subsidy. Only if wealth-maximization is an end by itself is the educational subsidy warranted (again, neglecting externalities).[36]

Summary and Conclusions

The question posed in this paper is an old one: Why attend school? Does the well-documented relationship between education and income result because schooling allows individuals to earn higher wages or because high income individuals purchase more schooling as they purchase more of all normal goods? The first part of the paper sets out a theoretical framework which treats education as a joint product, producing (potential) wage gains and utility, independent of the wealth increment, simultaneously. If education enters the utility function either positively or negatively, it is necessarily the case that the optimal level of education will differ from the wealth-maximizing level. It is thus misleading to ask how much of educational acquisition is due to consumption. The answer is, "all of it." The appropriate question is how large is the difference between the consumption optimum and wealth-maximizing level? This will be zero if education does not enter the utility function and will be equal to the attained level if education does not affect wealth.

The choice of general functional forms for the uitlity and educational production function allows one to ascertain the relevant magnitudes. The model, which uses information on parental education and occupation, number of siblings, race, and endowed (in the absence of schooling) measured

IQ, produces estimates of the optimal and wealth-maximizing levels of education, and the market value of education. The way in which these values are affected by differences in the background variables is also determined.

One of the most difficult problems encountered was to obtain a measure of endowed ability. A method is devised which nets out the effect of schooling on measured IQ independent of the simultaneity bias which results from potential causality between true IQ and the optimal level of schooling. The solution rests on the use of longitudinal data (the National Longitudinal Survey on young men is used) which follows individuals over a portion of their lifetimes.

The fourth section contains the results of the analysis. The findings can be summarized:

1. Most important is the finding that education is a bad, i.e., it enters negatively into the utility function. Individuals therefore stop short of the wealth-maximizing level of education by an average of 4.9 years. While high school graduation is the average quitting point, acquisition of an undergraduate degree is still a whorthwhile investment for the mean individual.

2. The wealth-maximizing level of education varies inversley with endowed IQ and directly with parental education. Although increases in both imply increased returns to a year of schooling, they also imply higher costs as well as the result of higher foregone earnings. Increases in endowed IQ effect wages to a larger extent than returns to education while the opposite is true for parental education. Similarly, the wealth-maximizing level of education is lower for blacks than for whites.

3. It is also the case that the simple correlation between endowed (as opposed to ex post measured) IQ and the level of attained schooling is zero. This suggests that there is no tendency for higher education to be dominated by the more intelligent. The partial relationship between IQ and optimal schooling is, in fact, negative. This results from two factors: First, the cost of schooling is higher to the more able due to higher foregone earnings. Second, the more able are richer and as such acquire less of schooling, a normal bad. Parental education is both in a simple and partial sense, positively associated with attained schooling. Similarly, the model predicts (as well as observes) that blacks will acquire less education than whites. This result attains even when corrections are made for differences in parental education and endowed measured ability. The difference is caused by much lower returns to schooling for blacks than for whites.

4. Although the wealth-maximizing level of education and actual level of education are positively related, variation in the former explains only 6-1/2% of variation in the latter. Actual levels of education move with the endowment variables in the way predicted by the analysis. This suggests that part of the residual is explained by differences in the consumptive optimum across individuals. It is true, however, that 97% of the sample stopped short of the wealth-maximizing level.

5. An educational subsidy of $6872 in present value at birth would cause the mean individual to acquire the wealth-maximizing level of education.

The model presented in this paper is empirically complicated and requires many assumptions before any conclusions are obtained. To the extent that extrapolation results in rough measures of important variables, the final conclusions must be regarded as less than definitive. On the other hand, the tests for robustness revealed that the results were quite stable even when important assumptions were changed substantially. Many checks on the reasonableness of the findings have also been provided and the model performed quite well by these criteria. The assessment of the finding's credibility, of course, rests with the reader.

FOOTNOTES

1. See Arrow (1972), Stiglitz (1973) and Wolpin (1974) for a more complete discussion of this hypothesis.

2. Levhari and Weiss (1974) offer an alternative explanation. They show that under conditions of uncertainty with decreasing absolute risk aversion, an increase in initial wealth will encourage investment in more human capital since the wealthier will buy more risky assets (human capital being one).

3. Schultz (1962) describes the benefits to education as divisible into three components: an investment component which results in an increase in an individual's measured wealth; a present consumption component such as the utility currently derived, say, from attending class; and a future consumption component which results from the fact that education improves one's ability to consume other goods later in life.

4. This amounts to assuming that additional years of schooling produce slightly more units of human capital where a unit of human capital is defined such that its rental price in current dollars is constant for each period of rental. The increase in the number of units of human capital is then assumed to exactly offset the decline in the value of human capital acquired later in life which results from a shorter payoff period. Note that this assumption does not imply a constant "rate of return" since the education production function, $f(X^P)$, allows the costs of producing education to rise with E.

5. This assumes that "education" is the appropriately defined good. Specifically, this requires that there are no possibilities of different types of education, some of which are more productive in the market, some of which are more productive in the non-market.

6. One is not indifferent between C and B, however, since measured wealth (i.e., ability to purchase goods) is the same at each point, but the amount of education consumed at C exceeds that of point B.

7. Ishikawa (1973) obtains a similar conclusion.

8. Some experimentation was done with other utility functions (Cobb-Douglas, among others), but all lacked the necessary properties and the results obtained were therefore non-sensical.

9. As Rosen (1973) points out, the assumptions of the human capital framework as applied empirically generally preclude the observation of different levels of schooling across individuals.

10. D is entered exponentially since D^{λ} would imply $E = 0$ for non-whites.

11. One might also argue that tastes differ across individuals, depending upon the values of the endowment variables. For simplicity, neutrality of tastes is assumed with respect to endowment variables. That is, although it is recognized that higher IQ individuals may derive greater absolute utility from both E and X^c, it is assumed that the ratio of marginal utilities (ΘX^c) is independent of K, M and D.

12. The child generally receives some parental transfers even in the absence of attending school. It is the amount by which this transfer increases when the child attends school which should be thought of as school cost financed by the parent.

13. Becker (1967) argues that schooling differences across individuals can be caused either by differences in returns to schooling or differences in costs. Treating P_E as variable addresses the former while incorporating parental financing attempts to deal with the latter. Differences in borrowing rates are ignored only to the extent that they do not operate through parental transfers.

14. The difference between observed E and E* does not accurately reflect the amount of education due exclusively to consumption. The true consumption component is $\hat{E} - E^*$ where $\hat{E}$ is the solution to (19) in terms of E for each individual. Not all of unexplained variation should be called consumption since some of it is the result of error in the calculation of E*.

15. Given that these individuals are young, there is a reasonable probability that they will return to school even if not currently attending. This may cause E to be mismeasured. Dealing with young individuals has the advantage, however, that X^p, the costs of schooling, can be more accurately measured than they can for older individuals whose schooling occurred further into the past.

16. Observations are selected for which $E > 0$. $E \geq 18$ is coded as $E = 18$.

17. "Better" is defined in the context of wage functions. It is not clear that it will do as well as a measure of the ability to produce education. In the absence of a superior measure, however, it must suffice.

18. If all individuals attended the same number of years of schooling, the problem would be less serious since the more able would learn these testable skills more readily.

19. Griliches and Mason (1972) are faced with much the same problem and this technique is not totally unlike theirs.

20. An alternative measure would be $\hat{KWW} \equiv KWW - \alpha_1 S_{66} - \alpha_2 (S_{66})^2$. This measure uses the information in the error term and would be superior if the error primarily reflected unestimated differences in measured

ability rather than unestimated differences in the effect of schooling on measured ability. This measure contains more information on the individual, but is also less likely to be free of schooling effects. The inclusion of age in the KWW equation changed none of the results. This is discussed in more detail in the final sections.

21. Neither $(S_{69})^2$, F, nor FI entered significantly into the regression. See Lazear (1974) for a discussion of the IS variable.

22. Nor is there any reason to expect that wage growth and average wage levels will be correlated if schooling does not affect wages. E.g., the more able may have flatter earnings profiles in the absence of schooling than the less able, even though the level of the former's wages are higher.

23. This specification is tantamount to the assumption of neutrality. That is, the effect of age, for example, on wage rates is treated here as being independent of schooling. If age is a proxy for on-the-job training, this says that schooling has no effect on the cost of investment in OJT because it increases efficiency in producing this form of human capital by the same amount as it increases the costs of producing it (the foregone earnings, primarily). Experimentation with functional forms revealed the absence of interaction effects so that neutrality in this context may not be far from the truth.

24. This problem is not as serious as it appears since the inaccuracies at later ages are to a large extent mitigated by the present value calculation which renders them small relative to the earlier component.

25. See Lazear (1975), p. 11.

26. This calculation appears to have ignored the effect of parental wealth on the child's endowed wealth. This is not the case. FI was allowed to affect wage rates directly in (22), but did not enter significantly. FI enters indirectly by altering $\widehat{KWW}$. The direct transfer of funds from wealthy parents to their children is taken into account by the introduction of F in (15).

27. The fact that only a small portion of the variation in wages is explained in (22") is not too disturbing. As long as the individual has no better information about the rest of wage variation than we do, it can be ignored since he will not take it into account in estimating his own endowment.

28. The assumption that direct costs are one-half foregone earnings was replaced by the assumption that they are zero through year 12 and half of foregone earnings beyond that. The results were not altered. This is discussed in more detail in the final section.

In addition, changing the discount rate to .05 from .10 did not qualitatively alter the results.

29. Equation (27) was replaced by $H(j) = .66(830 + 70_j)$ and the complete analysis was carried out. Again, the results changed very little (as is reported below). The robustness of the findings in light of such apparently substantial changes in assumptions is encouraging.

30. The data do not distinguish between a true value of zero and a zero value for insufficient information. This prevented the application of tobit analysis.

31. A linear probability regression with the same independent variables and the dependent variable being a dummy equal to 1 when $T_j > 0$ yielded similar results. The sample was restricted to those currently attending school so that the $T = 0$ was less likely to reflect insufficient information. No variable in this regression entered significantly and all coefficients (although biased) were very small.

32. Nor is this the result of a mechanical relationship. It is not true that because KWW is constructed by holding schooling constant that there must be no relationship between the two. There are two reasons: The coefficient on the schooling coefficient in the KWW regression is not obtained cross-sectionally. Given the construction of the sample for estimation of (21), the correlation between optimal (final) schooling and KWW is removed. Thus, taking this effect out says nothing about cross-sectional findings on $\widehat{KWW}$. Second, D, F, and FI enter significantly and are not held constant by Table 4. Even if the partial correlation between $\widehat{KWW}$ and E were zero, there would be no reason for the simple correlation to be zero.

33. This finding which suggests that the truly bright individuals are out in the real world earning money rather than ivory-towering their time away, is sure to please a large part of the population. Before conceding the issue, however, two points should be made. If it is true that the more able take the returns to schooling in non-pecuniary ways, then the measured returns understate true returns by more for the more able. This might lead to the prediction that optimal levels of schooling and ability are negatively correlated when the reverse is true. One wonders, of course, why it is that the more able should have a relative preference for non-pecuniary returns. Higher tax rates on higher income is the usual explanation. No additional justifications are offered here.

34. An attempt to stratify the sample into low and high education groups was only partially successful. Observations were split into these obtaining more than twelve years of schooling and those obtaining less than twelve years. For the low schooling groups, the following results were obtained: $\gamma = .1561$, $1/\theta = -23.98$, $\beta_0 = .000705$, $\beta_1 = -.0000089$, $\beta_2 = -.0000211$, $\beta_3 = -.0000747$. These coefficients are similar to those in the text, but yield $P_E = 3508$ (as compared to \$3571 for the group as a whole) and $E^* = 13.4$ (rather than 16.0). This is consistent

with the group being low-schooled individuals. For the high schooling group, only 300 observations were present. This resulted in $\gamma = .3961$, $\beta_1 = -.00000276$, $\beta_2 = -.0000061$, $\beta_3 = -.0000036$. Although the coefficients retain the same signs and qualitative implications, P_E becomes 7092 and E* a ridiculous 51 years! This is directly attributable to the high γ and presumably to the insufficient information in the few observations.

35. If the government could finance education at 10% while some individuals faced 15% borrowing rates then what was wealth-maximizing at the private level would fall short of what was wealth-maximizing at the social level.

36. Wealth-maximization could be a goal in itself if a country were trying to produce maximum output en route to a developed state. In such a case, leaving school prior to the wealth-maximizing level is analogous to failing to build a machine whose return exeeds cost.

REFERENCES

Arrow, K. "Models of Job Discrimination," in A. Pascal, ed., Racial Discrimination in Economic Life. Lexington: Lexington Books, 1972.

Becker, G. Human Capital and the Personal Distribution of Income. W. S. Wotinsky Lecture #1, University of Michigan, 1967.

Benham, L. "Benefits of Women's Education within Marriage," Journal of Political Economy, 82, 2 (March/April 1974 Supp.): 57-71.

Bowles, S. "Schooling and Inequality from Generation to Generation," Journal of Political Economy (May 1972 Supp.): 219-51.

Griliches, Z. "Research Expenditures, Education, and the Aggregate Agricultural Production Function," American Economic Review LIV (December, 1964).

__________. "Notes on the Role of Education in Production Functions and Growth Accounting." Unpublished manuscript, Harvard University, 1968.

__________. "Wages and Earnings of Very Young Men: A Preliminary Report." Unpublished manuscript, Harvard University, August 1974.

Griliches, Z. and W. Mason. "Education, Income and Ability," Journal of Political Economy (May 1972 Supp.): 74-103.

Ishikawa, T. "The Simple Jevonian Model of Educational Investment Revisited." Discussion paper 289, Harvard Institute of Economic Research, March 1973.

Lazear, E. "Schooling as a Wage Depressant." Working Paper #92, National Bureau of Economic Research, June, 1975.

__________. "Age, Experience, and Wage Growth." Report #7502, Center for Mathematical Studies in Business and Economics, University of Chicago, January 1975. (Forthcoming, American Economic Review).

Levhari, D. and Y. Weiss. "The Effect of Risk and Investment in Human Capital," American Economic Review LXIV, 6 (December 1974): 950-63.

Liebowitz, A. "Home Investments in Children," Journal of Political Economy 82, 2 (March 1974 Supp.): S111-S131.

Michael, R. "Education in Non-Market Production," Journal of Political Economy 81, 2 (1973): 306-27.

Michael, R. "Education and the Derived Demand for Children," Journal of Political Economy 81 (March 1973a Supp.): S128-S164.

Parsons, D. "The Cost of School Time, Foregone Earnings, and Human Capital Formation," Journal of Political Economy 82, 2 (March/April 1974): 251-66.

Rosen, S. "Income Generating Functions and Capital Accumulation," Harvard Institute of Economic Research Discussion Paper #306, 1973.

Rothenberg, T. and C. Leanders. "Efficient Estimation of Simultaneous Equation Systems," Econometrica 32 (January 1964): 57-76.

Schultz, T. The Economic Value of Education. New York: Columbia University Press, 1963.

Spence, M. "Market Signalling: The Informational Structure of Job Markets and Related Phenomena." Discussion paper of Kennedy School of Government, Harvard University, February 1972.

Stigletz, J. "The Theory of Screening, Education and the Distribution." Cowles Foundation Discussion Paper No. 354, Yale University, March 1973.

Welch, F. "Education in Production," Journal of Political Economy 78, 1 (January/February 1970): 35-59.

__________. "Education, Information and Efficiency." Unpublished Working Paper #1, National Bureau of Economic Research, 1973.

Williams, A. "Fertility and Child Mortality: Further Results." Unpublished manuscript, University of Chicago, September 1974.

Wolpin, K. "Education and Screening." Unpublished manuscript, University of Chicago, 1974.