

## FAMILY BACKGROUND AND OPTIMAL SCHOOLING DECISIONS

Edward Lazear\*

THE assumption that individuals "optimize" is one of the most important engines of economic analysis. Recent years have seen a number of attempts to examine this assumption in more detail. Leibenstein (1966) introduced, and a recent note by Stigler (1976) examines, the notion of "X-efficiency" or the ability to operate on the production function. Radner and Rothschild (1975) discuss the equilibrium of markets characterized by firms that operate according to rules of thumb. The industrial organization literature is full of analyses that assume behavior other than profit maximization. (See Simon (1957), Williamson (1966), and Marris and Wood (1971), for example.)

In this paper, another aspect of optimizing behavior is considered. Specifically, it asks whether variations in levels of attained schooling across groups can be explained by a model that assumes that capital markets are perfect and that individuals maximize wealth. It has long been recognized that attained levels of schooling are positively correlated with parental income.<sup>1</sup> This has led some to suggest that the ability to internally finance the acquisition of education tends to perpetuate schooling and income inequality across generations. Capital cost differences, it is argued, may be a crucial variable in the explanation of education and income variation.<sup>2</sup> The model set forth in this paper allows one to test for capital cost differences across income groups. Other things constant, evidence of very small differences is found. We can reject with 95% confidence that implicit borrowing costs to the

rich are as little as 0.25 of a percentage point lower than those for the poor. We can also reject with the same level of confidence that those differences are zero. The conclusion is, therefore, that with 95% confidence the poor face implicit borrowing costs that are greater than those for the rich, but greater by an amount less than one-quarter of a percentage point.

Implications for policy vary considerably with the source of intergroup differences. If the explanation of the poor's relatively low levels of attained schooling rests upon capital market imperfections, the wealth of these individuals (and GNP) could be increased by the provision of "subsidized" loans.<sup>3</sup> On the other hand, if the source of differential schooling levels is differential returns to schooling (because, say, of wage discrimination), subsidized schooling would not be the appropriate solution. Wealth of the disadvantaged group would be most efficiently increased by policies designed to reduce labor market discrimination.

The logic of the analysis proceeds as follows: First, a model is constructed that allows estimation of costs and returns to education for each individual, based on the assumption that all individuals face the same borrowing rates. Given costs and returns, one can obtain an optimal wealth-maximizing level of education for each individual. Differences between actually acquired and wealth-maximizing levels of education can then be calculated, and one can determine whether or not the residuals are systematically related to background variables. If, for example, low-income individuals have a consistently larger estimated wealth-maximizing level of education than actual level, one could conclude either that returns to schooling differed between groups or that capital market differences exist. The model allows these two explanations to be distinguished. Since differential re-

Received for publication November 18, 1977. Revision accepted for publication February 26, 1979.

\* University of Chicago and National Bureau of Economic Research.

The author wishes to thank Gary Becker, Charlie Brown, Gary Chamberlain, Reuben Gronau, James Heckman, Tom MaCurdy, Robert Willis, and Kenneth Wolpin for useful suggestions. Financial support from the National Institute of Mental Health, the Sloan Foundation, and the Rockefeller Foundations is gratefully acknowledged. Able research assistance was provided by Craig Hakkio.

<sup>1</sup> See Bowles (1972) or Jencks et al. (1972), for example.

<sup>2</sup> Becker (1967) provides a theoretical analysis of differences in capital costs and their implications across groups.

<sup>3</sup> This assumes, of course, that the government or agency providing the subsidized loans possesses information on the nature of the imperfection. A variant on this is suggested by Freeman (1975). He argues that it is the state that discriminates in the provision of schooling.

turns are caused by wage differences across groups, the wealth-maximizing level can take these labor market variations into account. Any residual variation will be due to factors other than differential wage rates, presumably capital cost differences.<sup>4</sup>

### I. A Wealth-Maximizing Model

In order to determine whether capital market differences across individuals result in different attained levels of schooling, it is necessary to build a model that takes other causes of schooling variation into account. That is, differences in schooling attainment can result from differences in cost or from differences in returns. In this model, we will take all differences in costs and returns into account, except for those that may result from capital cost variation. Initially, the model will assume that capital costs are the same across all individuals. If, in reality, poor individuals face higher borrowing costs than rich individuals, the model will overpredict education for the poor and underpredict for the rich. The systematic nature of the residuals will result from capital cost rather than wage discrimination differences because the model takes the latter into account directly. The effect of the former, which is unobservable, can be inferred from examination of the residuals.

The problem for the individual is to choose an optimal amount of education,  $E$ , given his endowment vector of attributes,  $S$ . These exogenous attributes are expected to affect the individual's cost of producing education,  $C(E, S)$ , as well as his returns to education. Let  $P_E(S)$  be the "selling price" of a year of schooling. It is interpreted as the present value of the amount by which earnings are increased over the lifetime as the result of acquisition of an additional year of schooling.  $P_E(S)$  varies across individuals with their attributes, but  $P_E$  is assumed to be constant for any given individual across all years of schooling.<sup>5</sup> If  $C(E, S)$  and  $P_E(S)$  are constructed

so that they are in present value units at the individual's time of birth, wealth is maximized by selecting  $E$  to maximize

$$W = EP_E(S) - C(E, S) + X^e(S) \quad (1)$$

where  $X^e(S)$  is endowed wealth in the absence of schooling. If  $C$  is strictly convex in  $E$ , monotonically increasing, and

$$\frac{\partial C(0, S)}{\partial E} < P_E(S),$$

then a maximum is obtained with  $E > 0$  when

$$P_E(S) = \frac{\partial C(E, S)}{\partial E} \equiv C_E(E, S). \quad (2)$$

Equation (2) states the obvious condition that marginal cost is equal to marginal return.

If one knows the function  $C(E, S)$ , an individual's endowment, and his attained level of schooling  $E$ ,  $P_E(S)$  can be inferred from (2), assuming optimization. This is illustrated in figure 1. Consider individual A with attributes  $S_A$ . The attributes define a cost function  $C(E, S_A)$ . One observes that the individual acquires  $E_A$  years of schooling. Since optimality insures that  $C_E(E_A, S_A) = P_E(S_A)$ ,  $P_E(S_A)$  is the slope of the line tangent to  $C(E, S_A)$  at point  $Q$ . Consider individual B with attributes  $S_B$ .  $S_B$  is identical to  $S_A$  except that the  $i^{\text{th}}$  component of  $S_B$  is one unit larger than the  $i^{\text{th}}$  component of  $S_A$ . B obtains  $E_B$  of education so  $P_E(S_B)$  equals the slope of the line tangent to  $C(E, S_B)$  at  $R$ . The difference  $P_E(S_B) - P_E(S_A)$  is now identified and is the market shadow value of attribute  $i$  in returns to education. The individual takes this value as given. Once his attributes are known, his  $P_E(S)$  is determined. He simply moves to the point on his cost curve where his market determined  $P_E(S)$  equals  $C_E(E, S)$ .<sup>6</sup>

Suppose it were the case that the  $C(E, S)$  function were known with certainty (say, from a

of human capital is defined such that its rental rate is constant. The increase in the number of units of human capital exactly affects the decline in value resulting from acquisition later in life. This assumption does not imply a constant rate of return, however, since costs may vary with  $E$ .

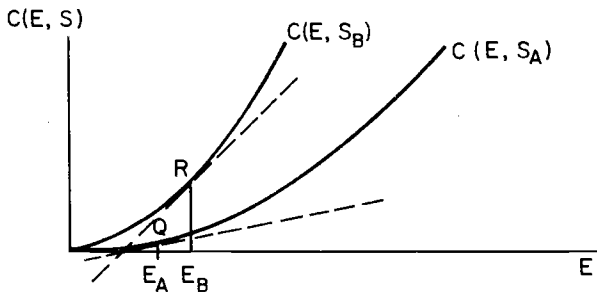
The assumption that  $P_E(S)$  is not a function of  $E$  is essential for identification. This potentially verifiable assumption cannot be independently tested in this model and remains one of its shortcomings. An entire life history of earnings and labor force participation appears necessary to answer that question.

<sup>6</sup> In the language of Rosen (1974), the individual takes the hedonic price surface as given. We are estimating the slope of that surface.

<sup>4</sup> The wealth-maximizing framework allows one to distinguish between labor and capital market differences as explanations for differential schooling levels. A simple regression of schooling on independent variables where nonwhite mean levels of the independent variables are multiplied by white regression coefficients will not distinguish the two. (See Gilman (1965) for an analysis along those lines.)

<sup>5</sup> This amounts to assuming that additional years of schooling produce slightly more units of human capital where a unit

FIGURE 1.—DETERMINATION OF OPTIMAL SCHOOLING LEVEL



previous study). The knowledge of  $C(E, S)$ , along with an individual's chosen  $E$ , allows identification of  $P_E(S)$ . That is, if the cost function is already known, one can infer marginal returns by examining the level of marginal cost which necessarily equals marginal revenue at the chosen equilibrium level of  $E$ . Thus, if the marginal cost function is already known or easier to estimate than the marginal return function, one can determine the latter from the former. Once each individual's  $P_E$  is known, we can estimate a "hedonic"  $P_E$  function by regressing  $P_E(S)$  on the characteristics,  $S$ :<sup>7</sup>

$$P_E = a_0 + a_1 S_1 + \dots + a_k S_k. \quad (3)$$

Here  $a_i$  is the shadow value of attribute  $S_i$ .

Furthermore, once  $P_E(S)$  and  $C(E, S)$  are known, one can work backwards through the model to obtain the wealth-maximizing level of education,  $E^*$ . Since  $C(E, S)$  is strictly convex and increasing in  $E$ ,  $C_E(E, S)$  is strictly increasing in  $E$  and is therefore invertible with respect to  $E$ . Thus,

$$E = C_E^{-1}(C_E, S). \quad (4)$$

At the wealth-maximizing level of education, condition (2) holds so that using (2) and (4), one obtains

$$E^* = C_E^{-1}(P_E(S), S) \quad (5)$$

or

$$E^* = g(S). \quad (6)$$

Thus, given a vector of attributes, (6) identifies their possessor's wealth-maximizing level of education.

It has been suggested that if costs were more

<sup>7</sup> There is no necessity that this be a linear function, and in the empirical analysis, various functional forms were tried. None differed significantly from the linear form.

easily observed than returns, the latter could be inferred from the former. Here, it is useful to specify the cost function. This function is not an estimated behavioral relationship but something that can be determined a priori once tuition and foregone earnings are known.

The primary cost of schooling consists of foregone earnings. Until one enters college, it can be argued that virtually all schooling costs consist of foregone earnings. Assume, therefore, that costs of grades one through twelve are simply the associated foregone earnings and that direct costs for grades beyond twelve are one-half the foregone earnings component. Then, the present value of the cost of  $E$  years of schooling is given by<sup>8</sup>

$$C(E, S) = \begin{cases} \sum_{j=1}^E F_j \left( \frac{1}{1+r} \right)^{j+5} & \text{for } E \leq 12 \\ \sum_{j=1}^{12} F_j \left( \frac{1}{1+r} \right)^{j+5} + (1.5) \sum_{j=13}^E F_j \left( \frac{1}{1+r} \right)^{j+5} & \text{for } E > 12 \end{cases} \quad (7)$$

where  $F_j$  are the foregone earnings associated with year  $j$ . Note that the discount factor is assumed to be the same across all individuals. The only cost differences that are accounted for are differences in foregone earnings across individuals. If borrowing costs do differ across individuals, the misstatement of the cost curve will lead to an overprediction of schooling levels for high finance cost individuals and an understatement for low cost ones.

The foregone earnings consist of hours spent in school times the price of time. Let us approximate hours spent in school by the following linear relationship<sup>9</sup> consistent with 25 hours per week at grade 1 to 58 hours per week at grade 18:

$$H(j) = 830 + 70j. \quad (8)$$

<sup>8</sup> This assumption was changed such that  $C(J, S) = 1.5$

$\sum_{j=1}^J F_j \left( \frac{1}{1+r} \right)^{j+5}$  for all  $J$ . The results were not significantly altered.

<sup>9</sup> This assumption was changed to  $H(j) = \frac{2}{3}(830 + 70j)$ . The results were not qualitatively altered.

The price of time can be estimated. On the margin, it is the opportunity wage that the individual could have earned had he opted to work during a given year. A wage function that includes years of schooling completed and the endowment variables is specified below:<sup>10</sup>

$$\ln W_{69} = b_0 + b_1 E_{69} + b_2 IQ + b_3 M + b_4 N + b_5 D + b_6 R + b_7 (Age) \quad (9)$$

where

$W_{69}$  is the hourly wage rate in 1969 in dollars per hour and

$N$  is the number of members in the individual's family in 1966,

$D$  is a dummy set equal to 1 for whites,

$R$  is a dummy set equal to 1 if the residence at age 14 was not in a Standard Metropolitan Statistical Area (SMSA),

$IQ$  is the score reported by the high school on the most recent IQ test, and

$M$  is the highest level of education completed by the mother.

Using the estimates from equation (9) in table 1, the current expected price of time during year  $j$  is then<sup>11</sup>

$$W^*_j = \exp[-.1655 + .04196(j - 1) + .01540(j + 5) + .00187(IQ) + .00775(M) - .01001(N) + .11286(D) - .11516(R)]. \quad (10)$$

Thus, upon substitution, equation (10) becomes

<sup>10</sup> Other functions (and functional forms) were tried. Specifically, father's income, schooling squared, schooling-IQ interactions, and a number of other endowment variables were added. They did not enter significantly. Nor did coefficients differ significantly between whites and nonwhites other than the intercept. The reader should not be particularly disturbed by the meaning of the price of time for young children and the necessary extrapolation to estimate it. Since all individuals in the sample have completed at least ninth grade, costs of previous years of schooling, even if estimated incorrectly, are approximately constant across individuals. Stated in other terms, it is marginal cost that does the work in this model rather than total cost; and marginal cost should be estimated accurately by this method, even if total cost is not.

<sup>11</sup> Note that family income is absent from equation (10). This is not necessary for identification. Even if family income affected wages, we could still estimate the amount by which our predictions of  $E$  were off due to capital imperfections, i.e., nonwage considerations. What is important, however, is that income did not enter the equation significantly. In particular, its coefficient was 0.00000362 with a standard error of 0.00000894. This implies an elasticity of 0.015 with a standard error of 0.038. The upper bound of 95% confidence interval on the income elasticity is therefore 0.092, which is small indeed.

$$C(E, S) = \begin{cases} \sum_{j=1}^E H(j) W^*_j(.01) \left(\frac{1}{1+r}\right)^{j+5} & \text{for } E \leq 12 \\ \sum_{j=1}^{12} H(j) W^*_j(.01) \left(\frac{1}{1+r}\right)^{j+5} + \sum_{j=13}^E H(j) W^*_j(.01) \times \left(\frac{1}{1+r}\right)^{j+5} (1.5) & \text{for } E > 12 \end{cases} \quad (11)$$

where  $C(E, S)$  is measured in present dollars. The discount rate, at this point, is assumed to be the same for all groups and equal to 10%. It should be pointed out that equation (11) gives the level of costs generated not only at the attained level of schooling, but at any level of  $E$  that we wish to consider for any individual. I.e., once the endowment vector is specified, the cost of any level of schooling may be computed.

One difficulty is that in order to obtain the marginal costs necessary for equations (4), (5), and (6), one must differentiate (11) with respect to  $E$ . This yields a rather unwieldy expression as the result of the awkward form of (11). However, equation (11) can be approximated by a function of much more manageable form. Suppose we approximate (11) by the following function:

$$C(E, S) = \eta E^\gamma IQ^\alpha M^\beta N^\delta \exp[\theta D + \lambda R]. \quad (12)$$

Then the optimization condition, equation (2), becomes

$$P_E(IQ, M, N, D, R) = \gamma \eta E^{\gamma-1} IQ^\alpha M^\beta N^\delta \exp(\theta D + \lambda R). \quad (13)$$

Upon substituting (3) into (13) and solving for the optimal  $E$ , we get

$$E^* = [(a_0 + a_1 IQ + a_2 M + a_3 N + a_4 D + a_5 R)^{-1} \cdot \gamma \eta IQ^\alpha M^\beta N^\delta \exp(\theta D + \lambda R)]. \quad (14)$$

Equation (14) identifies  $E^*$  once the endowment vector is specified. From (11), it is theoretically possible to solve algebraically for (12) and to express  $C(E, S)$  as the right-hand side of (12) plus a remainder. This is an unnecessarily difficult task. By specifying arbitrary vectors of

$[E, S]$ , one can generate cost as defined in equation (11). Then, by using OLS, we can estimate (12) where the left-hand side of (12) is the right-hand side of (11). The parameters of (12) are obtained directly from a regression of  $\ln C(E, S)$  on the log of the right-hand side of (12). Note that nothing behavioral has occurred so far.<sup>12</sup> We are simply approximating one function by a more manageable second function and asking the computer to assist us in obtaining the parameters of approximation. As the  $R^2$  approaches 1, the approximation becomes better. As can be seen in equation (12), table 1, the  $R^2$  from this procedure is 0.97, so the price we pay for a more manageable cost function is quite low.

To recap, in the preceding few paragraphs a model is constructed that allows estimation of the optimal (attained) level of schooling simply

<sup>12</sup> See Lazear (1977b) for more discussion of this technological relationship.

TABLE 1.—REGRESSION RESULTS

Independent Variable	Dependent Variables		
	Equation (9) $\ln W_{69}$	Equation (12) $\ln C(E, S)$	Equation (3) $P_E$
$IQ$	0.0018 (0.0010)		2.366 (0.104)
$M$	0.0077 (0.0049)		12.96 (0.52)
$N$	-0.0100 (0.0068)		-4.81 (0.71)
$D$	0.1128 (0.0397)	0.1095 (0.0034)	85.4 (4.5)
$R$	-0.1151 (0.0256)	-0.1151 (0.0021)	122.6 (2.8)
$E_{69}$	0.0154 (0.0078)		
$\ln E_{75}$		1.412 (0.007)	
$\ln IQ$		-0.00039 (0.00886)	
$\ln M$		0.0730 (0.0035)	
$\ln N$		-0.0123 (0.0020)	
Family Income			
Age	0.0419 (0.0045)		
Constant	-0.1655 (0.1285)	5.317 (0.034)	649.7 (10.9)
$R^2$	0.2103	0.973	0.738
SEE	0.340	0.044	58.5
$N$	771	1,769	1,769

Note: Standard errors are in parentheses.

by knowing the endowment vector. That is, we construct a non-linear reduced form equation which appears as equation (14). This equation is based on the assumption that all individuals face the same borrowing costs. If individuals from wealthy homes face lower borrowing costs than those from poor families,  $\hat{E}^*$  as computed by (14) will tend to be too high for the poor and too low for the rich. If no systematic residual is observed, we can conclude that the assumption that individuals face the same borrowing cost is not violated by the data.

## II. Estimation

The data used for this analysis came from the National Longitudinal Survey (1975) on young men, 14–24 years old in 1966. These data are described in detail elsewhere (see Griliches (1976) or Lazear (1976)), but they consist of a five-year study of about 5,000 males with information provided on work history, education, and family background. For the purposes of this study, it was necessary that the individuals have completed schooling. Therefore, all observations were dropped for persons who were currently (in 1975) attending school.<sup>13</sup> Similarly, observations were dropped for which there was incomplete information on the relevant variables. Attained schooling is defined as the highest grade of schooling completed in 1975.

The first order of business is to estimate the wage equation from which schooling costs are derived. Since the wage that is relevant is the wage during the individual's educational years, the wage rate in 1969 among the nonstudent population was used. This reduces the number of observations but yields a more appropriate dependent variable in the wage equation. For all other calculations, the larger sample derived from the 1975 data was used. The results of the wage equation estimation are found in equation (9), table 1. Second, the cost function approximation is derived as equation (12) in table 1. The arbitrary vectors  $[E, S]$  were generated by using actual values found in the NLS, where  $E$  is the level of schooling completed in 1975. No real values were necessary, however. As can be seen,

<sup>13</sup> This introduces the possibility of selection bias. By 1975, however, most individuals in the sample had completed school, so that the bias is likely to be small, at worst. See Heckman (1977) for a method which deals with this problem.

the fit is very good ( $R^2 = .97$ ) so that the continuous and differentiable form of  $C(E, S)$  given by the right-hand side of (12) is a good approximation to the unmanageable true cost which, as equation (11), becomes the left-hand side of (12). Finally, from (13), one can infer  $P_E$  once attained  $E_{75}$  is known for each individual. These values of  $P_E$  are then regressed on the endowment characteristics to obtain the hedonic returns-to-schooling function. This appears as equation (3) in table 1. An errors-in-variables problem involved in the estimation of (13) is discussed in an appendix (available from the author upon request).

The model is now complete. All parameters are identified so that an estimated  $E^*$  can be obtained for all individuals from equation (14). If the model gives unbiased predictions, one would expect that the mean level of the difference between actual  $E_{75}$  and optimal  $E^*$ ,  $E_{75} - E^*$ , should be approximately zero (actually 0.081 as the result of nonlinearities).<sup>14</sup> The first row of

<sup>14</sup> Actually, the expected value of  $E_{69} - E^*$  is slightly greater than zero as a result of nonlinearization of the error term. This is seen as follows: From equation (4),

$$E_{75} = C_E^{-1}(C_E, S)$$

and from equation (5)

$$E^* = C_E^{-1}(P_E(S), S).$$

Estimates of  $E^*$  are obtained by substituting  $\hat{P}_E(S)$  for  $P_E(S)$  in (5). Since  $C_E = \hat{P}_E(S) + \epsilon$ , we rewrite

$$E_{75} = C_E^{-1}(\hat{P}_E + \epsilon, S)$$

and

$$\hat{E}^* = C_E^{-1}(P_E, S).$$

Define  $B \equiv IQ^\alpha M^\beta N^\delta \exp(\theta D + \lambda R)$ . Then, from (9),

$$\hat{E}^* = \left( \frac{\gamma \eta B}{\hat{P}_E} \right)^{1/(1-\gamma)}$$

and similarly,

table 2 reveals that  $\mathcal{E}(E_{75} - E^*)$  equals 0.074 with a standard error of 0.045. The sample mean does not differ significantly from its expectation and so, for the sample as a whole, numbers consistent with unbiased predictions appear to be obtained. Table 2 also tells us that for the sample as a whole,  $\mathcal{E}(\hat{P}_E)$  is 1,023 where  $\hat{P}_E$  is the predicted value of  $P_E$  from equation (3). Since the marginal value of a year of education in present terms must equal the marginal cost, the cost of the final year of schooling is, in period zero dollars, 1,023. The final year occurs, on average, at age 18, so the age 18 dollar value of the cost =  $1,023 e^{(.11)18} = \$6,188$ . This indeed seems reasonable.

### III. Rich versus Poor

Up to this point family income has not been introduced explicitly. Other things constant, individuals from wealthy families have been treated as if they were indistinguishable from individuals from poor families. This feature enters at two points. First, family income<sup>15</sup> did not

$$E_{75} = \left( \frac{\gamma \eta B}{\hat{P}_E + \epsilon} \right)^{1/(1-\gamma)}$$

Since  $\gamma = 1.412$ , we can approximate  $E_{75} - \hat{E}^*$  by

$$E_{75} - \hat{E}^* \approx \frac{(\hat{P}_E + \epsilon)^2 - \hat{P}_E^2}{(\gamma \eta B)^2}$$

On the assumption that  $B$  is fixed, and since  $\epsilon$  is orthogonal to  $\hat{P}_E$ ,

$$\mathcal{E}(E_{75} - \hat{E}^*) \approx \frac{\mathcal{E}(\epsilon^2)}{(\gamma \eta B)^2}$$

Using estimates of  $\epsilon^2$  and  $\gamma \eta B$  from the sample, one obtains that

$$\mathcal{E}(E_{75} - \hat{E}^*) \approx .081.$$

<sup>15</sup> Family income is defined as the median income of the father's current occupation according to the 1960 Census. This is more likely to reflect permanent income or parental wealth than is current income.

TABLE 2.—CELL MEANS

Sample	Sample Size	$E_{75}$	$E^*$	$E_{75} - E^*$	$\hat{P}_E$
Total	1,769	13.77	13.69	.074 (.045)	1,023
Low-Income Third	590	13.14	13.21	-.073 (.077)	973
Middle-Income Third	589	13.49	13.59	-.105 (.082)	1,023
High-Income Third	590	14.66	14.26	.403 (.076)	1,082

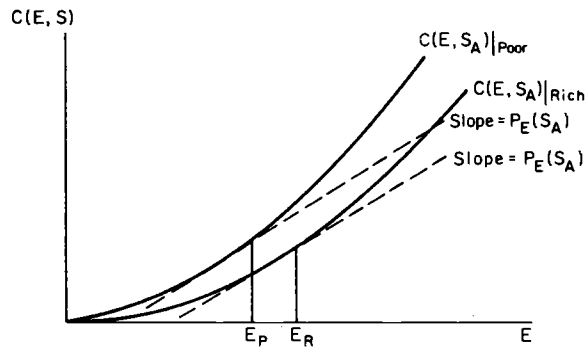
Note: Standard errors are in parentheses.

enter the wage equation (as an empirical fact rather than as an assumption). This means that there is no direct effect of family income on an individual's foregone earnings and cost of schooling. Further, it was assumed that all individuals faced the same interest rates throughout. Again, if individuals with wealthy parents can finance their education more cheaply than can those with less wealthy parents, one would expect the model to overpredict attained schooling levels for the poor and underpredict them for the rich. Otherwise stated, the model infers and predicts returns by saying that they are equal to marginal cost as measured in the absence of capital costs differences. But given  $S$ , the amount that employers pay workers for additional schooling is independent of workers' costs of acquiring it. Thus, although the true marginal returns function is the same, the true marginal cost functions would differ in a way not picked up in the model. This would result in a lower attained  $E_{75}$  than predicted for the high cost individual and a higher attained  $E_{75}$  than predicted for the lowest cost group. Table 2 provides evidence on this point. Specifically,  $\mathcal{E}(E_{75} - \hat{E}^*)$ , the mean of the difference between predicted "optimal" and actual levels of schooling, differs slightly, but statistically significantly, from its expectation ( $= 0.081$ ).  $E^*$  overpredicts attained levels of schooling for low-income individuals and underpredicts attained levels for high-income groups. This is consistent with the hypothesis that borrowing costs are higher for individuals from low-income families than for individuals from high-income ones. Thus, the assumption that all individuals face the same borrowing costs can be rejected at conventional levels of statistical confidence. But the more important question is, "How different are these borrowing costs across groups?"

To get at this issue, the question must be turned around. Suppose it is assumed that individuals who come from poor families do indeed face higher borrowing rates than those from wealthy homes. The situation would then be the one illustrated in figure 2.

Consider two individuals, A and B, who are similar in all respects except for family income. Since they have the same characteristics,  $P_E(S_A)|_{\text{poor}} = P_E(S_A)|_{\text{rich}}$ . That is, employers do not care about background characteristics that do not affect wages (and therefore productivity). Thus,

FIGURE 2.—OPTIMAL SCHOOLING: DIFFERENT BORROWING COSTS



the poor individual stops at  $E_P$  where  $C_E(E, S_A)|_{\text{poor}} = P_E(S_A)$ , and the rich individual stops at  $E_R$ . If  $E^*$  is calculated for the poor using the assumed high interest cost function, but the unique  $P_E(S)$  function for both rich and poor, the expectation of ( $\hat{E}^*$ ) will equal  $E_P$  if borrowing cost is in fact higher for the poor. Similarly, using the unique  $P_E(S)$  function, but the relatively low interest cost function for the rich, the expectation of ( $\hat{E}^*$ ) will equal  $E_R$ . If, on the other hand, the borrowing rate is in fact the same for both groups and equal to the assumed 10% of the rich, rich and poor individuals will stop at  $E_R$ . Using  $C_E(E, S_A)|_{\text{poor}}$  to estimate  $P_E(S_A)$  for the poor will yield a  $\hat{P}_E(S_A)$  that is too high for both groups (since only one  $P_E(S)$  function is estimated and it will include some observations that measure too high a value of  $P_E$ ). In fact,  $\hat{P}_E(S_A)$  will, since it is a weighted average, be smaller than  $C_E(E, S_A)|_{\text{poor}}$  but larger than  $C_E(E, S_A)|_{\text{rich}}$ . In this case, if the  $\hat{E}^*$  for the poor is calculated using the high interest cost curve, the expectation of  $\hat{E}^*$  will lie below  $E_R$ . But, for the rich, since  $\hat{P}_E$  is greater than true  $P_E$ ,  $\hat{E}^*$  will exceed  $E_R$ . Then if the higher finance cost assumption is correct, no systematic residuals will be observed. If incorrect,  $\mathcal{E}(E_{75} - E^*)$  will be positive for the poor and negative for the rich.

It is now necessary to respecify the cost function to reflect the wedge between lending and borrowing rates for the poor. The cost of the wedge depends upon the length of the repayment period. Suppose that it is assumed that the entire loan is paid back upon graduation. Then the cost of financing the  $j^{\text{th}}$  year of schooling in dollars of repayment year time is

$$\begin{aligned}
 &= F_j(1 + r_B)^{(E+5)-(j+5)} \\
 &= F_j(1 + r_B)^{E-j}
 \end{aligned}$$

where  $r_B$  is the borrowing rate. To convert this to present value, the lending rate should be used for discounting since it is at this rate that a dollar in year zero is turned into  $(1 + r_L)^t$  dollars in year  $t$  (by lending at that rate). Thus, the present value of the loan cost is

$$= \frac{F_j(1 + r_B)^{E-j}}{(1 + r_L)^{E+5}}$$

If  $r_B = r_L$ , then this becomes  $F_j/(1 + r_L)^{j+5}$  which is what is given in equation (7). Thus, for individuals from wealthy homes, cost is defined as in (7). For those from the lowest one-third<sup>16</sup> of family income homes, cost is defined as

$$C(E, S) = \begin{cases} \sum_{j=1}^E \frac{F_j(1 + r_B)^{E-j}}{(1 + r_L)^{E+5}} & \text{for } E \leq 12 \\ \sum_{j=1}^{12} \frac{F_j(1 + r_B)^{E-j}}{(1 + r_L)^{E+5}} + 1.5 \sum_{j=13}^E \frac{F_j(1 + r_B)^{E-j}}{(1 + r_L)^{E+5}} & \text{for } E > 12. \end{cases} \quad (7')$$

Letting  $r_L = 0.10$  and  $r_B$  vary from 0.10 to 0.15, the values were substituted in and equations (12) and (3) were re-estimated. Equation (12) was estimated separately for the poorest one-third and richest two-thirds since we are fitting different cost functions. However, as explained above, there is a unique  $P_E(S)$  function and therefore only one equation (3). Using these estimates,  $E^*$  is calculated for each individual. The summary statistics are contained in table 3.

In column 3 of table 3,  $r_B$  is assumed to be 0.15 for the low-income group and 0.10 for the higher income group. As predicted, the mean of the residuals is positive and large for the low-income group and negative and large in magnitude for the high-income group. This implies that we can reject at all conceivable levels of statistical significance that the borrowing costs to the poor are five percentage points higher than they are to the higher income group. The same is true when interest rate differences are assumed to be 0.1, 0.5,

and 0.25 percentage points different as in columns 4, 5, and 6 of table 3, respectively.

It is only when the difference is assumed to be one-tenth of one percentage point that we can no longer reject the null hypothesis that the two groups differ in borrowing cost. From the first section (table 2), we can reject with 95% confidence that the high-income and low-income groups face the same level of borrowing costs. But from table 3 we can reject with 95% confidence that the spread is 0.25 percentage points or greater. Therefore, we are very certain about the amount by which borrowing costs differ. They differ by an amount greater than zero but less than 0.25 of a percentage point with 95% confidence. The assumption that they differ by one-tenth of one percentage point (column 7 of table 3) seems to fit about exactly the expectation of  $\mathcal{E}(E_{75} - \hat{E}^*)$  equal to 0.081 for all groups.

#### IV. Some Additional Results

In an earlier version, a number of other findings were reported. That analysis<sup>17</sup> is merely summarized here:

One extension of the analysis considers some differences that arise when the sample is stratified by race. The most important finding in this regard is that when  $\mathcal{E}(P_E)$  is calculated by racial groupings,  $\mathcal{E}(P_E) = \$1,043$  for whites and only \$914 for blacks. Thus, the return to the marginal year of education for blacks is only 88% of that for whites. This is consistent with previous findings by Freeman (1974a) and Welch (1973). It is also found, however, that holding other things constant, blacks have higher optimal and actual levels of education than whites. This is because the lower returns to blacks are more than offset by the lower costs.<sup>18</sup>

Some of the specifications were altered in order to test the robustness of the model. The most important change was that the wage rate used in the calculations was altered to include the value of on-the-job training using the method of Lazear (1976, 1977a, and 1979). The main effect was an increase in  $\mathcal{E}(\hat{P}_E)$ ; the behavior of the residuals was not significantly altered. I.e., the difference between implicit borrowing costs for rich and poor remained very small.

<sup>16</sup> The actual figure is slightly less than one-third of the sample since there were a number of families with the tertial income level.

<sup>17</sup> Available from the author upon request.

<sup>18</sup> This finding is consistent with those of Freeman (1974b) and Jencks (1972).



TABLE 3.—CELL MEANS: NULL HYPOTHESIS  $r_L = .10$ ,  $r_B \geq .10$ 

(1) Sample	(2) Sample Size	(3) $\mathcal{E}(E_{75} - \hat{E}^*)$ $r_B = .15$	(4) $\mathcal{E}(E_{75} - \hat{E}^*)$ $r_B = .11$	(5) $\mathcal{E}(E_{75} - \hat{E}^*)$ $r_B = .105$	(6) $\mathcal{E}(E_{75} - \hat{E}^*)$ $r_B = .1025$	(7) $\mathcal{E}(E_{75} - \hat{E}^*)$ $r_B = .101$	(8) <sup>a</sup> $\mathcal{E}(E_{75} - \hat{E}^*)$ $r_B = .10$
Total	1,769	-1.781 (0.108)	.0005 (.0600)	.059 (.046)	.071 (.046)	.072 (.046)	.070 (.046)
High Income (family income $\geq \$4,188$ )	1,221	-4.059 (0.087)	-.494 (.056)	-.164 (.056)	-.008 (.056)	.081 (.055)	.145 (.054)
Low Income (family income $\leq \$4,155$ )	548	3.296 (0.126)	1.099 (.084)	.557 (.081)	.249 (.081)	.054 (.081)	-.097 (.082)

Note: Standard errors are in parentheses.

<sup>a</sup> This column uses separate cost functions so the results differ slightly from those found in table 2.

A second change was to respecify equation (11) such that foregone earnings were assumed to be zero during the first seven years of school (or approximately until age twelve). The entire model was re-estimated and all conclusions remained intact. This is to be expected, since early foregone earnings costs are quite unimportant in affecting estimates of the marginal cost of schooling, which drives most of the analysis.

Finally, in an earlier version, a different part of the sample was used and  $E_{69}$  was the relevant variable rather than  $E_{75}$ . In addition, the wage equation was linear rather than log-linear. The results in that draft are virtually indistinguishable from those in this draft.<sup>19</sup>

## V. Summary and Conclusion

This paper attempts to build a model of optimal schooling acquisition based on the premise that individuals are wealth-maximizers. By postulating a specific schooling cost function, one can infer the marginal costs and therefore returns to education. This return varies across individuals, and the model allows estimation of these variations with respect to specified endowment variables.

The primary task of the framework is to ascertain whether one's genetic and environmental endowment affects the ability to obtain schooling. It is easily established that endowment is important as a schooling determinant because it affects foregone earnings. The more subtle question, however, is, "Are particular groups with given endowments discriminated against by a capital

market mechanism?" The answer to this question appears to be, "Only very slightly." The assumption that wealthy and less wealthy individuals borrow at the same rate does result in systematic overprediction of the poorer individual's attained schooling level. However, we can reject with 95% confidence that the difference between implicit borrowing costs equals or exceeds one-quarter of one percentage point.

## REFERENCES

- Becker, Gary S., *Human Capital and the Personal Distribution of Income*, W. S. Woytinsky Lecture No. 1, University of Michigan, 1967.
- Bowles, Samuel, "Schooling and Inequality from Generation to Generation," *Journal of Political Economy*, Supp. 82, 2 (May 1972), 57-71.
- Freeman, Richard, "Implications of the Changing Labor Market for Minorities," in M. Gordon (ed.), *Higher Education and the Labor Market* (New York: McGraw-Hill, 1974a).
- , "Socio-economic Mobility and Black-White Economic Differences in the 'New Market' for Black Labor," Harvard Institute for Economic Research Discussion Paper No. 377 (Aug. 1974b).
- , *The U.S. Discriminatory System*, unpublished manuscript, 1975 (forthcoming book).
- Gilman, Harry, "Economic Discrimination and Unemployment," *American Economic Review* 55 (Dec. 1965), 1077-1096.
- Griliches, Zvi, "Wages of Very Young Men," *Journal of Political Economy* 84 (Aug. 1976), 569-586.
- Heckman, James J., "Sample Bias as a Specification Error," *Econometrica* 47 (Jan. 1979), 153-162.
- Jencks, Christopher, et al., *Inequality* (New York: Basic Books, 1972).
- Lazear, Edward P., "Age, Experience, and Wage Growth," *American Economic Review* 66 (Sept. 1976), 548-558.
- , "Schooling as a Wage Depressant," *Journal of Human Resources* 12 (Spring 1977a), 164-176.
- , "Education: Consumption or Production," *Journal of Political Economy* 85 (June 1977b), 569-597.
- , "The Narrowing of Black-White Wage Differentials

<sup>19</sup> Available from the author upon request.

- Is Illusory," *American Economic Review* 69 (Sept. 1979), 553-564.
- Leibenstein, Harvey, "Allocation Efficiency vs. 'X-Efficiency,'" *American Economic Review* 56 (June 1966), 392-415.
- Marris, R., and A. Wood, *The Corporate Economy* (Cambridge, Mass.: Harvard University Press, 1971).
- Radner, Roy, and Michael Rothschild, "On the Allocation of Effort," *Journal of Economic Theory* 10 (June 1975), 358-376.
- Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy* 82 (Jan./Feb. 1974), 34-55.
- Simon, Herbert A., *Models of Man* (New York: John Wiley & Sons, 1957).
- Stigler, George J., "The Existence of X-Efficiency," *American Economic Review* 66 (Mar. 1976), 213-216.
- Welch, Finis, "Black-White Differences in Returns to Schooling," *American Economic Review* 63 (Dec. 1973), 893-907.
- Williamson, J. D., "Profit, Growth, and Sales Maximization," *Economica* 33 (Feb. 1966), 1-16.