On Dogmatism in Human Capital Theory

by

Alan S. Blinder[*]

ON DOGMATISM IN HUMAN CAPITAL THEORY

Alan S. Blinder

I.   Introduction

Every scholar in the fields of human capital or income distribution
is indebted to Jacob Mincer [12, 13] for his pioneering work in integrating
these two branches of economic theory. His contributions have been so
forceful and original that there is now a danger that he may have succeeded
too well. Tools like the human capital earnings function, which in Mincer's
hands are subtle and useful aides to empirical research, can become blunt
instruments which retard rather than advance knowledge, if wielded with
dogmatism rather than insight. In particular, there seems to be a viewpoint
afloat -- which I certainly would not attribute to Mincer, but which seems
to characterize Mark Rosenzweig and Jack Morgan [15] -- that Mincer has
"proven" that the logarithm of earnings is a linear function of years of
schooling (S) and a quadratic function of a variable "j" defined as:

(1) $j = A - S - 5,$

where A is age.

Rosenzweig and Morgan accuse me, quite correctly, of violating this
doctrine. They maintain, quite incorrectly, that my coefficient estimates
are therefore biased. Briefly, their argument runs as follows:

(i)   Blinder's regressions utilize an incorrect functional form.

(ii)   In particular, he uses A and A^2 as regressors, though it
is known from Mincer's work that the true functional form uses j and j^2
instead.

(iii) When we run the two functional forms on a set of data different from Blinder's, we find: (a) that the schooling coefficient for white men is almost the same in each specification; (b) but the schooling coefficient for white women is substantially higher in the Mincer version.

(iv) We therefore conclude that Blinder underestimates the returns to education for women, and correspondingly overestimates the extent of discrimination.

Is there any merit to this argument? I certainly would never quarrel over (i). No one has revealed the "true" functional form to me. However, (ii) strikes me as a gross overstatement of what Mincer has accomplished. What he has done is taken a set of assumptions, some of which are demonstrably false but which may nonetheless be reasonable approximations, and deduced from them a precise functional form for regression analysis. In Section II, I detail these assumptions, and point out the difficulties with some of them. The argument is mainly theoretical which, of course, leaves open the question of the empirical usefulness of the assumptions. Mincer's work alone demonstrates that they are indeed useful, and a recent paper by Heckman and Polachek [10] underscores this opinion.[1]/

As to item (iii), I have serious reservations about Rosenzweig and Morgan's interpretation of their own results. Moreover, what they offer as "Blinder's equation" is not my equation at all -- both the data base and the functional form are different. But, even waiving that, I shall show that calling my coefficients biased estimates of their parameters is just a misuse of language. In Section III, I defend my 1973 paper against Rosenzweig and Morgan's critique. While it is easy to show (given the right assumptions) that my estimates may be biased, inconsistent, or have any

other bad property that can be imagined, I remain unconvinced that they have demonstrated that "...Blinder's results are likely to be a significant overstatement of the size of the schooling component of the male-female earnings differential attributable to discrimination." [15, p.

II.   Schooling, Experience and Mincer

Through an ingenious set of simplifying assumptions, Mincer [13] derives the following estimating equation:

$$(2) \quad \log Y_t = a_0 + a_1 S + a_2 j + a_3 j^2 + u ,$$

where Y_t is earnings at age t, j is defined by (1), and u is a stochastic error. This expression, which I shall hereafter call the "Mincer equation," should be understood for what it is -- a useful empirical approximation to the complex dependence of earnings on schooling and post-school investments in human capital. Mincer certainly recognizes it as such, and, in fact, offers several alternative regression models. [13, pp. 83-92]

I.1 Mincer's Assumptions

What, then, are the assumptions which lie behind (2)? Here I shall list them briefly, with page references to Mincer [13], and derive (2). Then, in the next section, I shall comment on their validity. I reiterate that my only purpose is to debunk the view that (2) is the "true" model, not the view that it may be useful in empirical work.

   A1: In the absence of post-school investments, an individual with S years of schooling would have a flat age-earnings profile. (pp. 8-10)

   A2: In the absence of post-school investments, the present discounted value of lifetime earnings would be equal for all individuals, regardless of how long they stay in school. (pp. 10-11)

By definition, the present value of earnings for a person with S years of schooling is:

$$V_S = \int_S^L E_t e^{-rt} dt \ ,$$

where E_t is earnings and L is the age of retirement. Under A1, we can substitute a constant, E_S, for E_t , so that A2 implies:

$$(3) \quad \frac{E_s}{r} (e^{-rS} - e^{-rL}) = \text{constant.}$$

<u>A3:</u> The number of years spent at work, N, is independent of the number of years spent in school. (pp. 8-11)

Since

$$(4) \quad L = N + S,$$

this means that each additional year of schooling postpones retirement by exactly one year.[2/] Substitution of (4) into (3) and simplification leads to:

$$E_S = E_0 e^{rS} \ ,$$

or,

$$(5) \quad \log E_S = \log E_0 + rS \ ,$$

so that in the absence of post-school investments, competition would make log earnings a linear function of years of schooling.

But there are post-school investments ("on-the-job training"). To accomodate these, Mincer distinguishes among the following concepts:

E_t = potential earnings at age t

Y_t = actual earnings at age t

C_t = human investment (in dollars) at age t $\equiv E_t - Y_t$.

k_t = investment ratio at age t $\equiv C_t / E_t$.

He then assumes:

> A4: The "rate of return" to all post-school investments in human capital is a constant, ρ , in the sense that investing ΔC_1 at $t = t_1$ increases potential earnings by $\Delta E_t = \rho \Delta C_1$ for all $t > t_1$. (p. 19)

Given A4, the following identity holds:

$$E_t = E_{t-1} + \rho C_{t-1} = E_{t-1}(1 + \rho k_{t-1}).$$

Solving this recursion formula yields:

$$E_t = E_S \prod_{i=S+1}^{t-1} (1+\rho k_i) ,$$

or,

$$\log E_t = \log E_S + \sum_{i=S+1}^{t-1} \log(1+\rho k_i) .$$

For small ρk_i , this can be approximated by:

$$\log E_t = \log E_0 + rS + \rho \sum_{i=S+1}^{t-1} k_i ,$$

where (5) has been used. In continuous time, this translates to:

$$(7) \quad \log E_t = \log E_0 + rS + \rho \int_0^X k(\tau)d\tau ,$$

where τ is a dummy variable of integration, and X ("experience") measures time spent at work since the end of schooling.

> A5: The investment ratio during the post-school investment period, $k(\tau)$, declines linearly with τ , beginning at k_0 at $\tau=0$ and reaching zero at $\tau=T$. (p. 85)

In symbols, this states:

$$(8) \; k(\tau) = k_0(1 - \frac{\tau}{T}) .$$

Substituting this into (7) and integrating gives:

$$(9) \quad \log E_t = \log E_0 + rS + \rho k_0 X - \frac{\rho k_0}{2T} X^2 ,$$

which is almost the desired result. The only hitch is that actual and potential earnings are not equal during the on-the-job training period, i.e., for the first T years of work. Instead:

$$Y_t = E_t(1-k_t),$$

or,

$$\log Y_t = \log E_t + \log(1-k_t) .$$

Substituting both (8) and (9) into this expression yields the human capital earnings function:

$$(10) \quad \log Y_t = \log E_0 + rS + \rho k_0 X - \frac{\rho k_0}{2T} X^2 + \log(1-k_0+\frac{k_0}{T} X)$$

But this is not quite the Mincer equation. To obtain the latter, he needs:

A6: The function $f(X) = \log(1 - k_0 + (k_0/T)X)$ can be
approximated by a second-order Taylor series expansion. (pp. 90-91)

This enables Mincer to write:

$$(2') \quad \log Y_t = a_0 + a_1 S + a_2 X + a_3 X^2 + u ,$$

which is amenable to empirical analysis using linear regression techniques on some data sets. However, both Mincer [13] and Rosenzweig and Morgan [15] use Census data, which does not offer any direct measurement of accumulated work experience. So Mincer introduces:

A7: During schooling, no time is spent in the employed labor
force. After schooling, all time is spent in the employed labor
force. (p. 84)

This allows Mincer to replace the variable X in (2') -- which is observable in principle but generally not observed in practice -- by the observed variable j defined by (1). Thus he finally arrives at (2).

I.2 Evaluation of the Assumptions

Let me now consider the applicability of each of these assumptions
to the real world, and, in particular, to the contemporary United States.
I see no compelling reason to believe A1. It is true that exponential
depreciation can easily be accomodated.[3/] But who is to say that depreciation
of earning power is exponential? As a rough approximation to facilitate
theoretical or empirical work: fine. As an article of faith: never.
Even if exponential depreciation is accepted, it must be remembered that
the dependent variable is earnings, not wage rates, and the former are also
affected by labor supply decisions. Presumably, then, A1 also requires a
flat age-hours profile. There are excellent theoretical and empirical
reasons to question the veracity of this assumption.[4/]

Assumption A2 seems even more dubious. First of all, present values
would only be equated across schooling levels if (a) all individuals had
equal access to capital markets; (b) all individuals were of equal ability,
both in producing income and in producing human capital; (c) jobs at different
schooling levels did not differ as to riskiness or nonpecuniary benefits.
This should be enough to shake one's faith in A2. But, even if all this is
waived, a technical question arises. Why should it be the present discounted
values excluding post-school investments which are equated by competition?
People presumably are aware of the opportunities for investing on the job.
In perfect markets, then, it should be present discounted values inclusive
of OJT that are equated. This, of course, carries the absurd implication
that all persons receive the same lifetime earnings.

Mincer (pp. 8-9) offers some empirical evidence in support of A3, and the assumption may well be roughly correct. Still, in a human capital model extended to include labor supply decisions, the age of retirement (L) would be an endogenous variable. The theoretical question then would be: In such a model, would a parameter change which lengthens the optimal schooling period by dS also shorten the optimal retirement period by dL=dS? It would be nice to have a theoretical model which exhibits this property, but I know of none.

Since lifetimes are finite, what Mincer calls the "rate of return" to OJT is not an "internal rate of return" as conventionally defined. That is, it is not the interest rate which equates the discounted value of future benefits to the current costs. Since as age advances the earning span $\underline{remaining}$ declines, Mincer's assumption A4 means that the internal rate of return falls with age.$^{5/}$ Optimal behavior in a perfect capital market with constant interest rates would not result in such a behavior pattern. Instead, the fraction of potential earnings invested would decline with age in such a way as to equate the $\underline{marginal}$ internal rate of return with the interest rate at each moment.$^{6/}$ Thus optimal behavior would make what Mincer calls the "rate of return" rise with age. Of course, for young workers, the remaining earning span is probably long enough to treat as infinite for all practical purposes. But A4 can cause problems in the analysis of older workers, where the finiteness of life becomes more crucial.

Whether or not the investment ratio, k_t, declines linearly, as assumed in A5, is, of course, an empirical issue. However, as Mincer correctly indicates (pp. 13-17), the time profile of k_t should be deduced from optimizing behavior. Haley's [9] recent explicit solution to the Ben-Porath

model [4] gives a nonlinear solution. In my research on the subject I have never encountered a model, nor even a single special case, that leads to linear decline in k_t. Remember, I do not wish to argue that linearity is not a useful empirical approximation, only that other approximations may also be useful.

In a sense, A6 is not an assumption at all. Any function has a Taylor series representation. I only wish to point out that the function $f(X)$ is definitely not quadratic, so the Taylor series is just an approximation. Equation (2') may well be an excellent approximation to equation (10), but to rule alternative functional forms out of court makes no sense. Note also that the Taylor series approximation destroys the interpretation of the coefficients a_2 and a_3 in (10) as simple functions of k_0, ρ and T.

II.3 Experience versus "j"

Mincer's assumption A7 merits special consideration. Clearly, investigators afforded the luxury of having actual measures of experience have no use for Mincer's clever proxy variable, j. Malkiel and Malkiel [11], for example, had one of the best microdata sources imaginable and therefore did not need j. Interestingly, they experimented with it anyway, and found that j was a fairly good proxy for males, but not for females.[7/] Most data sets, however, are not so rich. No doubt Mincer's proxy is therefore useful in many applications. When will it be useful? As A7 suggests, it should do rather well for groups with continuous work histories -- uninterrupted by childbearing, service in the armed force, spells of unemployment and the like. How one might go about identifying these people in the absence of actual work histories is a good question. However, I think few

economists would quarrel with using the proxy for prime-age white males.
Using j for females, however, is hazardous, as both Mincer and Polachek [14]
and Rosenzweig and Morgan [15] clearly recognize. Even if attention is
restricted to women who have never married, as Rosenzweig and Morgan's is,
the problem remains if females either (a) suffer more unemployment, or
(b) have lower labor-force participation than males.[8]

What happens to regression estimates if j is used as a proxy for X
in inappropriate circumstances? Letting e denote the discrepancy between
X and j (i.e., $j = X + e$), it is apparent that measurement error attaches
to two of the regressors in the Mincer equation, j and j^2. The situation
is more complicated than the classical errors-in-variables case because
(a) two variables are measured with error, and the measurement errors are
highly correlated; (b) each e is nonnegative, so the mean error is positive.
In general, all that can be established is that every coefficient in the
regression will be biased. No general results on the direction of the bias
can be proven, but the simple case of a single variable measured with an
error of zero expected value can help us make an educated guess. In that
case it is known that, regardless of what other regressors appear in the
equation, the coefficient of the variable measured with error will be biased
towards zero. Suppose only schooling and j (as a proxy for X) appeared in
the regression. Then the coefficient of j would be biased downward, so
the coefficient of schooling would have to be biased upward. Of course,
the Mincer equation is more complicated than this, but Malkiel and Malkiel's
results bear out our intuitive guess. It can be seen from Table 1 that
measurement error hardly changes the constant or the coefficient of experience

squared. However, it substantially reduces the experience coefficient while raising the schooling coefficient. And this is true for men as well as for women.

TABLE 1

Example of Bias from Errors in Variables

| | Schooling | Coefficient of: | | Constant |
		j or X	j^2 or X^2	
Males				
Equation using X	.081	.054	-.001	8.70
Equation using j	.091	.044	-.001	8.67
Females				
Equation using X	.066	.059	-.001	8.50
Equation using j	.078	.037	-.001	8.57

SOURCE: Malkiel and Malkiel [11], Tables 1 and 2.

Ironically, while Rosenzweig and Morgan suggest that my schooling coefficient is biased downward because I eschew the use of j, this evidence leads to the conjecture that theirs is biased upward because they use it!

Some results from my 1973 paper further illustrate just how bad the j proxy can be for women.[9/] Owing to equation (1), the derivative,

$$\frac{\partial \log W}{\partial j} \bigg|_S = \text{constant}$$

which Rosenzweig and Morgan want me to estimate is necessarily equal to the derivative,

$$\frac{\partial \log W}{\partial A} \bigg|_S = \text{constant}$$

which I do in fact estimate. My value for the latter for white females is [5, p. 451] -.011 + .00018A, which is negative up to about age 61.

I think this suggests caution in interpreting the coefficient as the
effect of experience on earnings. I certainly never invited such an
interpretation.

II.3 Earnings versus Wage Rates

One important difference between my work and the Mincer equation,
a distinction with Rosenzweig and Morgan gloss over much too lightly,
is that my dependent variable is logW, not logY. They glibly switch from
logW in their equations (1) and (2) to logY in their equation (3), without
so much as a mention of the change. The difference is much more than a
debating point.

The version of human capital theory developed by Mincer [12, 13],
Becker [1,2] and others assumes that each individual maximizes lifetime
earnings. This, of course, makes it impossible for the model to say anything
about hours of work.$\underline{10}/$ Mincer is fully aware of this shortcoming. He
notes in [13]:

> "To the extent that hours of work vary over the life
> cycle, the profile of annual earnings is affected...growth and
> decline of earning capacity is likely to induce a corresponding
> pattern of hours of work...Hence, the growth of observed annual
> earnings leads to overestimates of investments in human capital
> or of rates of return... The analysis of the relation between
> hours of work and human capital investments is not theoretically
> integrated into the present model." (pp. 22-23)

If, in fact, it is a utility function defined over goods, leisure, and
perhaps other variables, which is "really" being maximized, what sense can
be made of the income maximization model? At first one might hope that the
theory could be rescued as a model of the determination of wage rates, i.e.,
of potential earnings, rather than of actual earnings. Could it be that
individuals first maximize their lifetime "full income," and then allocate

this between leisure and consumption? No, because the optimal level and pattern of human investment depends on how hard the individual wants to work later in life. When labor supply is endogenous, the "rate of return" on human capital is not a valid proscriptive device, though it can always be calculated ex post.

There are two more promising routes. First, if the utility function depends on consumption and leisure, and if the entire age-leisure profile is fixed exogenously, and if the utility function is separable in consumption and leisure, then each individual will first maximize his lifetime income (actual, not potential), and then worry about allocating it over his life cycle. These may be an inordinately restrictive set of circumstances. A better route, it seems to me, is to develop a full-blown utility maximization model, determine the optimal life profiles of human investment and labor supply simultaneously, and then see whether the investment profile is more or less as predicted by the income maximization model. Yoram Weiss and I do this in [7], and find that, for those phenomena which the income maximization model is capable of addressing, its implications are broadly consistent with the utility maximization model.

Ignoring labor supply in theory, Mincer adopts the empirical expedient of adding $\log H$ (H=weeks worked per year) to equation (2), and Rosenzweig and Morgan follow his precedent. What can be said of this ad hoc procedure? Obviously,

$$(11) \quad \log Y = \log W + \log H.$$

Suppose we have some model (such as the Mincer equation) for $\log W$:

$$(12) \quad \log W = \sum_k \beta_k Z_k' \,,$$

where the Z_k' are a set of regressors, and some other model of hours of work:

$$(13) \quad \log E = \sum_k \gamma_k Z_k'' + \sigma \log W ,$$

where the Z_k'' are another set of regressors. We can, of course, substitute (12) into (13) and then both of these into (11) to derive a regression for $\log Y$:

$$(14) \quad \log Y = \sum_k \delta_k Z_k ,$$

where Z is the union of Z' and Z''. But care must be exercised in interpreting the coefficients. For example, if variable Z_1 enters both Z' and Z'', then:

$$\delta_1 = \gamma_1 + (1+\sigma)\beta_1 ,$$

so both human-capital formation and labor-supply responses are embodied in δ_1.

In any case, $\log H$ certainly does not appear in (14). So what is the meaning of the coefficient of $\log H$ estimated by Rosenzweig and Morgan (as by Mincer) in the equation:

$$(15) \quad \log Y = \sum_k \alpha_k Z_k + \phi \log H ?$$

Presumably, it connotes the partial derivative $\frac{\partial \log Y}{\partial \log H}$ when all the Z's are held constant. But if the Z_k''s are constant, then W should be constant by (12). And if the Z_k'''s are also constant, then H is constant by (13). Can any meaning be attached to a partial derivative which varies H while simultaneously holding all the determinants of H constant? Euqation (15) only makes sense if $\phi = 1$, for then it is just (12).

II.4 Where Do We Go from Here?

Where do all these theoretical speculations lead us? Certainly not to discard all of Mincer's important work. But rather, I think, to reconstitute the regression analysis with wage rates as the dependent variable,[11] and to be less doctrinaire about the functional form. An eclectic model, conforming to the spirit, though not the letter, of Mincer might be:

$$(16) \quad \log W = f(S, X_1, X_2, \text{other variables}),$$

where X_1 denotes "relevant" (to the present job) experience, and $X_2 = X - X_1$ denotes all other experience. As Malkiel and Malkiel [11, pp. 694-695] correctly point out, there is every reason to believe that the two should affect wages differently.[12] The equation which I estimated in my 1973 paper is one of many imaginable variants of (16) -- not the best that could ever be devised, just the one that made the most sense to me at that stage of my life cycle. Of course, I was not blessed with data on X_1 and X_2. I did, however, have measurements of one important part of X_1: length of time on the present job. Though Rosenzweig and Morgan [15, p. n] chide me for using this information in lieu of j, I still do not view it as a mistake. X_2 can, of course, be written:

$$X_2 = j - X_1 - \text{time unemployed} - \text{time out of the labor force}.$$

Since I lacked data on the last two components, and had only partial data on X_1, I simply used A as a proxy for both X_2 and the remainder of X_1. I might just have well have used j, but I doubt that it would have changed things much (see below). A long list of "other variables" was offered in my 1973 paper. Again, alternative lists might be as good or better. But the strict human capitalist's dogma that nothing matters save S, j and j^2 is just too much to swallow.

III. Bias, Bias, Who's Got the Bias?

At least provisionally, it seems that either my 1973 equation, or the Mincer equation (Rosenzweig and Morgan's equation (2)), or the equation which Rosenzweig and Morgan falsely attribute to me (their equation (1)), are all candidates for empirical representations of (16). In this light, let me examine their criticism of my work.

As previously noted, my specification allows the dependence of $\log W$ (not $\log Y$) on schooling to follow an arbitrary step function. This is far more flexible than the Mincer equation, which imposes linearity. In fact, my estimated coefficients do not look at all linear. But waiving that, and also waiving the fact that wages are not earnings, suppose I estimated a schooling coefficient with the interpretation:

$$\text{(a)} \quad \frac{\partial \log Y}{\partial S}\Big|_A = \text{constant}$$

while they wanted me to estimate a schooling coefficient with the interpretation:

$$\text{(b)} \quad \frac{\partial \log Y}{\partial S}\Big|_j = \text{constant} \quad .$$

It hardly needs saying that these are simply not the same things. How could one ever be a biased estimate of the other? "My" coefficient measures the marginal impact on $\log Y$ when S increases, <u>for people of the same age</u>. Theirs considers raising S with j constant. This means that <u>both age and years of schooling rise together</u>. Either or both may be of interest, but it is little wonder that the latter gives bigger numbers.

This line of reasoning makes clear how one can use either their equation (1), which I write as:

$$\text{(17)} \quad \log Y = \beta_0 + \beta_1 S + \beta_2 A - \beta_3 A^2 + \beta_4 \log H \quad (\beta_i > 0),$$

or their equation (2), which I write as:

$$(18) \quad \log Y = b_0 + b_1 S + b_2 J - b_3 J^2 + b_4 \log H \quad (b_i > 0) \; ,$$

to get alternative estimates of the same partial derivative. Consider

first expression (a). In (17), it is just β_1. Using (18), it is

computed as $b_1 - b_2 + 2b_3 J$. Based on Rosenzweig and Morgan's equations,

the two estimates of the same parameter (evaluated at the mean j in

the case of equation (18)) are shown in Table 2. Similarly, it is easy to

compute expression (b) from equation (17) as $\beta_1 + \beta_2 - 2\beta_3 A$. The two

Table 2

Marginal Effect of S on logY, Given A

	Males	Females	Difference
From equation (17)	.070 (.004)	.108 (.016)	−.038
From equation (18)	.066	.105	−.039

SOURCE: Regressions in [15], Table II.
Note: The numbers in parentheses are standard errors.

sets of estimates of (b) are given in Table 3. I submit that the differences

Table 3

Marginal Effect of S on logY, Given j

	Males	Females	Difference
From equation (17)	.081	.128	−.047
From equation (18)	.078 (.004)	.123 (.016)	−.045

SOURCE: Regressions in [15], Table II.
Note: The numbers in parentheses are standard errors.

between these two sets of estimates are trivial enough to be ignored. Furthermore the gaps between the male and female coefficients -- which are what matter for measuring discrimination -- are virtually indistinguishable in the two equations.

To be sure, I do not claim that the two models are identical, only that their resemblance is striking. To see exactly how they differ, suppose that the Mincer equation, (18), is true. Substituting j out from (1), and simplifying gives $\log Y$ as a function of A and S:[13]/

$$(19) \quad \log Y = \gamma_0 + \gamma_1 S + \gamma_2 S^2 + \gamma_3 A + \gamma_4 A^2 + \gamma_5 A \cdot S + \gamma_6 \log H$$

where

$$\gamma_0 = b_0 - 5b_2 - 25b_3 , \quad \gamma_3 = b_2 + 10b_3 ,$$
$$\gamma_1 = b_1 - b_2 - 10b_3 , \quad \gamma_5 = 2b_3 ,$$
$$\gamma_2 = \gamma_4 = -b_3 , \quad \gamma_6 = b_4 .$$

So, if the Mincer equation holds, equation (17) mistakenly omits two variables: S^2 and $A \cdot S$. My own equation, of course, allows for non-linearity in S. But it does omit the interaction. I plead nolo contendere to one count of variable-ommission, but doubt that it did much harm. It is worth noting that Rosenzweig and Morgan, or anyone who uses the Mincer equation, omit all the "other variables" of equation (16). If my results, and those of several others, make any sense at all, perhaps twenty or thirty relevant variables are thereby dropped. I suggest that the resulting bias may be very serious indeed.

Furthermore, Section II documents why I am not convinced that (18) is the true model of the world. Suppose I take a more eclectic approach. A rather general representation of (16), suitable for data sets where only

S and A are actually observed, would be:

$$\log Y = g(A,S) + \gamma_6 \log H \ ,$$

where weeks worked are again appended in Mincer's <u>ad hoc</u> way. Now expand $g(A,S)$ in a Taylor series up to second-order terms.[14] What do you get? Just equation (19), but without the restrictions on the parameters. Thus an eclectic economist can, indeed should, view both (17) and (18) as special cases of (19). Equation (17) -- which Rosenzweig and Morgan attribute to me -- assumes $\gamma_2 = \gamma_5 = 0$. Equation (18) -- the Mincer specification -- assumes $\gamma_2 = \gamma_4 = -\frac{1}{2}\gamma_5$. Each of these is, in principle, a pair of testable restrictions on a general model. However Rosenzweig and Morgan test neither set.

While I lack the information necessary to perform the F-tests, Rosenzweig and Morgan have kindly supplied me with the sum of squared residuals from their regressions.[15] They are given in Table 4. Equation (17) has (slightly) smaller residuals in each case. Since each imposes two linear restrictions on (19), it is clear that (17) "wins" the F-contest for both males and females.

<div align="center">

Table 4

Sums of Squared Residuals
</div>

	Males	Females
Equation (17)	1435	257
Equation (18)	1447	259

SOURCE: Private correspondence from M. R. Rosenzweig

That is, if the data permitted rejecting the restrictions in (17), they would also reject the restrictions in (18); but the converse is not true. In a word, had Rosenzweig and Morgan performed the F-tests, they could not have accepted (18) while rejecting (17).

Of course, the same is true, *pari passu*, of my equation, since it uses S in far more general form than the quadratic. I hasten to add that I derive little comfort from the fact that (17) "wins" on this particular data set. With different data, perhaps even with the data used in my 1973 paper, the Mincer equation might prevail. My point is that the issue is one of empirical usefulness, not theoretical purity.

IV. Some Corrections

Let me take this opportunity also to correct two misleading statements and a typographical error in my 1973 paper.

At the start of their paper, Rosenzweig and Morgan quote my statement that "men earn much larger wage increments for advancing to higher educational levels..." [5, p. 448]. I regret that this is slightly misleading. As the coefficients in my Table A-1 (p. 452) indicate, men gain more from advancing to sixth grade, from graduating from high school, and from attaining the college degree. However, women gain more (in percentage terms) from entering ninth grade, entering college, and achieving an advanced degree.[16/] I should have said that, because of the educational distribution of white women, my expression for the portion of the wage differential attributable to differing coefficients gives more weight to the male advantages. [5, p. 439]

In checking my calculations on this point, I discovered a typographical error in Table A-1 [5, p. 452]. The coefficient in the white female equation for advancing to grades 6-8 should be -0.179, not +0.179.

Finally, I erroneously gave the impression in footnote 3, p. 438, that the decomposition technique I offered is the only one which yields a ready economic interpretation for every term. Ronald Oaxaca points out to me, quite correctly, that I could have used the low-group's wage equation to evaluate the endowment differences and written:

$$\overline{Y}^H - \overline{Y}^L = \sum_j \beta_j^L (\overline{X}_j^H - \overline{X}_j^L) + \sum_j \overline{X}_j^H (\beta_k^H - \beta_j^L) + \beta_0^H - \beta_0^L \quad ,$$

which is just as valid as the decomposition I employed.

1. They consider a variety of functional forms for the dependence of log earnings (or log wage rates) cn S and j , and conclude that Mincer's specification turns in the best overall performance.

2. An alternative assumption, which is equivalent for this purpose, is that the length of life, L , is infinite.

3. Under exponential depreciation, the earnings profile would be:

$$E_t = E_S e^{-\delta(t-S)}$$

So discounted earnings would be:

$$V_S = E_S \int_S^{N+S} e^{-\delta(t-S)} e^{-rt} dt = E_S e^{\delta S} \int_S^{N+S} e^{-(r+\delta)t} dt$$

$$= \frac{E_S e^{\delta S}}{r} e^{-(r+\delta)S}(1-e^{-rN}) = \frac{E_S e^{-rS}}{r}(1-e^{-rN}) ,$$

from which (5) follows.

4. See [6, Ch. 3], [7], [16].

5. The relation between the two rates is exceedingly simple. The internal rate of return, denoted by i is defined implicitly by:

$$1 = \int_0^R \rho e^{-it} dt = \frac{\rho}{i}(1-e^{-iR}) ,$$

where R is the number of years remaining in the earning span. Clearly, $\rho > i$. Further, as $R \to \infty$, $i \to \rho$, so Mincer's ρ is the rate of return that would be realized in an infinite lifetime. It is easy to see $di/dR > 0$. Just define the benefits from a dollar of investment as:

$$B(i,R) = \int_0^R \rho e^{-it} .$$

Obviously, $\partial B/\partial i < 0$ while $\partial B/\partial R > 0$. Thus if R rises, i must also rise to compensate. Conversely, for a given i, ρ must fall as R rises. Note that Mincer's "error" is quite trivial when R is large (i.e., for young men). However, for older men the differences can be substantial.

6. Insightful readers may find this result too obvious to require proof, but a proof is given in [7].

7. The accuracy of j as a proxy should not be overstated, even for men, especially in view of the very small biases which Rosenzweig and Morgan claim I have. See Table 1.

8. Rosenzweig and Morgan provide a vivid example of the confusion that can arise from losing sight of the difference between X and j. They are surely correct in stating that "in a sample containing married women, the mean work experience of men will be greater than that of women." (p.   ). But they then mistakenly assert that in my sample this is also true of j : "the sample work experience of males measured by $A - S - 5$ exceeds that of females." (p.   ) This is simply untrue. As my Table 3 [5, p. 448] suggests, the differences between the sexes in mean A and mean S are trivial. In their sample, which includes only never-married women, mean j is indeed lower for females than for males. But this is because the women are younger on average. (See their Table I, p.   .)

9. My female sample includes all household heads, regardless of whether they were ever married. So these women have a much spottier work history than those studied by Malkiel and Malkiel or by Rosenzweig and Morgan.

10. For an exception to this, see Becker and Ghez [3].

11. Prime-age white males suffer little unemployment and rarely withdraw voluntarily from the labor force. Thus the distribution of hours of work in this subpopulation is concentrated around 2,000 hours per year. Since $Y = HW$, if H is almost a constant, Y will be nearly proportional to W. Thus Mincer's empirical work may still be meaningful in this context.

12. In fact, they detect no effect of X_2 on earnings. [11, p. 696]

13. Rosenzweig and Morgan [15] perform a similar exercise in deriving their equation (3). They have an error, but it is of no consequence either for their critique or for my response.

14. This is the econometric specification advocated for production functions by Christensen, Jorgenson and Lau [8] as a very flexible functional form.

15. I wish to thank them for this information.

16. Incidentally, this points out how forcing linearity in schooling can distort things.

REFERENCES

1. G. S. Becker, Human Capital. New York: Columbia University Press for National Bureau of Economic Research, 1964.

2. _____, Human Capital and the Personal Distribution of Income. W. S. Woytinsky Lecture No. 1. Ann Arbor: University of Michigan, 1967.

3. _____ and G. R. Ghez, The Allocation of Time and Goods over the Life Cycle. New York: Columbia University Press for NBER, forthcoming.

4. Y. Ben-Porath, "The Production of Human Capital and the Life-Cycle of Earnings." Journal of Political Economy 75: (August 1967): 353-365.

5. A. S. Blinder, "Wage Discrimination: Reduced Form and Structural Estimates." Journal of Human Resources 8:4 (Fall 1973): 436-455.

6. _____, Toward an Economic Theory of Income Distribution. Cambridge, Mass.: The MIT Press, 1974.

7. _____ and Y. Weiss, "Human Capital and Labor Supply: A Synthesis," Industrial Relations Section Working Paper #, Princeton University, 1974.

8. L. R. Christensen, D. W. Jorgenson and L. J. Lau, "Conjugate Quality and the Transcendental Logarithmic Production Function." Econometrica 39 (July 1971): 255-256.

9. W. J. Haley, "Human Capital: The Choice between Investment and Income." American Economic Review 63:5 (December 1973): 929-944.

10. J. Heckman and S. Polachek, "Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship." Journal of the American Statistical Association 69:346 (June 1974): 350-354.

11. B. G. Malkiel and J. A. Malkiel, "Male-Female Pay Differentials in Professional Employment." American Economic Review 63:4 (September 1973): 693-705.

12. J. Mincer, "Investments in Human Capital and Personal Income Distribution." Journal of Political Economy 66: (August 1958):

13. _____, Schooling, Experience and Earnings. New York: Columbia University Press for NBER, 1974.

14. _____ and S. Polachek, "Family Investments in Human Capital: Earnings of Women." Journal of Political Economy 82:2:II (March 1974): S576-S108.

15. M. R. Rosenzweig and J. Morgan, "Wage Discrimination: A Comment."
         Journal of Human Resources.

16. Y. Weiss, "On the Optimal Lifetime Pattern of Labour Supply."
         Economic Journal 82 (December 1972): 1293-1315.