

NBER WORKING PAPER SERIES

PROGRESS IN HUMAN CAPITAL ANALYSES  
OF THE DISTRIBUTION OF EARNINGS\*

Jacob Mincer

Working Paper No. 53

CENTER FOR ECONOMIC ANALYSIS OF HUMAN BEHAVIOR AND SOCIAL INSTITUTIONS  
National Bureau of Economic Research, Inc.  
204 Junipero Serra Blvd., Stanford, CA 94305

August 1974

Preliminary; Not for Quotation

NBER working papers are distributed informally and in limited number for comments only. They should not be quoted without written permission.

This report has not undergone the review accorded official NBER publications; in particular, it has not yet been submitted for approval by the Board of Directors.

\*This paper was presented at the Royal Economic Society Conference on the Personal Distribution of Incomes and Property, Lancaster, England, July 1974.

This research was funded by a grant to the NBER from the National Science Foundation (GS-31334) and by a contract from the U.S. Department of Labor (L-73-135) for research on the determinants of the distribution of income and earnings. The opinions expressed herein are those of the author and do not necessarily reflect the views of the National Science Foundation or the Department of Labor.

May, 1974

PROGRESS IN HUMAN CAPITAL ANALYSES  
OF THE DISTRIBUTION OF EARNINGS

1. Introduction

According to a popular adage economists study choice behavior, while sociologists explain why there are no choices to be made. In this light, the label of economics as a "dismal" science is surely misplaced. In the same light, the traditional studies of income distribution, a field with which economists are becoming increasingly concerned, must be described as basically sociological. That is, the traditional<sup>1</sup>

---

<sup>1</sup> Prior to the rehabilitation of Adam Smith.

---

approaches tend to stress differences in opportunity, ability, and chance as conditions largely unaffected by human choice.

The ascendancy of the human capital approach can be viewed as a reaction of economists to this non-economic, though certainly not irrelevant, tradition. In stressing the role played by individual and family optimizing decisions in human capital investments, important aspects of income determination are brought back within the mainstream of economic theory and within the power of its analytical and econometric tools.

Investment in human capital can take the form of expenditures on education, job training, health, information, and migration--to list

some of the major categories. Such expenditures of resources of time, money, and effort tend to augment an individual's earning capacity and thus can be viewed as investments, the augmentation of earnings being the return on them. Investment activities are undertaken by the individual and by his family within the constraints of genetic endowment, parental wealth, and access to educational and market opportunities.

Economics is the analysis of constrained choices. Whether the range and significance of these choices is "large" or "small" in the context of study of income distribution is a question amenable to research, not a matter to be left to ideological preconceptions. Nor is investment in human capital the only element of choice in the analysis of income distribution. Adam Smith, no stranger to this part of the world, listed a number of aspects of job choices which affect the distribution of labor incomes. These, he said, "are the principal circumstances which so far as I have been able to observe, make up for a small pecuniary gain in some employments, and counter balance a great one in others: first, the agreeableness or disagreeableness of the employments themselves; secondly, the easiness and cheapness, or the difficulty and expense of learning them; <sup>thirdly,</sup> the constancy and inconstancy of employment in them; fourthly, the small or great trust which must be reposed in those who exercise them; and fifthly, the probability or improbability of success in them."<sup>2</sup>

---

<sup>2</sup> Adam Smith [1937], p. 106.

---

Nonpecuniary aspects of wages, instability of employment, uncertainty of success, and problems of trust have been analyzed by

economists, in a rather fragmentary fashion.<sup>3</sup> Far more work needs to

---

<sup>3</sup> Cf. Friedman and Kuznets [1945], Friedman [1953], Weiss [1972], Becker and Stigler [1974].

---

be done on each of the topics suggested by Smith. The emphasis on human capital investments--his point number two--should not distract our attention from these aspects of work choices. Nevertheless, it appears that the subject of human capital investments lends itself to a more systematic and comprehensive analysis of wage differentials, than each of the other factors. Perhaps also the current prominence of the subject derives from a historical context: Students of economic growth were the first to recognize the importance of human capital in analyzing the modern evolution of industrial development.

The following is a description of research in the distribution of labor incomes in which human capital theory serves as an organizing principle. It is, in part, a sequel to my 1970 survey and, in part, a report of ongoing research of my own and of others. Again, the emphasis on human capital is not to be read as a denial of other aspects of choice listed by Adam Smith, or of the "sociological" factors, which are best viewed as constraints on choices, rather than as mutually exclusive hypotheses. Put differently, the research reported below does not inquire into all the forces and factors affecting the distribution of income. Far more modest, the question is: what is the role and impact of human capital investment decisions on the distribution and structure of earnings. Though the question is partial, the theoretical framework of the human capital approach is flexible. It is not a single, rigid

model but a way of thinking, capable of development in scope and complexity. The appropriate concept of income based on human capital is labor income, and the recipient unit the individual worker. Ultimately, an inclusion of non-employment incomes and aggregation of individual incomes into family or household incomes will be needed for the analysis of the distribution of total household incomes.

I stop short of such aggregation in this report. Human capital theory applies most directly to labor incomes. Since labor income is by far the major component of personal income, except perhaps at the far ends of the distribution, its analysis is the task of priority--particularly for labor economists.

## 2. Earnings Profiles

The basic conceptual and observational unit of human capital analysis is the lifetime earnings stream of the individual, not just his earnings during a limited, say annual, period of time. Earnings at any given time are viewed as a return on--a rental value of--the human capital stock, the "skill level" which the individual has accumulated. Since the size of the capital stock changes over the life cycle, growing by means of investment and declining because of depreciation and obsolescence, earnings change correspondingly over the life cycle. The characteristic age profile of earnings shows rapid growth during the first decade of working life, subsequent deceleration of growth and a leveling in the third and fourth decade. This is true when average earnings of "homogeneous" cohorts are studied over time, net of economy-wide growth trends and net of short-run fluctuations. Individual earnings profiles differ,

even within such groups, in height (level), rate of growth (slope), and the rate of change in the latter (curvature). In simplest terms, the personal or size distribution of earnings is viewed as a distribution of the earnings profiles of the individual members of the labor force.

Thus the distributional analysis starts from its micro-economic building block--the analysis of the individual income stream, the earnings profile. The parameters of the individual earnings curve, its level, slope and curvature, acquire specific economic interpretations in the light of human capital analysis. The analysis of earnings distribution then reduces to an analysis of the distribution of these parameters in the population.

The economics in this analysis is to be found in the process by which the individual earnings curve is generated. This process is analyzable as an optimizing decision of the individual (and his family) about the allocation of investments in his human capital stock over his life cycle. Such optimization models were pioneered by Ben Porath [1967] and by Becker [1967]. The models are undergoing continuous refinement, but their essence is brought out in these early formulations. Briefly, rational allocation requires that most of the investment in the person be concentrated at younger ages. The investments may increase before adolescence, but will continue at a diminishing rate throughout much of a person's working life. Investments are not incurred all at once in a short and early period, even though this would maximize the remaining payoff period and total returns. This is because marginal costs of producing human capital rise within the period. The solution is to stagger investments over time at an eventually diminishing rate--both

because benefits decline as the payoff period, the remaining working life, shortens and as opportunity costs of time, which is an input in the learning process, are likely to rise over the individual's working life.

This reasoning applies to gross investments in human capital. It also applies to net investments, provided the depreciation rate is fixed or a positive function of time over the life cycle.

Since earnings are proportional to the level of the human capital stock, they rise at an eventually diminishing rate and decline when net investments become negative, if at all, in old age. The typical working life earnings profile is therefore concave, at least in percent terms. Its average level is a positive function of total net investments added to the initial endowment. Its rate of growth at any time is a positive function of the net amount invested in the prior period, and the degree of concavity depends on how rapidly investments decline over time.

According to a popular alternative view, the individual earnings curve is basically an intrinsic age phenomenon: it reflects productivity changes due to inherent biological and psychological maturation, leveling off in the middle years and declining later because of declining physical and intellectual vigor. In the language of human capital, this view explains the earnings profile by the depreciation rate alone: the rate is negative in early years, zero in middle life, and positive in later years.

There is evidence, however, to indicate that this inherent age factor affects earnings only to a minor degree during the usual working

life. In data where age and work experience are statistically separable, the earnings curve is found to be mainly a function of experience, more than of age, in terms of both its location in the life cycle and the sizes and signs of its growth rates. Earnings profiles differ by occupation, sex, and other characteristics in systematic ways not attributable to the aging phenomenon.

Another interpretation of the shape of the earnings profile as a "learning curve" or a reflection of growth of abilities with age and experience known as "learning by doing" is not at all inconsistent with the human capital investment interpretation, provided it is agreed that opportunities for learning are not costless. That is, given differential learning options among jobs and no insuperable barriers to labor mobility, present values of earnings among the various learning options will tend toward equalization among workers with similar capacities. Thus labor mobility will impose opportunity costs of learning, by reducing initial earnings of the steeper profiles below the initial earnings of the flatter profiles. The relevant labor mobility applies, of course, to workers with similar qualifications, that is, the same level of human capital stock prior to entry into the labor market. This kind of investment in human capital via job mobility in the labor market is to be distinguished from job training, formal or informal, on a given job or "job ladder." But the analysis in human capital terms is the same. I, therefore, prefer the term "post-school investments," which encompasses both aspects of job investments, to the more narrow term which has come to be known as the "on-the-job training hypothesis."

I now proceed to an exposition of the earnings function which



summarizes in equation form the various categories of human capital investments as determinants of earnings profiles. Thus far, the categories have been broad, couched in life cycle intervals, such as schooling and post-school investments, and most recently pre-school or "home" investments. Future progress in the analysis of the earnings function lies in successive refinements of content in these categories. Progress in the broadest sense will require a development of structural relations to include determinants as well as consequences of human capital investments.

### 3. Earnings Functions

A brief development of an individual's earnings function is as follows:

Let  $C_{t-1}$  be the dollar amount of net investment in period (t-1) while "gross" earnings, that is, earnings from which the investment expenditures are not netted out, are  $E_{t-1}$ . Let  $r_{t-1}$  be the rate of return on this particular instalment of investment, and assume--for simplicity--that  $r$  is the same in each period.<sup>4</sup>

---

<sup>4</sup> As with many other simplifying assumptions, this one can be relaxed, given the purpose and the data.

---

Then:

$$E_t = E_{t-1} + rC_{t-1} \quad (1)$$

Progressive substitutions for  $E_{t-1}$  lead to:

$$(2) \quad E_t = E_0 + r \sum_{j=0}^{t-1} C_j, \quad \text{where } E_0 \text{ is the initial earning}$$

capacity, a person's earnings if no subsequent investments were made in him. If  $E_0$  originates at age 0, we can view it as the return on his

genetic endowment. If the starting point is later, it is a mixture of genetic and environmental influences. If the latter can be thought of as investment activities, for example by parents in preschool children, and separated out as such, it would be useful to include them in the second term of equation (2).

Clearly, data on the individual instalments of investment are not easily observable, except for formal schooling and training programs, which are only a part of the story. Even so, it is years of school attainment and not dollar costs for which data are abundant. For this reason alone, and for others to be mentioned later, it is preferable to express the right-hand variables in the earnings function in terms of "time spent in investment" rather than in dollar magnitudes. This is accomplished by viewing the ratio of investment expenditure to gross earnings as a time-equivalent amount of investment:

$$\text{Define } K_t = \frac{C_t}{E_t} \quad (3)$$

If  $t$  is a given year and  $K_t=20\%$ , this means that 20% of the year's gross earnings was spent in investment. If the costs of investments are only time costs, then  $K$  does, in fact, represent the fraction of the year spent in investment activities.

Substituting (3) in (1), we have

$$(4) \quad E_t = E_{t-1}(1+rK_{t-1}), \quad \text{and by recursion}$$

$$E_t = E_0 (1+rK_0) (1+rK_1) \dots (1+rK_{t-1})$$

With  $rK$  a relatively small number, a logarithmic approximation is

appropriate, and:

$$\ln E_t = \ln E_0 + r \sum_{j=0}^{t-1} K_j \quad (5)$$

Some investments are in the form of schooling, others take the form of pre-school care, job training, job mobility, medical care, acquisition of information, and so forth. At this stage of development of the earnings functions, the K-terms have been segregated into two categories, namely schooling and post-school investments.

Thus (5) can be written:

$$\ln E_t = \ln E_0 + r_s \sum_{i=0}^{s-1} K_i + r_p \sum_{j=0}^{t-1} K_j \quad (6)$$

where  $i$  runs over years of schooling,  $j$  over years of post-school experience.  $K_i$  are investment ratios during the school period, and  $K_j$  thereafter. The subscripts at  $r_s$  and  $r_p$  indicate that, in principle, the average rates of return on schooling may differ from the average rates of return on post-school investments.

Function (6) is specified in terms of net investment ratios (K). Net investments can be decomposed into gross investment and depreciation as follows: Let  $C_{t-1}^*$  be the dollar amount of gross investment in period  $t-1$ ,  $\delta_{t-1}$  the depreciation rate of the stock of human capital, hence of earnings  $E_{t-1}$  during that period, and  $K_t^* = \frac{C_t^*}{E_t}$ , the gross investment ratio.

$$\text{Then } E_t = E_{t-1} + r C_{t-1}^* - \delta_{t-1} E_{t-1}$$

$$\text{and } \frac{E_t}{E_{t-1}} = 1 + r K_{t-1}^* - \delta_{t-1} = 1 + r K_{t-1}, \text{ by equation (5).}$$

Therefore,  $rK_t = rK_t^* - \delta_t$ , and function (6) can be written:

$$\ln E_t = \ln E_0 + \sum_{i=0}^{s-1} (r_s K_i^* - \delta_i) + \sum_{j=0}^{t-1} (r_p K_j^* - \delta_j) \quad (7)$$

Earnings functions (6) or (7) must be adapted for empirical purposes in at least two respects.

First, the dependent variable  $E_t$  which I term "gross earnings" or "earnings capacity" is the earnings figure that would be observed if the individual stopped investing in himself in period  $t$ . Continued investment means, however, that "net" earnings ( $Y_t$ ) are smaller than  $E_t$  by the amount invested  $C_t$ . For practical purposes, I equate observed with "net earnings."<sup>5</sup>

---

<sup>5</sup> Note that observed earnings, as they are usually reported in statistical accounts, would equal "net" earnings if  $C_t$  consisted only of opportunity costs. Since direct expenditures are not usually "netted out," observed earnings overstates net earnings somewhat. Given the importance of opportunity costs in human capital investments, observed earnings more closely approximate the "net" than the "gross" concept.

---

Since  $Y_t = E_t (1 - K_t)$ , the earnings functions can be written:

$$(8) \quad \ln Y_t = \ln E_t + \ln (1 - K_t), \text{ substituting the appropriate expressions from (6) or (7) for } \ln E_t.$$

Next, the investment ratios  $K_t$  or  $K_t^*$  have to be given empirical content. In the schooling stage  $K_i^* = \frac{C_i^*}{E_i}$ , where  $C_i^*$  are: foregone earnings ( $E_i$ ), plus tuition and cost of living differential attributable to schooling,

minus student earnings and student aid. Without knowing the  $C_i^*$  for each individual, we know that  $K_i^*$  is not far from unity during school years, and this is a convenient approximation.<sup>6</sup> In the post-school stage, we

---

<sup>6</sup> It appears from 1960 U.S. data on college students that, on average, student earnings plus scholarship roughly paid for tuition. Even if true on average, this assumption is worth relaxing when data are available, as has been done in the recent work by Solmon [1972], Wachtel [1973], Johnson and Stafford [1973], and Leibowitz [1974]. The correction for quality requires relaxing the assumption that  $K_i = 1$  during school years. This requires expenditure data which differ among schools for the numerator of  $K_i$ . Of course, expenditure data do not fully capture quality, particularly in the public school system. Still, accounting for variation in expenditures among schools was significant in the empirical analyses in the references cited above.

---

have only the theoretical hypothesis that  $K_j^*$  declines after completion of schooling, when it was close to unity. Positive earnings upon entry into the labor force mean that  $K_j^* < 1$ , and since  $C_j^*$  eventually declines to zero, so must  $K_j^*$ . Note that a monotonic decline of  $K_j^*$  is not inconsistent with an initial constancy or even increase in  $C_j^*$ . This means that concavity of logarithmic earnings profiles is not inconsistent with initial linearity or even convexity of dollar earnings profiles.

My own experiments with specifying  $K$  or  $K^*$  as functions of time proved that the simplest linear specifications fit as well as other forms.

Recalling that  $K_i^* \approx 1$  and putting  $K_j^* = K_0^* + \frac{K_0^*}{T^*}t$ , when  $T^*$  is the length

of working life, we get:

$$(9P) \quad \ln Y_t = \ln E_0 + (r_s - \delta_s)s + (r_p K_0^* - \delta_p)t - \frac{r_p K_0^*}{2T^*} t^2 + \ln (1 - K_t^*)$$

Alternatively, in terms of net investment:

$$(10P) \quad \ln Y_t = \ln E_0 + r_s s + r_p K_0 t - \frac{r_p K_0}{2T} + \ln (1 - K_t)$$

Here  $T$  is the period of positive net investment, so that  $T < T^*$ . Thus the peak of earning capacity is reached some time before the end of working life when depreciation nullifies or outstrips gross investment.<sup>7</sup>

---

<sup>7</sup> Putting  $\frac{d \ln E_t}{dt} = 0$ , timing of peak earning capacity is  $t_p = (1 - \frac{\delta}{rK_0^*})T^*$ .

As gross investment continues to decline, the peak of observed earnings will appear approximately  $\frac{1}{r}$  years later. [Mincer, 1974, p. 21].

---

An alternative specification which I used is a geometric decline in the investment profile:  $K_j^* = K_0^* e^{-\beta j}$  which leads to:

$$(9G) \quad \ln Y_t = \ln E_0 + (r_s - \delta_s)s + \delta_p t + r_p \frac{K_0^*}{\beta} (1 - e^{-\beta t}) + \ln (1 - K_t^*)$$

or, in net investment terms:

$$(10G) \quad \ln Y = \ln E_0 + r_s s + r_p \frac{K_0}{\beta} (1 - e^{-\beta t}) + \ln (1 - K_t)$$

Here,  $\beta$  is the annual percent rate of decline in the investment ratio  $K$ , and the function (9 or 10G) is a Gompertz function, well known in studies of industrial growth.

To the extent that hours of work vary over the life cycle, the

profile of annual earnings is affected. Since capacity wage rates first grow and later decline before retirement, they are likely to induce a corresponding pattern of hours of work supplied to the market. This assumes that over the life cycle the substitution effect of changes in earning power due to human capital investments dominates the wealth effect in the labor-leisure choice for a given individual. Incidentally, hours of work are likely to peak before observed wage rates do because (as noted in footnote 7 above) capacity wage rates decline before observed wage rates do, given human capital depreciation.

While the life cycle profile of hours of work can be explained, in part, by human capital investments, much of it is exogenous or transitory. Since annual earnings are a product of wage rates and hours,  $\ln Y_t = \ln W_t + \ln H_t$ , a logarithmic hours variable should be attached to the earnings function for standardizing purposes. Short-run variation in time worked per year is not uncommon for a given individual, and it is much greater across individuals.<sup>8</sup>

---

<sup>8</sup> By ignoring individual experience during the working life--our variable  $t$ , and the actual amount of working time during the year ( $H$ ), Jencks [1972] was left with a huge, unexplained variance in his analyses of income distribution, after experimentation with a large number of evidently less important variables.

---

#### 4. Empirical Analysis of Male Earnings, U.S. 1960

With a quadratic approximation for the last term  $\ln(1-K_t)$  in the earnings functions (9) and (10), the following simplest statistical

estimating equation relates accumulated human capital to earnings at each point in the working life:

$$(11P) \quad \ln Y_t = b_0 + b_1 s + b_2 t + b_3 t^2 + v$$

$$\text{where } b_0 = \ln E_0 - K_0 \left(1 + \frac{K_0}{2}\right); \quad b_1 = r_s$$

$$b_2 = r_p K_0 + \frac{K_0}{T} (1 + K_0); \quad b_3 = - \left[ \frac{r_p K_0}{2T} + \frac{K_0^2}{2T^2} \right]$$

Standardizing for hours worked during the year:

$$(12P) \quad \ln Y_t = b_0 + b_1 s + b_2 t + b_3 t^2 + b_4 \ln H + U$$

$b_4 = 1$ , if hours and hourly wage rates are uncorrelated. If  $b_4 \neq 1$ , the coefficients of the other variables in (11) will differ from those in (10a).

Similarly, for the Gompertz function, with gross investment ratios:

$$(11G) \quad \ln Y_t = b_0 + b_1 s + b_2 t + b_3 X_t + b_4 X_t^2 + v$$

$$\text{where } X_t = e^{-\beta t} \quad b_2 = -\delta$$

$$b_0 = \ln E_0 + \frac{r_p K_0^*}{\beta} \quad b_3 = - \frac{r_p K_0^*}{\beta} - K_0^*$$

$$b_1 = r_s \quad b_4 = - \frac{K_0^{*2}}{2}$$

Standardization for hours worked during the year is obtained in an equation (12G) by adding the term  $b_5 \ln H$ , as in (12P).

In my just-published study, equations (11) and (12) were applied



to data of the 1/1,000 sample of the 1960 U.S. Census. The multiple regressions were run on individual earnings of over 30,000 white, urban males, nonstudents of pre-retirement age, who had some earnings in 1959. The results are shown in Table 1.

---

Table 1

---

Findings of regression analysis

1. The regressions shown in Table 1 perforce assume the same  $K_t$ -profile for each individual (more precisely the same product  $rK_t$ ). Despite the relegation of the unobservable individual differences in  $r$  and in  $K_t$  into the residual variance, the two variables  $s$  and  $t$  alone explained about 30% of total inequality.

Among the other findings, the following are noteworthy, keeping in mind the human capital interpretation of estimated parameters:

2. A negative coefficient for  $s^2$  suggests, as was found by others, that the rate of return to schooling diminishes at higher levels of schooling. However, the significance of the  $s^2$  term vanishes once the employment variable ( $W$  = weeks worked) is added to the regression. Thus, it appears that when earnings are measured in wage rates, there is no decline in rates of return at higher schooling levels. Therefore the differences in employment during the year almost fully account for the higher rates of return at the lower levels of schooling when annual earnings are compared.

3. The negative coefficient of the interaction term ( $st$ ) shows an apparent convergence of (logarithmic) experience profiles in the cross-

TABLE 1

REGRESSIONS OF INDIVIDUAL EARNINGS ON SCHOOLING (s),  
EXPERIENCE (x), AND WEEKS WORKED (W)  
(1959 annual earnings of white, nonfarm men)

	Equation Forms	R <sup>2</sup>
S(1)	In Y = 7.58 + .070s (43.8)	.067
P(1)	In Y = 6.20 + .107s + .081t - .0012t <sup>2</sup> (72.3) (75.5) (-55.8)	.285
P(2)	In Y = 4.87 + .255s - .0029s <sup>2</sup> - .0043ts + .148t - .0018t <sup>2</sup> (23.4) (-7.1) (-31.8) (63.7) (-66.2)	.309
P(3)	In Y = f(D <sub>s,t</sub> ) + .068t - .0009t <sup>2</sup> + 1.207 ln W (13.1) (10.5) (119.7)	.525
G(1a)	In Y = 7.43 + .110s - 1.651x <sub>it</sub> (77.6) (-102.3)	.313
G(1b)	In Y = 7.52 + .113s - 1.521x <sub>it</sub> (74.3) (-101.4)	.307
G(2a)	In Y = 7.43 + .108s - 1.172x <sub>it</sub> - .324x <sub>it</sub> <sup>2</sup> + 1.183 ln W (65.4) (-16.8) (-10.2) (105.4)	.546
G(2b)	In Y = 7.50 + .111s - 1.291x <sub>it</sub> - .162x <sub>it</sub> <sup>2</sup> + 1.174 ln W (65.0) (-3.5) (-16.0) (107.3)	.551
G(3)	In Y = f(D <sub>s,t</sub> ) + 1.142 ln W (108.1)	.557
G(4)	In Y = 7.53 + .109s - 1.192x <sub>it</sub> - .146x <sub>it</sub> <sup>2</sup> - .012t + 1.155 ln W (-2.4)	.556

NOTE: Figures in parentheses are t ratios. R<sup>2</sup> = coefficient of determination; S = linear form; P = parabolic form; G = Gompertz form; D<sub>s,t</sub> = dummies for schooling and experience; x<sub>it</sub> = e<sup>-1.5t</sup>, x<sub>it</sub> = e<sup>-1.0t</sup>; W = weeks worked during 1959.

section. Again, this term becomes insignificant when weeks worked are included in the regression.

4. Tentative estimates of  $r_p$ ,  $K_0$  and  $\delta$  are somewhat more secure using the (G) functions:  $r_p$  is about 12%,  $K_0$  is 40-50%, and  $\delta$ , the depreciation rate, 1.2% per year. The  $K_0$  coefficient is disturbingly high. It may confound a maturation phenomenon (negative value of  $\delta$  at early ages) or is due to some other misspecification.

One data problem which affects these estimates is the absence of direct information on the start of work experience. The average age of completion of schooling was taken as an estimate of  $t_0$ . This defect is remediable with appropriate data.

5. Adding variation in weeks worked ( $\ln W$ ) to the equations raises their explanatory power to 55%. The coefficient at ( $\ln W$ ) is significantly larger than unity, suggesting a positive correlation between weeks worked and weekly earnings within schooling and experience cells.

No doubt, one can maintain that the variables schooling ( $s$ ), years of work experience ( $t$ ), and hours of work during the year ( $H$ ) are rather obvious determinants of earnings, not requiring any analytical structures such as human capital theory. Though I have no objections to empirical fishing expeditions--indeed, they are quite useful--the analytical structure provides guidance both for specification of variables, equations, and equation forms and, most important, it provides interpretation--a set of insights into the earnings structure linked together by a story, albeit tentative and partial. Its viability becomes apparent in the way the story "hangs together" and in the way the beginning of

it leads to fuller developments in a disciplined fashion.

In this brief discussion of the findings I want to emphasize the features contributed by the model--in distinction to ad-hoc analyses which utilize similar variables among many others. I will not apologize for the simplicity of a single equation with at most three substantive and imperfectly measured variables: a first step must precede all the others. I will defer to the later pages the description of the steps which follow.

The contribution of the human capital model to the empirical analysis can be seen in a number of features:

1. The earnings function expresses the earnings profile as an individual growth curve. The Gompertz curve, for example, is a familiar empirical representation of industrial growth. That it fits a personal growth curve is not a coincidence, since the staggered investment interpretation is suitable in both cases.

2. The coefficients of the function represent estimates of (average) rates of return and of volumes of investment in schooling and after schooling. It appears from calculations based on the estimates that the rates of return and the total dollar volumes of investment are of similar magnitude in both categories.

3. The experience variable ( $t$  or  $X$ ) is a proxy for the investment profile ( $K_t$ ). It is measured in years of labor market experience to represent cumulated investments in job training and in job mobility. Ad-hoc analyses of earnings single out age. As already noted, the distinction between age and experience is important: it helps explain earnings profiles of workers who differ in levels of education.

More educated workers of the same age have less work experience, since they enter the labor market later. Since growth rates of earnings reflect investment rates and these are a (negative) function of experience, growth rates of earnings are stronger for the more than for the less educated workers at given ages. This is the essence of the universally observed and econometrically bothersome age-schooling interaction effect. With experience as an explicit variable, the interaction effects on earnings vanish.

A more dramatic example of the difference between age and experience appears in the analysis of earnings of women, whose labor market experience is frequently interrupted. Age is an especially poor substitute for experience in that case. For the analysis of earnings of workers whose work experience is discontinuous the human capital earnings function is well suited, provided work history data are available. This analysis is described in a later section.

4. The form of the earnings function depends on the units in which the independent variables are expressed. If dollar earnings are of interest, schooling and post-school investment variables must be expressed in dollar volumes--a difficult task. The more readily available time variables require a semi-logarithmic equation form, that is, earnings must be expressed in logarithms, while schooling and experience enter arithmetically. The additional advantages of the semi-log formulation is that it eliminates the interaction between  $s$  and  $t$ , and it provides analyses of relative inequality, which is of greater interest than absolute dispersion.

The explanatory power of the semi-log form when schooling and

experience are expressed in years has been shown to be superior to the arithmetical or double-log specification of the same variables. [Heckman and Polachek, 1974].

The role of schooling and the concept of "overtaking"

Simple correlations between earnings and years of schooling are quite weak. Moreover, in multiple regressions when variables correlated with schooling are added, the regression coefficient of schooling is very small. This leads to a view, which is regaining currency, that schooling matters very little, insofar as earnings are concerned. (Ironically, an opposite view, sometimes held by the same people, is that schooling may matter in earnings but has little to do with learning.) The human capital approach suggests that such conclusion is too hasty.

In the 1960 U.S. data previously referred to, the simple coefficient of determination between log-earnings and years of schooling for the whole sample was merely 7%. Standardizing for effects of age doubles the coefficient in age groups 35-44, but the coefficient weakens in younger and older groups. The human capital framework suggests that the proper standardization is by years of experience, but even then we find that the correlation differs a great deal depending on which years-of-experience groups we consider. What stage of experience is the

---

Table 2

---

most appropriate for observing effects of schooling, least contaminated by other factors? The answer of the human capital model is: during the first year of experience, if no further investments in human capital were

TABLE 2

CORRELATION OF LOG EARNINGS WITH SCHOOLING  
WITHIN EXPERIENCE OR AGE GROUPS

Years of Experience	Coeff. of Det. ( $r^2$ )			Coeff. of Det. ( $r^2$ ) (3)
	All <sup>a</sup> (1)	Year-round (2)	Years of Age	
1-3	.31	.25		
4-6	.30	.27	20-24	.02
7-9	.33	.30	25-29	.04
10-12	.26	.30	30-34	.11
13-15	.20	.25	35-39	.14
16-18	.17	.20	40-44	.16
19-21	.16	.18	45-49	.12
22-24	.13	.17	50-54	.12
25-27	.13	.15	55-60	.09
28-30	.12	.14	60-64	.08
31-33	.07	.14		
34-36	.05	.07		
37-39	.07	.09		
Aggregate	.07	.08	Aggregate	.07

SOURCE: 1/1,000 sample of the U.S. Census, 1960.

a. All workers, including both year-round and those whose work was part time, seasonal, or otherwise intermittent.

undertaken beyond schooling. In that case, by equation (6) and with the approximation  $K_1 = 1$ , indeed

$$\ln E_s = \ln E_0 + r_s s \quad (13)$$

would be the best specification. The existence of post-school investments, however, makes  $E_s$  unobservable. Instead we observe  $Y_s = E_s - C_{s0}$  when  $C_{s0}$  is the amount invested in the first year of work experience. Earnings are initially smaller than  $E_s$ , but as they rise with experience they eventually reach the level  $E_s$ . Given a rate of return on human capital investments equal to  $r$ ,  $E_s$  is the level of a horizontal earnings flow whose present value is equal to the present value of the actual earnings stream  $Y_t$ , both discounted at the rate  $r$  at the beginning of working life. Since the coefficients of the earnings function provide estimates of  $r$ , these intersection ("overtaking") points were found to locate at a little less than a decade of experience.<sup>9</sup> The overtaking period, of course,

---

<sup>9</sup> It is easily shown a priori that the "overtaking" year of experience  $\hat{t} \leq \frac{1}{r}$ , if post-school investments do not increase over time:

$$Y_{\hat{t}} = E_s + r \sum_{j=0}^{\hat{t}-1} C_j - C_{\hat{t}}$$

For  $Y_{\hat{t}} = E_s$ , the remaining terms on the right must be zero.

Hence  $r \hat{t} \bar{C} = C_{\hat{t}}$ ,  $\hat{t} = \frac{1}{r} \cdot \frac{C_{\hat{t}}}{\bar{C}}$ , where  $\bar{C}$  is the average amount

invested per year over the first  $\hat{t}$  years of experience.

Thus  $\hat{t} \leq \frac{1}{r}$ .

---



differs among individuals depending on their initial post-school earning capacities  $E_s$ , rates of return, and their post-school investment profile. Using average profiles of schooling groups, a rough central tendency was located within the 7-9 year interval.

To repeat, theoretically, the correlation between initial gross earnings and schooling would clearly bring out the effect of schooling, and would decay with each successive year of experience, unless post-school investments were perfectly correlated with schooling investments. The distribution of initial gross earnings is not observable, but it is roughly approximated by the distribution of observed earnings at the "overtaking stage" of experience.

At that stage the coefficient of determination represents an estimate of the fraction of earnings inequality that is attributable to differences in schooling. It should be higher than the correlation with initially observed earnings, if the exact "overtaking" point could be found for each individual, and should decay thereafter. This pattern is observed in Table 2, and the coefficient of determination at "overtaking" exceeds 30%.

The inequality of earnings at the overtaking stage amounts to about 70% of aggregate inequality in the 1960 sample. If the remaining 30% are largely attributable to individual differences in post-school investments (including depreciation), and a third of the inequality at "overtaking" is due to schooling differentials, together a half of total inequality of observed earnings can be attributed to the distributions of schooling and post-school investments. The 50% figure is probably

an understatement. It fails to reflect differences in investments in quality of schooling,<sup>10</sup> and it does not take account of the differences

---

<sup>10</sup> There is a growing literature on this subject which far exceeds the bounds of this paper. See note 6 for human capital approaches. My rough estimate based on this work is that the inclusion of schooling quality would raise the explanatory power of human capital by over 5%.

---

in hours worked during the year which are induced by differences in human capital investment. The true residual, which is less than half of observed inequality, is due to individual differences in rates of return to human capital investments, transitory variation in employment during the year, and a portmanteau of everything else which we may call "chance."

#### The earnings structure

There are several prominent features of the statistical distribution of earnings (and income) which are repeatedly observed in temporally and regionally differing data. Aggregate skewness and the growth of inequality with age are the best known. Shapes of distributions cross-classified by schooling and experience are less familiar, and perhaps less stable. These characteristic features of earnings distributions have puzzled observers since detailed statistical data became available. Partial explanations, largely of the "random shock" variety, have been proposed.

In the human capital model, most of these features can be explained by the correlation between the stock of human capital at any stage of the life cycle and the volume of subsequent investment. This

correlation is understandable, if factors of ability and of opportunity which affect individual investment behavior tend to persist over lengthy periods of a person's life. For example, the absolute growth of dollar earnings with experience is greater at higher schooling levels.

Since the slope of the earnings profile at time  $t$  reflects investment in the prior period, this relation is an example of the persistence of levels of investment in schooling and afterwards.

Several other implications of the positive correlation between successive instalments of investment in human capital in dollar terms can be observed. Dollar profiles of earnings "fan out" with experience and, a fortiori, with age, both across and within schooling groups. Dollar variances in these groups, therefore, increase with experience and with age. Similarly, because the dispersion of dollar schooling costs increases with the level of schooling, variances of earnings increase with level of schooling. Since mean earnings increase with experience and with schooling, there is a positive correlation between means and variances in age and schooling subgroups of the earnings distribution. This correlation contributes to the appearance of positive skewness in the aggregate earnings distribution. This factor is independent of, and in a way more basic than, the shape of the distribution of schooling, which in the past also contributed to the positive skewness in earnings.

According to the relation  $\ln E_s = \ln E_0 + rs$ , earnings (at "overtaking") tend to be positively skewed even if the distribution

of years of schooling  $s$  is symmetric.<sup>11</sup> Distribution of years of schooling

---

<sup>11</sup> This is a sufficient condition. For a precise formulation see Mincer [1974, p. ].

---

tends to be positively skewed when the average level of schooling is low and to become symmetric at higher levels. In the U.S. skewness in the distribution of schooling has turned negative in the younger cohorts. So the shape of the distribution of schooling is no longer an important factor in explaining the persistence of positive skewness in the distribution of earnings in the U.S. Since the level of schooling in the U.S. is among the highest, aggregate skewness in earnings in most countries follows a fortiori.

If we define relative skill differentials by percent differentials in wage rates among schooling groups having comparable years of experience, we find that these are almost invariant over the working life. Since the logarithmic experience profiles of wages are concave, this finding implies that relative wage differentials among schooling groups increase with age. However, within schooling groups, relative wage dispersions, measured by variances of logs, show somewhat different profiles, depending on the level of schooling. When plotted against age, all are U-shaped along at least some portion of the curve, and clearly so at the center of the schooling distribution, that is, for the high-school group. For the post-high-school group, the profile is mainly increasing. Within lower schooling groups, it first decreases and then levels off.

Both the wage differentials between schooling levels and the

inequality patterns within the middle levels of schooling reflect a negligible correlation between post-school learning capacity and time-equivalent post-school investment. This same lack of correlation underlies the invariance between experience and relative wage differentials among schooling groups. The phenomenon arises if experience profiles of post-school investments, in time-equivalent units, are not systematically different among schooling groups. Put another way, it arises when the elasticity of post-school investments (in dollars) with respect to post-school earning capacity is, on average, unitary across schooling groups. Within schooling groups, however, the elasticity of investment with respect to earning capacity appears to increase with schooling level: it is less than 1 at lower levels and greater than 1 at higher levels.

The size of the elasticities and the systematic positive relation between schooling level and elasticity of investment with respect to earning capacity raise questions for further research. In this connection, it is noteworthy and suggestive that very similar patterns are found in studying the consumption function: The "long-run" elasticity of saving with respect to income is not clearly different from 1, and the "short-run" or cross-sectional elasticity increases with schooling level [Solmon, 1972].

##### 5. Human Capital Versus Stochastic Models

In stochastic theories of income distribution  $\epsilon_1$  is interpreted as year-to-year individual fluctuation in earnings and the whole structure of earnings is explained by a stochastic process that is attributed to

this "random shock"  $\epsilon_i$ . These models specify that:

$$v_i = \ln Y_i = \ln Y_0 + \sum_{j=1}^t \epsilon_j, \quad (14)$$

where the  $\epsilon_j$  are homoscedastic and mutually independent. This leads to a monotonically increasing log variance as a function of  $t$  (age or experience), and a positively skewed aggregate distribution (log-normal or Pareto, depending on differences in assumptions). But, as we have seen, the prediction that logarithmic variances of income grow monotonically and equally in all skill (schooling) groups is largely incorrect.

The greater and richer explanatory power of the human capital model need not preclude some validity in the random shock approach. Moreover, some of the predictions are similar: log variances of earnings do grow in some schooling groups and over certain phases of the working life. Even so, the same empirical phenomena are differently interpreted in the two models. In the stochastic models temporal variation in income is interpreted as chance variation. In contrast, in human capital models, much of the temporal variation in earnings is viewed as a systematic and persistent consequence of cumulative investment behavior. Discrimination between the two views can be sought in so-called panel correlations of earnings of the same cohort in two different time periods.

If we follow the earnings experience of a cohort  $m$  years after the initial year  $t$ , the random shock model implies that: (1) log variances will increase by the same amount  $\sigma^2(\epsilon)$  each year, so that:

$$\sigma^2(\ln Y_{t+m}) = \sigma^2(\ln Y_t) + m\sigma^2(\epsilon) \quad ; \quad (15)$$

and (2) panel correlations, that is, correlations between  $\ln Y_t$  and  $\ln Y_{t+m}$ ,

will decay continuously as the interval  $m$  is widened:

$$R^2 (\ln Y_t, \ln Y_{t+m}) = \frac{\sigma^2(\ln Y_t)}{\sigma^2(\ln Y_{t+m})}, \quad (16)$$

and

$$\frac{1}{R^2} = 1 + m \left[ \frac{\sigma^2(\epsilon)}{\sigma^2(\ln Y_t)} \right]. \quad (17)$$

According to the random shock model, both variances and the reciprocals of the coefficients of determination should increase linearly with the time interval  $m$ . We have already seen a contradiction in that the profiles of variances are not linear. If it could be assumed that the profiles are linear, the steeper slope at the higher schooling level implies a greater importance of random shock there, that is, a larger  $\sigma^2(\epsilon)$ , hence a more rapid decay of panel correlations in the higher schooling groups (since  $\sigma^2(\epsilon)/\sigma^2(\ln Y_t)$  would be larger at higher schooling levels). Again, this implication is not substantiated in Table 3, which is based on a 1959 survey of the Consumers Union Panel<sup>12</sup> and contains

---

<sup>12</sup> There were 4,191 usable responses in the recall data. Over half of the respondents were college graduates. For a detailed description of the data, see Juster [1964].

---

panel correlations ( $R^2$ ) and their inverses ( $1/R^2$ ). Data on past earnings from which the correlations were calculated are based on recall of respondents. Recall data probably contain a great deal of error, which may affect the level and pattern of the coefficients of determination. In an attempt to minimize this error, correlations of earnings at  $t$  and  $t+m$

TABLE 3

## PANEL CORRELATIONS OF MALE EARNINGS, BASED ON CONSUMERS UNION PANEL, 1959 SURVEY

Initial Year (t)	Years of Schooling														
	12 or Less			13-15			16			17 or More			All		
	2	7	11	2	7	11	2	7	11	2	7	11	2	7	11
	<b>Coefficients of Determination (R<sup>2</sup>)</b>														
4	.989	.227	.312	.911	.444	.518	.854	.302	.376	.803	.441	.316	.822	.388	.430
7	.951	.220	.268	.852	.324	.265	.691	.383	.381	.760	.430	.388	.752	.426	.348
9	.711	.491	.279	.800	.396	.483	.712	.598	.527	.800	.461	.381	.785	.503	.453
12	.837	.654	.498	.907	.648	.616	.889	.581	.552	.897	.528	.679	.878	.578	.586
15	.846	.520	.412	.816	.684	.507	.932	.538	.615	.873	.555	.608	.824	.630	.608
18	.818	.588	.399	.898	.604	.591	.918	.652	.662	.887	.652	.739	.898	.681	.714
21	.899	.483	.498	.839	.699	.643	.874	.771	.755	.925	.771	.596	.871	.716	.658
24	.828	.419	.403	.931	.764	.688	.966	.768	.715	.908	.868	.637	.930	.788	.648
27	.902	.682	.744	.935	.801	.860	.955	.765	.757	.982	.794	.419	.952	.793	.781
Average	.864	.476	.423	.876	.596	.574	.865	.595	.593	.870	.611	.529	.856	.611	.580
	<b>Reciprocals of R<sup>2</sup></b>														
Average of															
t = 4, 7	1.031	4.475	3.468	1.135	2.669	2.852	1.308	2.960	2.641	1.280	2.301	2.872	1.272	2.462	2.598
All	1.165	2.451	2.588	1.143	1.834	1.914	1.170	1.845	1.896	1.155	1.744	2.051	1.172	1.735	1.833
t = 12	1.170	1.845	2.129	1.129	1.440	1.576	1.085	1.503	1.498	1.097	1.505	1.683	1.128	1.451	1.515

NOTE: Earnings at t years of experience are correlated with earnings at t + m years of experience; t + m is in 1959 for each of the cohorts; m = 2, 7, or 11, as indicated in the column headings.



years of experience were observed only in those cohorts whose experience did not exceed  $t+m$ . Thus, only rows in Table 3 pertain to given cohorts. Years of experience were provided by respondents as time elapsed since they first entered full-time employment.

Despite the unpredictable effects of errors in such data, there are two features in the table that are noteworthy: (1) As the interval  $m$  is widened from two to seven years, the correlation declines sharply when the panel base  $t$  is in the first decade of experience. The decline is much milder thereafter. (2) When the interval  $m$  is widened further, from seven to eleven years, the decline in correlation, if any, is negligible. The growth in  $1/R^2$  is not linear, particularly over the earlier decades of experience. These findings are clearly inconsistent with the random shock model. They do seem reasonable in the light of the human capital model: panel correlations bracketing the overtaking stage would be expected to be relatively weak, but stronger thereafter. When the interval brackets the overtaking point, we are correlating

$$\ln E_s + \sum_{j=0}^{t-1} (rK_j - K_t)$$

with

$$\ln E_s + \sum_{j=1}^{t+m-1} (rK_j - K_{t+m})$$

By definition, the post-school investment component of earnings is negative before overtaking and positive thereafter. The bracketing, therefore, introduces a negative correlation between the investment components of earnings, which weakens the panel correlation. Indeed, if

all  $E_s$  were equal this correlation would be negative.<sup>13</sup> The sharp decel-

---

<sup>13</sup> John Hause [1974] attempted to test the implied negative partial correlation of the human capital model in two longitudinal samples. The data were those collected by A. Husén in Sweden and by D. C. Rogers in the U.S. The correlation was erratic in the first set, for which a very short time span (4-5 years) was used. It was negative in the second set.

Paul Taubman [1974] uses the NBER-TII sample and uses a simple instead of a partial correlation. The simple correlation need not be and is not negative when earning capacity at "overtaking" is not fixed.

---

eration or even halt in the decline of correlations beyond a seven-year span is not implausible: beyond "overtaking," the ranking of individual earnings acquires a long-run stability, though disturbed by short-run, "transitory" fluctuations.

In contrast to the purely stochastic models of income distribution, which are not capable of covering much ground, two sophisticated studies of earnings combine earnings functions which have much in common with the human capital approach, with rigorous stochastic specifications of the residual variation:

M. M. Fase [1970] studied a sample of Dutch incomes. In his model an individual starts his career at the age of  $s$  years. The annual rate of increase in the salary is assumed to decline linearly with age, starting at  $s$ . In addition, a random disturbance drawn from a log-normal distribution contributes to the annual change in the logarithm of income. Fase acknowledges the human capital interpretation of his model, though it did not originally motivate his work.

In A. Klevmarcken's [1972] study of Swedish earnings data, the human capital concepts are well recognized. His treatment of the stochastic component is less restrictive than in Fase. One of the conceptual improvements is the distinction between "physical" and "active" age<sup>14</sup>--

---

<sup>14</sup> Separation of the two effects was attempted more intensively in a more recent study of Klevmarcken and Quigley [1973].

---

that is, between age and labor market experience. Another, econometric improvement is the pooling of cross-section and cohort data.

## 6. The Occupational Wage Structure

Occupational "skill differentials" in wages are commonly measured by the percentage difference between adult male wage rates in sets of pairs of narrowly defined occupations. The choice of pairs, the definition of wages, and the changing skill contents make the interpretation of such comparisons and of trends in them as trends in relative factor prices rather uncertain. The often steep rise of earnings with age (experience) suggests that the differing age distribution among occupations is another source of ambiguity in these measures.

An acceleration of upward trends in schooling raises the average age in the lower schooling and skill groups and lowers it in the upper groups. This produces an apparent narrowing of relative wage differentials, which may be misinterpreted as a relative price change, plausibly resulting from changes in relative supplies. But the apparent compression of the wage differential may simply be an artifact.

According to my analysis, even standardization for age is generally insufficient. The appropriate analysis must take into account the occupational experience profiles as well as the schooling component of occupational skill, at the very least. An interesting analysis of this sort was carried out by C. M. Rahm [1971] who fit earnings function to mean earnings in over 500 detailed occupations listed in the 1960 U.S. Census. His results are shown in Table 4 below.

Table 4

Mean Earnings Regressions across DetailedOccupational Groups of Males

U.S. Census, 1959

(Standard errors in parentheses)

	<u>R<sup>2</sup></u>
(1) $\ln Y = 7.005 + .128s$	.29
(.009)	
(2) $\ln Y = 5.112 + .151s + .164t - .0031t^2$	.58
(.007) (.013) (.0003)	
(3) $\ln Y = .491 + .106s + .094t - .0017t^2 + 1.513 \ln W$	.71
(.007) (.012) (.0003) (.097)	

Rahm thus explained 60-70% of the inter-occupational relative wage differentials with these few variables. The effects of schooling and of experience across occupations was similar to the estimates for individuals in our Table 1. In large part, then, occupation can be

viewed as a composite of skills acquired in schooling and on the job.

Rahm ran similar regressions on 116 less detailed occupations which were comparable in the Census years 1940, 1950, and 1960. The results were similar, over the years, though the  $R^2$  were inflated by aggregation,<sup>15</sup> and so were the coefficients of experience  $t$ . The only

---

<sup>15</sup> The  $R^2$  for 1959 were .83 in (2) and .87 in (3).

---

coefficient which changed since 1939 was that on schooling. It dropped from 14% in 1940 to 10% in 1950 and 1960. The well-known decrease in earnings inequality which occurred between the first two periods is evidently associated, in part, with this decline in the rate of return to schooling, since the variance in schooling did not change much (it narrowed slightly).<sup>16</sup>

---

<sup>16</sup> The other part is due to the narrowing of the variance in the weeks-worked variable. Cf. Mincer and Chiswick [1972].

---

Rahm's analysis assumes the same effects of schooling and of experience across the various occupations, so the parameter estimates are average effects. An analysis in which these effects are allowed to differ would be desirable for a number of purposes, not the least of which is an insight into differential post-school job skill investments.

## 7. Regional Differences and Temporal Changes in Income Distributions

By taking variances of both sides of the earnings function (10), relative income inequality can be expressed as a function of means, var-

iances and covariances in schooling, experience, and employment, and of parameters such as the average rates of return estimated as coefficients in the earnings function. Using data on adult males in the various states of the U.S., Chiswick [1974] estimated earnings functions within the regions and related the differences in the means, variances, correlations, and rates of return to differences in aggregate inequality between regions. With this variance formulation of the earnings function Chiswick was able to explain 85-90% of the inter-regional differentials in relative income (or earnings) inequality. The most important explanatory variables were the (average) rate of return on human capital (education) and variances in the distribution of employment (weeks worked), in the age, and in the schooling distribution.

Basically the same approach was taken by Mincer and Chiswick [1972] in a time-series analysis of annual changes in income inequality in the U.S. between 1939 and 1969. For males in the 25-64 age group there were no perceptible net trends in inequality between 1949 and 1969. 87% of the annual variation in inequality during this period was explained by changes in the distributions of employment, age, and schooling. The rate of return was assumed fixed in the annual data. On the other hand, the strong decline in inequality between 1939 and 1949 was attributable, in large part, to the decline in the variance of employment--a correlate of the decline in the level of unemployment--, and to a decline in the rate of return to schooling, observed in other data.<sup>17</sup>

---

<sup>17</sup>Cf. reference to Rahm [1971] in the preceding section.

---

A sensitivity analysis of the variance form of the earnings

function suggests that even large changes (in standard deviation units) in the distribution of schooling and age have minor effects on annual earnings inequality and that change in the distribution of employment, which is a cyclical variable, has stronger effects. The strongest effects are produced by changes in rates of return.

Going back in time prior to 1939, the narrowing of income inequality over the first half of this century is consistent with apparent declines in rates of return, as fragmentary evidence on both phenomena indicates. Growth of income, insofar as it leads to consumption-motivated growth in the demand for education, tends to depress rates of return and, thereby, to narrow inequality. At the same time, however, the growth in incomes is a result of growth in market productivity which probably generates growing demands for skills in the labor market. In the two decades prior to 1970 such growth must have been strong to keep the rates of return and inequality rather stable.

Though, apart from its effects on rates of return, the secular growth of education has minor effects on inequality, it has several distinguishable effects on the observed earnings structure: (1) Growth of schooling appears to be associated with a decline in the dispersion and in the positive skewness of the distribution of schooling. This is largely due to a natural (zero) or legislated (positive) lower limit on years of schooling. The decreased dispersion in the distribution of schooling in 1960 may also be a lagged effect of the narrowing inequality of parental income that was observed before 1950. The distribution of earnings within age groups in current data reflects the effects of a mild secular narrowing in the dispersion of schooling and of a stronger

reduction in its skewness, to the point where it is now negative.

(2) Acceleration of schooling trends produces additional reductions in inequality: the meaning of the upward trends in education is that the level of education is higher in young than in old age groups. This offsets, in part, the age variation in earnings, which is due to the growth of experience with age. Another consequence is that the relative numerical importance of the young and least educated and old and most educated groups becomes smaller the more rapid the upward educational trends. But these are precisely the groups within which the inequality in earnings is largest [Mincer, 1974]. Therefore, the stronger the upward trend in schooling, the smaller the aggregate inequality in earnings. It can be shown [Mincer, 1974] that, if growth in schooling ceased and the distribution of schooling in each age group remained the same as among young earners with less than a decade of work experience in 1960, aggregate earnings inequality in the U.S. would become 10% larger than it was in 1960. (3) Secular trends in education also affect the distribution of income indirectly via effects on the composition of the labor force and the resulting distribution of employment. The lengthening of schooling and increased enrollment produced a growing intermittent student labor force. The growth of education of women contributed to a growing female labor force which is also frequently intermittent. Growth of part-period and part-time work widens the dispersion of employment, which tends to widen the inequality in annual earnings. Thus, when all earners (men, women, and teenagers) are included, inequality in annual earnings has indeed widened in the U.S. in the past two decades. This is not true, however, in full-time or hourly earnings of men, nor



is it true when inequality is measured across family units rather than across persons.<sup>18</sup>

---

<sup>18</sup> Indeed, the inequality of income among families has had a slight downward trend, related to the growing labor force participation of married women. For a similar finding in Britain, see H. Lydall <sup>[1959]</sup> ~~[1970]~~. The analysis of this phenomenon is outside the scope of this paper. It involves the consideration of labor supply and human capital decisions within the family context.

---

#### 8. Earnings of Women

In the post-school stage of the life cycle much of the accumulation of earning power takes place in the labor market. The experience variable in the earnings equation acts as a representation of the sequence of job investments (ratios) of labor force participants. Where past work experience of men can be measured without much error in number of years elapsed since leaving school, such a measure of "potential work experience" is clearly inadequate<sup>19</sup> for workers whose labor force exper-

---

<sup>19</sup> Such a measure was used in earnings functions of women by R. Oaxaca [1973].

---

ience is discontinuous. Therefore, a basic requirement for the analysis of earnings of women is direct information on their work histories.

The 1967 National Longitudinal Survey of Work Experience<sup>20</sup>

---

<sup>20</sup> For a description of the NLS Survey, see Shea, Spitz, and Zeller [1970].

---

carried out by the U.S. Department of Labor contains work histories and other characteristics of a large sample of women who were between 30 and 44 years old in 1967. The heretofore unavailable opportunity afforded by the NLS, albeit on a retrospective basis, was exploited by Mincer and Polachek in a recent study [1974].

According to the data, less than 50% of the mothers worked in 1966, but close to 90% worked some time since they left school, and two-thirds returned to the labor market some time after the birth of the first child, but not necessarily permanently. In contrast, women without husbands and without children spent close to 90% of the years in continuous labor market activities. In terms of chronology, the life cycle of married women features several stages which differ in the nature and degree of labor market and home involvement. There is usually continuous market work prior to the birth of the first child. The second stage is a period of nonparticipation related to child bearing and child care, lasting 5-10 years, followed by intermittent participation before the youngest child reaches school age. The third stage is a more permanent return to the labor force for some, though it may remain intermittent for others. In the data which were obtained from women who were less than 45 years old, only the beginning of the third stage was visible.

A number of implications about human capital investment behavior of women are plausible, given expectations of such lifetime work patterns by the average woman. Of course, the expectations and the patterns are changing, and women's human capital investment behavior can be expected to change correspondingly. But, leaving aside trends,<sup>21</sup> we hypothesize that:

---

<sup>21</sup> In my view societal expectations and rhetoric follow rather than precede

the changing facts. The basic fact in the trends is the rise in real market wages. If these rose more rapidly than productivity in the household, as seems plausible, the upward trends in labor force participation of women and downward trends in fertility, and related changes in the family, become intelligible. Cf. Mincer [1962] and [1963].

---

1. The smaller the expected lifetime participation in the labor market the less the investment (or vocational) aspects of women's formal education, and the less the acquisition of job training at work compared to men with comparable education.

2. During the period of child-bearing and child care, prolonged nonparticipation may cause the skills acquired at school and at work to depreciate.

3. There is likely to be a stronger expectation of prospective continuity of employment after the children reach school age. At this stage women who return to the labor market may have strong incentives to resume investments in job-related skills.

4. These conjectures imply that the investment profile of married women is not monotonic, as was hypothesized for men. There is a gap which is likely to show negative values (net depreciation) during the child-bearing period. Labor market investments of never-married women are likely to exceed those of married women, even initially, assuming lesser expectations of marriage. At the same time they are smaller than investments of men, since some of the never-married women had expectations of marriage. The continuity of their work experience suggests a declining investment profile, as in the case of men.

5. The implications for earnings profiles are clear: Earnings profiles of men are steepest and concave, of childless women less so, and of mothers double-peaked with least overall growth.

These implications are consistent with the empirical findings in which we obtained parameter estimates of earnings functions especially adapted for the purpose of the analysis. A brief exposition of these "segmented" earnings functions is appropriate, as their usefulness extends beyond the particular application.

In order to adapt the earnings function to persons with intermittent work experience we break up the post-school investment term ~~into~~ ~~successive~~ into successive segments of participation and nonparticipation as they occur chronologically. In the general case with  $n$  segments we may express the investment ratio:

$$K_i = a_i + b_i t, \quad i = 1, 2, \dots, n$$

$$\text{and } \ln E_t = \ln E_0 + rs + r \sum_{i=1}^n \int_{t_i}^{t_{i+1}} (a_i + b_i t) dt. \quad (18)$$

Here  $a_i$  is the initial investment ratio,  $b_i$  is the rate of change of the investment ratio during the  $i^{\text{th}}$  segment

$$(t_{i+1} - t_i) = e_i = \text{duration of the } i^{\text{th}} \text{ segment.}$$

Note that in (18) the initial investment ratio refers to its projected value at  $t_1 = 0$ , the start of working life. In a work interval  $m$  which occurs in later life there is likely to be less investment than in an earlier interval  $j$ , though more than would be observed if  $j$  continued at its gradient through the years covered by  $m$ . In this case,  $a_m$  in equation (18) will exceed  $a_j$ .

Alternatively,  $a_j$  and  $a_m$  can be compared directly in the formulation:

$$\ln E_t = \ln E_0 + rs + r \sum_{i=1}^n \int_0^{e_i} (a_i + b_i t) dt \quad (19)$$

since  $a_i$  is the investment ratio at the beginning of the particular segment  $i$ .

While the rate of change in investment  $b_i$  is likely to be negative in longer intervals, it may not be significant in shorter ones. Since the segments we observe in the histories of women before age 45 are relatively

short, a simplified scheme is to assume a constant rate of net investment throughout a given segment, though differing among segments. The earnings function simplifies to:

$$\ln E_t = \ln E_0 + rs + r \sum_i a_i e_i \quad (20)$$

Where  $(ra_i) > 0$  denote positive net investment (ratios), while  $(ra_i) < 0$  represent net depreciation rates, likely in periods of nonparticipation.

The question whether the annual investment or depreciation rates vary with the length of the interval is ultimately an empirical one. Even if each woman were to invest diminishing amounts over a segment of work experience, those women who stay longer in the labor market are likely to invest more per unit of time, so that  $a_i$  is likely to be a positive function of the length of the interval in the cross-section.

Thus, even if  $K_{ij} = a_{ij} - b_{ij}t$  for a given woman  $j$ , if  $a_{ij} = a_j + \beta_j t$  across women, on substitution, the coefficient  $b$  of  $t$  may become negligible or even positive in the cross-section. On integrating, and using three segments of working life as an example, earnings functions (18), (19), (20) become:

$$(18 \text{ a}) \quad \ln E_t = a_0 + rs + r[a_1 t_1 + \frac{1}{2} b_1 t_1^2 + a_2 (t_2 - t_1) + \frac{1}{2} b_2 (t_2^2 - t_1^2) + a_3 (t - t_2) + \frac{1}{2} b_3 (t^2 - t_2^2)]$$

$$(19 \text{ a}) \quad \ln E_t = a_0 + rs + r[a_1 e_1 + \frac{1}{2} b_1 e_1^2 + a_2 e_2 + \frac{1}{2} b_2 e_2^2 + a_3 e_3 + \frac{1}{2} b_3 e_3^2]$$

$$(20 \text{ a}) \quad \ln E_t = a_0 + rs + r[a_1 e_1 + a_2 e_2 + a_3 e_3]$$

In this example,  $t$  is within the last (third) segment, and the middle segment,  $e_2 = h$ , is a period of nonparticipation or "home time." The signs of  $b_i$  are ambiguous in the cross-section, as already indicated; the coefficients of  $e_1$  and of  $e_3$  are expected to be positive, but of  $e_2$  (or  $h$ ) negative, most clearly in (20a).

The equations for observed earnings ( $\ln Y_t$ ) differ from the equations shown above by a term  $\ln(1-K_t)$ . ~~as was shown in the comparison of equations~~  
~~(20a) and (20b)~~ With  $K_t$  relatively small, only the intercept  $a_0$  is affected, so the same form holds for  $\ln Y_t$  as for  $\ln E_t$ .

It will help our understanding of the estimates of depreciation rates to express earnings function (9a) in terms of gross investment rates and depreciation rates:

$$\begin{aligned} \ln E_t &= \ln E_0 + \sum_i (r K_i^* - \delta_i) = \\ &= \ln E_0 + (rs - \delta_s) + (r K_1^* - \delta_1) e_1 + (r K_h^* - \delta_h) h \\ &\quad + (r K_3^* - \delta_3) e_3 \end{aligned} \tag{20b}$$

This formulation suggests that depreciation of earning power may occur not only in periods of nonparticipation ( $h$ ), but at other times as well. On the other hand, market-oriented investment, such as informal study and job search, may take place during "home time," so that  $K_h^* > 0$ . Positive coefficients of  $e_1$  and  $e_3$  would reflect positive net investment, while a negative coefficient of  $h$  is an estimate of net depreciation. If  $K_h^* > 0$ , the absolute value of the depreciation rate  $\delta_h$  is underestimated.

Findings:(1) Investment and Earnings Profiles

Life histories of women who worked in 1966 were segmented into five intervals: three of these were periods of work experience and two were of non-market activity. According to equation ( ) the coefficients attached to these intervals ( $e_i$ ) represent estimates of gross investment ratios minus depreciation.

Investment magnitudes implied by these coefficients were lowest for married women with children, higher for women without children, and highest for never-married women. They were lower for women who worked less than half of their post-school years than for those who worked more.

The relations with level of schooling appeared also to be positive, though less clearly.

The investment profile of never-married women was declining, as indicated by a negative coefficient of the experience variable in the earnings function. On the other hand, mothers over age 35 who have returned to the labor market showed higher coefficients for the current than for the prematernal interval. Presumably, their current market work is expected to last longer than the previous periods of work experience.

The coefficients for the two periods of non-participation were negative, indicating a net depreciation rate amounting to 1.5% per year, on average, and increasing with educational level. This was pronounced <sup>the</sup> for <sup>^</sup> coefficient attached to the uninterrupted interval of non-participation lasting several years, which followed the birth of the first child. The length of these "home time" intervals was related to numbers of children.



<sup>t<sub>10</sub></sup>  
 Induction of numbers of children as an additional variable in the earnings function therefore had no significant (negative) effect, except in the small subgroup of highly educated women.

(2) Inequality in Earnings of Women

Judging by  $R^2$  in Table 5, the earnings function is capable of explaining 25-30 per cent of the relative (logarithmic) dispersion in wage rates of white married women and about 40% of the inequality in the rather small sample of wage rates of single women in the 30-44 age group, who worked in 1966. The earnings function is thus no less useful in understanding the structure of women's wages than it is in the analysis of wages of males.

The dispersion of hours worked during the survey year is much greater among married women,  $\sigma^2(\ln H) = .75$ , than among men,  $\sigma^2(\ln H) = .11$ . The (relative) dispersion in annual earnings of women is, therefore, dominated by the dispersion of hours worked. This factor is also important in the inequality of annual earnings of single women and of men of comparable ages, but much less so. It is not surprising, therefore, that the inclusion of hours worked in the earnings function, raises the coefficient of determination from 28% in the hourly wage equation to 78% in the annual earnings equation of married women, from 41% to 65% for single women, and from 32% to 50% for men.

The lesser inequality in the wage rate structure of working married women than in the structure of male wages is probably due to lesser average, and hence lesser variation in, job investments among individuals. At the same time the huge variation in hours, reflecting intermittancy and part-

TABLE 5

Earnings Inequality and Explanatory Power of Wage Functions

(White Married Women, Single Women, and Married Men, 1966)

		$\sigma^2(\ln W)$	$R_w^2$	$\sigma^2(\ln Y)$	$R_Y^2$	$\sigma^2(\ln H)$	n
All Married Women		.22	.28	.97	.78	.75	1,140
By education	S < 12	.17	.21	.81	.76	.64	435
	12-15	.18	.17	.92	.78	.74	622
	+16	.17	.16	.77	.74	.60	83
Single Women		.30	.41	.62	.66	.32	138
Married Men		.32	.30	.43	.50	.11	3,230

$\sigma^2(\ln W)$  = variance of (log) wages

$\sigma^2(\ln Y)$  = variance of (log) annual earnings

$\sigma^2(\ln H)$  = variance of (log) annual hours of work

$R_w^2$  = coefficient of determination in wage rate function

$R_Y^2$  = coefficient of determination in annual earnings function

time work as forms of labor supply adjustments creates an annual earnings inequality among women which exceeds that of men. However, the meaning of that inequality both in a causal and in a welfare sense must be seen in the family context. As was shown elsewhere,<sup>22</sup> the inclusion of female

---

<sup>22</sup> J. Mincer [1974].

---

earnings as a component of family income narrows the relative inequality of family incomes compared to that of incomes of male family earners.

(3) The Sex Differential in Wages

A comparison of earnings functions of women with functions of men of the same age (30-44) shown in Table 5 permits a rough accounting for the relative wage differential between men and women which was close to 40% between husbands and wives, and over 10% between men and single women. The pronounced differences in levels of the independent variables were in work experience. These differences alone (assuming the same coefficients for men and women) accounted for close to a half of the wage gap. Indeed, for a subgroup of women whose attachment to the labor force was continuous the wage gap was reduced by nearly 60%. Whether the remainder is mainly a matter of discrimination or of lower investment rates--as we interpreted the coefficients--will remain debatable. The evidence described above does not support a view that the association of coefficients with prospective work experience is fortuitous.<sup>23</sup>

---

<sup>23</sup> Cf. studies of earnings of professional women based on NSF data, Johnson and Stafford [1973].

---

TABLE 6

Experience and Depreciation Coefficients

White Married Women, Single Women, and Married Men, Age 30-44, 1966

	<u>Variables</u>				
	<u>S</u>	$\hat{e}$	$h_c$	$h_o$	$e_3$
Married Women					
Regression Coefficients	.063	.012	-.015	-.006	.009
Means	11.3	9.6	6.7	3.5	3.2
Single Women					
	<u>S</u>	<u>e</u>	$e^2$		$e_3$
Regression Coefficients	.077	.026	-.0006		.009
Means	12.5	15.6	258		8.0
Married Men					
	<u>S</u>	<u>e</u>	$e^2$		
Regression Coefficients	.071	.034	-.0006		
Means	11.6	19.4	409		

S = years of schooling

$h_c$  = "home time" following birth of first child

$h_o$  = other "home time."

$e_3$  = current job tenure

$\hat{e}$  = 2SLS estimate of total work experience

Sources: Women, NLS, 1967.  
Men, SEO, 1967.

9. The Segmented Earnings Function and the Analysis<sup>of</sup> Job Mobility

The experience term in the earnings function can be segmented in more than one way. The segmentation into periods of market and non-market activity was useful in analyzing earnings of women. Another useful segmentation is into a sequence of intervals of job tenure. This is a novel application of the earnings function for the study of labor mobility. In this framework the gain from labor mobility, to the extent that the latter is voluntary,<sup>24</sup> can be seen as a package containing both immediate

---

<sup>24</sup>This can be distinguished in the data, which report whether job change followed a quit or a lay-off.

---

gains in wages as well as job-investment options. Investment rates, producing rates of change in earnings can be inferred from coefficients of the job segments. Returns to costs of (geographic) mobility can also be distinguished, with appropriate data.

Several sets of longitudinal data, including the previously referred to NLS, the NBER-TH, and the Coleman-Rossi sample are currently analyzed in this fashion by my associates at the National Bureau. The NLS and the NBER-TH are panels with some retrospective information, mainly on work histories, while the Coleman-Rossi sample is fully retrospective, both in work experience and in earnings.

The analysis of work histories has now been applied to expand the earnings functions of men aged 45-59 in the NLS data. Preliminary results indicate relations between labor mobility, job stability, and earnings.

Longer duration of job tenure is associated with larger investments on the job. This suggests that worker self-investments - which the earnings function is supposed to reflect in its coefficients - are positively related to employer investment in workers. The longer duration of job tenure is likely to be an outcome or corollary of such joint investments which are firm specific,<sup>25</sup> according to the theory formulated by Becker [1964].

---

<sup>25</sup>The role of specific training in earnings functions was analyzed in an excellent study of Japanese data by M. Kuratani [1972]. It was also explored in a study of differences between occupational earnings of men and women by L. Landes [1974].

---

Though job mobility apparently enhances earnings at younger ages, it is associated with lesser earning and growth at older ages. Part of the explanation may be the differential mix of quits and layoffs in the two age groups.

The specific training hypothesis creates certain biases in the coefficients of the segmented function, upward on the current job coefficient and downward on the preceding ones. Research on the estimation of these biases is currently proceeding at NBER.

The explanatory power of earnings functions which take account of work histories is significantly greater (close to 10% in the NLS data) than that of the function which utilizes undifferentiated total experience in addition to schooling.

10. Background Variables: Ability and Opportunity.

(a) Ability and Screening

The repeatedly observed positive correlation between educational attainment, socioeconomic background, and assorted measures of ability poses an obvious question: Would more educated people earn more in any case, or is the added income really a product of the schooling process? This question has motivated a great deal of research for the purpose of estimating biases in the schooling coefficients when these variables are, presumably improperly omitted. These efforts have taken the form of adding various measures of ability (such as IQ) and of family socioeconomic status (measured in a variety of ways) to the earnings function. The variables themselves, of course, are not very reliable measures and their inclusion in a single equation is theoretically inappropriate. At any rate, in these studies the inclusion of measured ability reduces the coefficient of schooling from 5 to 35 percent, depending on the data and the measures used.<sup>26</sup>

---

<sup>26</sup>cf. Gintis [1971], Griliches and Mason [1972], Hause [1972], Taubman and Wales [1974].

---

Aside from the proliferation of types of ability measures (from IQ to AFQT) there is a problem with the age at which the measure was taken. As is well known, these measures grow over time with age and with the early growth of human capital (until late adolescence). In a notable study, Griliches and Mason [1972] estimated that the coefficient of schooling is reduced by 7-10%, if the correction allows for ability prior to schooling, while post-school measured ability was 4-13 additional percentile points higher for each additional year of schooling. Thus, if post-school "ability" measures are used, the downward bias in the

schooling coefficient is exaggerated by almost 100%.

The role of ability in affecting earnings has been interpreted as indirect and direct. Indirectly, it is an input to the production of human capital, either as an initial stock or as an efficiency parameter. Moreover, it affects earnings directly: the more capable graduates of the same schools earn more. Thus when the ability variable is entered together with schooling in the earnings function, the coefficient of schooling is reduced, but both coefficients remain significant.

According to an alternative hypothesis, the productivity effects on earnings derive mainly from ability, not from schooling. So ability is what really matters, and the indirection is only apparent, because schooling serves merely as a screen conveying the information about relevant abilities, or other desirable characteristics of job applicants, to employers. In principle, the productivity and screening functions of schooling are not mutually exclusive in a world of imperfect information, given that ability is an <sup>input</sup>~~aspect~~ in the educational process. The controversy, if any, is on the relative importance of the productivity and screening functions of schooling in affecting earnings. Unless screening is a deliberate device for monopolization its effect on earnings can neither be major nor durable, once productivity of the worker can be directly observed. Moreover, the characteristics for which schooling serves as a screen could be discovered by means of direct interviewing and testing much more cheaply than by expenditures of many years and tens of thousands of dollars on an average education.



Markets for testing would surely spring up if such tremendous savings were possible.

Even the advocates of screening as the major function of schooling do not maintain that the screen is permanent. If it were, the correlation of schooling with earnings would be fixed at all levels of work experience. This is obviously false, as our Table (2) shows. If the sorting effect is temporary, the correlation should be strongest at the outset of work experience and decay progressively, and perhaps quite rapidly thereafter. A test of discrimination from the human capital model is possible here, because the correlation implied by it is somewhat different. The human capital model suggests that the correlation between gross earnings and schooling will decline with experience, because of individual differences in post-school job investments. Investments are larger in earliest years, so the correlation between schooling and observed earnings (= gross earnings minus investment) is biased downward, but the downward bias decreases with experience. At the "overtaking" point, observed earnings approximate initial gross earnings, so the correlation of schooling with observed earnings would be strongest at that point, if it could be precisely determined for each individual. Despite the imprecision in empirically determining individual "overtaking" points, our Table 2 shows that the correlation between schooling and earnings does not start to decay before the first decade of experience.

(b) Opportunity, Family Background, and "Home Investments"

In considering the opportunity or socio-economic background variables it is more difficult, than in the case of ability, to visualize direct effects

on earnings,<sup>27</sup> unless what is meant is racial discrimination, class collusion, or nepotism. The indirect effects may run via genetic inheritance, which would show in ability traits already discussed, or in the quality of the early environment including purposive investment by parents in the early human capital stock of their children.

---

<sup>27</sup>Direct effects in consumption and on property income via gifts and bequests are, of course, observable.

---

Most of the studies which include family background variables in the earnings equations report small effects, net of the human capital variables such as schooling and experience. On the other hand, the background variables, which usually include parental occupation, education, and numbers of siblings, are shown to be significant predictors of the child's educational attainment.

The Griliches-Mason work [1972] previously noted shows the importance of the indirect effect in another way: Their measure of schooling is partitioned into school completed prior to serving in the armed forces and subsequent schooling. By controlling for background variables, the early schooling coefficient is reduced about 25 percent, but the post-armed service schooling effect shows a negligible reduction. The impression that the link between background and earnings is largely the effect of background on the informal and formal learning environment is further supported by a number of studies which have shown that the partial correlation between son's earnings and mother's education is higher than the partial correlation between son's earnings and

father's schooling especially if parental income is controlled for.<sup>28</sup>

---

<sup>28</sup>The coefficients are not very different without such control. Evidently father's education is more strongly correlated with family income than is mother's education. See Hunt [1963], Hill and Stafford [1974], Coleman [1974], Leibowitz [1974], Parsons [1974].

---

The major importance of indirect compared to direct effects of family background shifts the focus of attention to parental efforts toward accumulation of the human capital stock of their children. Much of this accumulation takes place in the home, particularly during the preschool stage of the life cycle, as well as later. It appears that the education of parents is a significant variable even after controlling for family income and numbers of siblings. This suggests that aside from money expenditures on schooling, the quality and quantity of time parents spend with children may be viewed as inputs in the child quality (human capital). The time inputs are mainly those of the mothers who take the major child care responsibilities, and who reduce their market activities to engage in them. The reduction in <sup>their</sup> earnings which results from this reduction of time otherwise spent in the labor market is a direct measure of the opportunity cost of these investments. Estimates of these costs are feasible, given data on women's wages, and their child-care activities in the home.

An illustrative calculation of such opportunity investment costs of child care was performed on the 1966 NLS data on earnings and work histories of women.<sup>29</sup> On average, each additional child caused over

---

<sup>29</sup>Mincer and Polacheck [1974], Table 9.

---

2 years of interruption of work in the labor market. In 1966 dollars the opportunity costs per child ranged from \$7,000 for mothers with less than high school education to \$17,000 for mothers with college education or more. In money terms, the difference in investment in pre-school children among the higher and lower education group was equivalent to a difference of 2-3 years of schooling.

Time cost of the mother is only a partial measure of parental investments in preschool children. In principle, measures of such investments might be incorporated in the earnings function framework. From such functions we could tell more clearly whether these preschool investments have an independent effect on earnings, beyond affecting school attainment of the child. In any case, the greater earnings - and presumably also greater consumption capacities of children which are associated with the higher level and quality of education they attain, may be viewed in the family context as part of the return on the education of mothers.

We can visualize the early production function of a child's human capital as containing three inputs: the genetic endowment of the child, parental contributions of market goods, and of their own time. Thus far, research by economists has been confined to the estimation and valuation of parental time inputs. We are only at the beginning of research efforts into (1) the nature and scope of parental efforts, given genetic and economic constraints, (2) the productivity of these efforts in adding to the human capital stock of children, and (3) the relative importance of parental contributions in the ultimate level of the capital stock achieved by the children.

Evidence on (2), that is on the relation between parental time inputs and measures of child development, achievement, and earnings is scant. Some evidence was found by A. Leibowitz [1974] in the very special Terman Sample in a simplified recursive scheme. Briefly, she found that (a) parental time inputs as well as education of the mother affected the child's IQ measure, (b) once IQ and both parental educations are taken into account the time input measures have no further effect on educational attainment of the child, and (c) once education and experience of the adult son or daughter is taken into account, the parental variables are of little consequence in affecting earnings.

The findings are still very fragile in this emerging economic analysis of the role of the family in the formation of economic capacities of children. Its potential payoff is clearly promising. It lays the groundwork for a deeper exploration of income distribution both among families and across generations.

(c) Toward a Fuller Specification of an Equilibrium System

The positive association between earnings, human capital variables, ability, and socio-economic status raises questions of cause and effect which still need to be resolved. The system can be viewed in two stages: (1) Ability and opportunity factors affecting the accumulation of human capital, and (2) The effect of human capital on earnings. The earnings functions represent the second, or proximate stage of the system. In terms of ultimate determinants, the background variables (stage 1) must be brought into the analysis. It is clear, however, that the procedure of putting such variables along with human capital variables in the same earnings equation is not correct from a structural point of view. If investment in human capital results not from random behavior but from optimizing behavior, the estimates of "independent" effects of each of

the variables are biased if all are included in the earnings function.

According to Becker's optimization analysis [1967], the earnings function results from two simultaneous structural relations in the human capital market. These are demand functions ( $D_i$ ) which relate individual investments to marginal rates of return and supply functions ( $S_i$ ) which relate the volume of funds that can be obtained for human capital investments to their marginal costs. Of course, worker demand for self-investment ( $D_i$ ) is, in part, derived from employer demand for the worker's human capital. Optimizing behavior implies that the volume of individual investments, the magnitude <sup>of</sup> marginal and average returns, and therefore the volume <sup>of</sup> earnings are simultaneously determined by the intersection of demand and supply curves.

The biases resulting from a disregard of the optimizing aspects of behavior can be illustrated as follows: For a fixed level of human capital there must be a perfect negative correlation between ability and opportunity: persons with greater ability invested no more than others only because their opportunities were inferior, and conversely. Thus, adding both sets of background variables alongside with measures of human capital volumes is meaningless. If only one set of such variables is added, the explanatory power of the equation will be increased, but the "independent" effects of the background variables are still misestimated. In principle, the appropriate statistical procedure is a simultaneous equations model that could "identify" the opportunities and capacities functions, including the effects on both functions of background and human capital accumulation. Experimentation with empirically feasible specifications is a task of great urgency. It is a major subject of our current research efforts, as is the analysis of the recursive structure of human capital investment paths which lead

from family background and parental efforts to the next generation's lifetime earnings.

## II. Retrospect and Prospect

The human capital analysis reported in this review represents initial attempts to broaden the scope of earnings functions beyond the exclusive attention formerly paid to school education. The econometric models used are still deliberately parsimonious. Further expansion of variables and equations are steps to be guided by current and future theoretical development and empirical experimentation.

The simple forms described in the paper provided insight into the distribution of earnings among individuals, males as well as females, occupational wage differentials, effects of job mobility, regional differences, and temporal changes in income distributions. The power of the analysis was increased by the refinement ("segmentation") of the post-school investment category. Of course, we need to remember that it is not time spent in the labor market, but the volume of investment activity taking place during that time which determines earnings. Analyses of individual differences in post-school investment behavior, will require richer panel data, in addition to job histories.

Moving to an earlier stage of the life cycle, consideration was given to the notion and promise of analyzing parental investment in children, particularly preschoolers. Additionally, the initial (genetic?) levels of the human capital stock, subsequent investments in health, and the life cycle of human capital depreciation, including the important

problem of obsolescence, all deserve special analytical and empirical attention.

The distinction between annual earnings and wage rates has not been sufficiently emphasized in the work here reviewed. The proper analytical distinction requires a marriage of labor supply and human capital theory, which is also needed in moving from personal to family distributions of income.

The single earnings equation is basically a reduced form. There is a need to estimate the structural relations involving demand and supply in the market for human capital investment funds and in the labor market. Background factors of ability and opportunity which determine investments in human capital will need to be specified in a systematic fashion. Another and related structure that should be elucidated is the recursive chain leading from family background to lifetime earnings of the next generation.

Work has started on each of these topics at different stages of sophistication. It is encouraging to find that researchers in various countries have found the human capital approach, mapped out here, quite useful for analyzing their own societies and economies.<sup>30</sup>

---

<sup>30</sup>References are to completed studies of Fase [1970], Klevmarken [1972], Kuratani [1973], and Levy-Garboua [1973] who analyzed Dutch, Swedish, Japanese, and French data respectively.



## REFERENCES

- Becker, G. S., Human Capital, Columbia Press, 1964.
- \_\_\_\_\_. "Human Capital and Personal Distribution of Income," Woytinsky Lecture, University of Michigan, 1967.
- Becker, G. S., and Stigler, G. J., "Law Enforcement, Malfeasance, and Compensation of Enforcers," The Journal of Legal Studies, University of Chicago Law School, January, 1974.
- Ben-Porath, Y., "The Production of Human Capital and the Life-Cycle of Earnings," Journal of Political Economy, August, 1967.
- Chiswick, B. R., Regional Analyser of Income Inequality, NBER, 1974.
- Chiswick, B. R., and Mincer, J. "Changes in Personal Income Inequality in the U.S. Since 1939," Journal of Political Economy, Part II, June, 1972.
- Coleman, J., "Processes of Change in Occupation and Income," forthcoming, 1974.
- Fase, M. M., An Econometric Model of Age-Income Profile, Rotterdam University Press, 1970.
- Friedman, M., "Choice, Chance, and the Personal Distribution of Income," Journal of Political Economy, August, 1953.
- Friedman, M., and Kuznets, S., Incomes in Independent Professional Practice, NBER, 1945.
- Gintis, H., "Education Technology and the Characteristics of Worker Productivity," American Economic Review, May, 1971.
- Griliches, Z., and Mason, W., "Education, Income, and Ability," Journal of Political Economy, May/June, 1972.
- Hause, J. C., "Earnings Profiles: Ability and Schooling," Journal of Political Economy, May/June, 1972.
- \_\_\_\_\_. "The Covariance Structure of Earnings," NBER Working Paper, 1973.
- Heckman, J., and Polachek, S., "The Functional Form of the Income-Schooling Relationship," NBER Working Paper, 1972.
- Hill, R., and Stafford F., "Family Background and Lifetime Earnings," Conference on Research in Income and Wealth, May, 1974.

- Hill, R., and Stafford, F., "Allocation of Time to Preschool Children," Journal of Human Resources, 1974.
- Hunt, S. J., "Income Determinants for College Graduates," Yale Economic Essays, Fall, 1963.
- Jencks, C. et al., Inequality, New York: Basic Books, 1972.
- Johnson, G., and Stafford, F., "Social Returns to Quantity and Quality of Schooling," Journal of Human Resources, Spring, 1973.
- Klevmarcken, A., Statistical Methods for the Analysis of Earnings Data, Stockholm, 1973.
- Kuratani, M., "The Distribution of Earnings in Japan," Ph.D. Thesis, Columbia University, 1972.
- Landes, L., "Male-Female Occupational Wage Differentials," Ph.D. Thesis, Columbia University, 1974.
- Leibowitz, A., "Production in the Household," American Economic Review, May, 1974.
- \_\_\_\_\_. "Home Investments in Children," Journal of Political Economy, Part II, May, 1974.
- Levy-Garboua, L. "Rémunère-t-on Les Études?" Consommation, No. 3, 1973.
- Lydall, H., "The Long Term Trends in the Size Distribution of Income," Journal of the Royal Statistical Society, Series A, Part 1, 1959.
- Mincer, J., "Labor Force Participation of Married Women," in Aspects of Labor Economics, NBER, 1962.
- \_\_\_\_\_. "Market Prices, Opportunity Costs, and Income Effects," in Measurement in Economics, Stanford University Press, 1963.
- \_\_\_\_\_. "The Distribution of Labor Incomes: A Survey," Journal of Economic Literature, March, 1970.
- \_\_\_\_\_. Schooling, Experience, and Earnings, Columbia University Press, 1974.
- Mincer, J., and Polachek, S., "Earnings of Women," Journal of Political Economy, Part II, March/April, 1974.
- Oaxaca, R., "Sex Discrimination in Wages," in Discrimination in Labor Markets, Princeton, 1973.
- Parsons, D., "Intergenerational Wealth Transfers," Working Paper, Ohio State University, March, 1974.

- Rahm, C. M., "The Occupational Wage Structure," Ph.D. Thesis, Columbia, 1971.
- Shea, J. R., Spitz, R. S., and Zeller, F. A., Dual Careers, Columbus, Ohio, 1970.
- Smith, A., The Wealth of Nations, Modern Library, N.Y., 1937 edition.
- Solmon, L., "Education and Saving Behavior," NBER Working Paper, 1972.
- Solmon, L., and Wachtel, P., "Effects of Schooling Quality on Earnings," NBER Working Paper, 1972.
- Taubman, P., and Wales, T., Higher Education: An Investment and a Screening Device, NBER, forthcoming, 1974.
- Taubman, P., "Schooling, Ability, . . . , and Lifetime Earnings," Conference on Research in Income and Wealth, May, 1974.
- Weiss, Y., "Investment in Graduate Education," American Economic Review, December, 1971.