NBER WORKING PAPER SERIES

HUMAN CAPITAL AND LABOR SUPPLY: A SYNTHESIS

Alan S. Blinder*

Yoram Weiss*

Working Paper No. 67

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

January 1975

*Alan S. Blinder, Princeton University
Yoram Weiss, NBER and Tel Aviv University

The research reported herein was partially supported by funds from the National Bureau of Economic Research.

I. Introduction

It is by now widely recognized that investment decisions play a major role in the determination of individual age-earnings profiles. Several studies, most notably Mincer [1974], rely on investment in human capital as a leading <u>single</u> hypothesis from which many of the observed regularities in earnings data are derived.

The investment hypothesis, in its simplest form, can be summarized by the equation:1/

$$Y = F(K,\dot{K}) \quad (1)$$

where Y is observed earnings, K is potential earnings ("human capital"), and $\dot{K}$ is its rate of change. Competitive equilibrium in the labor market places restrictions on the signs of the partial derivatives. Specifically, so long as human capital contributes to earnings (i.e., $\frac{\partial F}{\partial K} > 0$), the individual must give up current earnings in order to enhance future earning opportunities (i.e., $\frac{\partial F}{\partial \dot{K}} < 0$). It is typically assumed that, at each point in his life, the individual faces a spectrum of earnings-investment combinations ("jobs"), and selects the one which is optimal.

Equation (1) highlights the formal similarity to the investment problem of the firm in the presence of adjustment costs (see Eisner and Strotz [1963]). However, the fact that human

---

1/ This precise representation of the model is due to Rosen [1973], but similar formulations are implicit in Becker [1964] and Ben-Porath [1967].

capital is embodied in the individual and cannot be sold leads to several important differences. First, since the property rights to human capital cannot be transferred, the finiteness of life plays a central role in human investment. By contrast, even a firm which plans to cash in its assets after a fixed number of years will behave "as if" its horizon were infinite so long as its assets are marketable. Second, since either the utilization or formation of human capital requires the sacrifice of leisure (a specific consumption good), it is not possible in general to separate the consumption and investment problems as is done in the theory of the firm.[1/] Finally, and perhaps most importantly, neither human capital nor the investment therein are normally observable. Thus investment functions, such as those of Jorgenson [1963] and others, are not applicable to human capital theory. If the model fails to generate testable predictions about observable variables like wage rates and earnings, then it is not a very fruitful one.

All of these features make the human investment problem much more difficult theoretically than the firm's problem. Lack of data on human capital leads to still another difficulty: it is hard to distinguish empirically among the effects of investment

---

1/ See for example, Hirshleifer [1958].

and a myriad of other influences on actual earnings. The most obvious of these, of course, is the choice of work intensity, which affects both current earnings (labor-leisure choices) and the rate of human capital accumulation (training-leisure choices).

The purpose of this paper is to present a simple life-cycle model of investment in human capital in which leisure choices are explicitly incorporated. In so doing, we integrate two previously disparate branches of life-cycle theory: models of labor supply with exogenous wages,[1/] and models of human capital formation with exogenous leisure.[2/] Of course, to accomplish this, we must posit utility maximization as the individual's goal rather than income maximization.

Apart from the direct interest in the interaction of labor supply and human investment over the life cycle, such a model is needed to test the robustness of the widely-used wealth-maximization models of human capital accumulation. For example, a standard implication of these models is that a period of specialization in investment (interpreted as schooling), if it

---

1/ See, for example, Weiss [1972], Blinder [1974, Ch. 3], Heckman [1974]. The paper by Weiss allows for endogeneity of wages due to learning by doing, but not for human investment.

2/ See, for example, Ben-Porath [1967], Weiszacker [1967], Sheshinski [1968]. Formally, our approach also embraces life-cycle consumption theory, but we take pains to separate this from the other two problems, and have little to say about consumption.

exists, will occur only at the beginning of life.[1] However, one can imagine that when schooling involves foregone leisure the existence of pure time preference might lead a utility-maximizing individual to postpone his education. Such possibilities appear explicitly in our model. Another important implication of the wealth-maximization model is that the fraction of time spent investing falls throughout the post-schooling investment period. This turns out to be generally true in our model as well, though some exceptions are noted.

To our knowledge, there have been four previous attempts (all unpublished at this writing) to integrate human capital and labor supply as we do here. The treatment by Becker and Ghez [1972] is the most general and explores the widest variety of issues. But it is also the least ambitious in that they generally content themselves with stating and interpreting the first-order conditions. Our model can be viewed as a special case of theirs, but a case which is pushed much farther. The three other studies[2] at least attempt to analyze the shape of the optimal plan, and naturally adopt simplifying assumptions in order to do so. Typically, the rate of investment in human capital and the supply of labor are related to some key variable such as the stock of human capital or its shadow price (both unobserved). However, since the

---

[1] On this, see Weiss [1971], Ishikawa [1973], and also the references cited in the preceding footnote.

[2] Heckman [1975], Landsberger and Passy [1973], Stafford and Stephan [1973].

models generally do not fully determine the behavior of these endogenous variables, a disturbing and unnecessary amount of ambiguity remains. Further, except for Heckman [1975], virtually no attention is paid to periods of specialization such as schooling or retirement. This seems a major omission.

The model presented here is more general than previous work on the subject (with the exception of Becker and Ghez), and yet generates many more concrete conclusions. Among the restrictions which our model places on optimal plans are:

(a) Several distinct patterns in investment, work and leisure may arise, depending on the subjective rate of impatience (i.e., the discount rate for future utilities).

(b) We consider the case where the rate of impatience is "small" (in a sense to be defined later) to be the leading case. We show that in this case specialization in schooling can only occur at the beginning of life, while retirement can only come at the end.

(c) Schooling is followed by a period of on-the-job training, during which the fraction of potential earnings and time devoted to human capital formation declines monotonically.

(d) Investment reaches zero some finite time before retirement (or death, if there is no retirement). Thus there is a finite interval of "pure work" with no investment late in life. This econometric finding of Mincer [1974] is explicitly ruled out by the Ben-Porath [1967] formulation of the problem (which is followed by Heckman [1975] and by Stafford and Stephan [1973].

(e) The demand for leisure over the life cycle is "U-shaped", with a tendency to decline during schooling and the early part of OJT, and thereafter to rise.

(f) Wages rise to a single peak, which occurs <u>after</u> the peak in hours of work.

Of course, we do not pull these rabbits out of the proverbial hat. Like other investigators, we have to make some simplifying assumptions of which two seem most crucial. First, (1) is given the special form:

$$(2) \qquad Y = F(K,\dot{K}) = Kg(\dot{K}/K), \; g' < 0.$$

The nature of this important function is the subject of the next section. Here we only wish to point out that the formulation is quite similar to that of Rosen [1973] who dealt with the special case $F(K,\dot{K}) \equiv K - C(\dot{K})$, $C'(\dot{K})>0$. In both specifications, K is defined as earning capacity, i.e., the amount the individual would earn in the absence of investment. While Rosen orders jobs by their <u>absolute</u> rate of increase of earning potential ($\dot{K}$), and deducts costs <u>additively</u>, we order jobs by their <u>proportionate</u> rate of growth in potential earnings ($\dot{K}/K$), and deduct costs <u>multiplicatively</u>. In principle, either representation is as good as the other, but the multiplicative version leads more naturally to the logarithmic wage functions encountered so frequently in empirical work (Weiss [1974]).

Our second important assumption is that, for a given input of time, the quantity of human capital created is <u>proportional</u> to the stock of human capital. In the terminology popularized by Ben-Porath [1967], the "human capital production function" is homogeneous of degree one in K. This specification which is suggested by Mincer's work [1974], simplifies the mathematics considerably. Thus it is both "realistic" and convenient, a combination encountered all too rarely in mathematical economics.[1/]

[1/] We do not wish to oversell this assumption. Heckman's [1975] empirical work certainly does not support it.

2. <u>The Earnings-Investment Frontier</u>

Basic to any model of human capital is the market constraint delimiting the combinations of current earnings and human capital formation which are obtainable from a given stock of human capital. This function, which we call the "earnings-investment frontier", is defined implicitly by (1). It is clearly downward sloping. Several frontiers, corresponding to different levels of K, are depicted in Figure 1. (If there is depreciation, $\dot{K}$ can be negative.) Our special form (2) enables us to divide both axes of Figure 1 by K, so that a *single* frontier, applicable to *any* K, can be drawn. Thus it plays for us the role that constant returns to scale plays in growth theory. The frontier is sketched in Figure 2, where the new symbol, x, has replaced $\frac{\dot{K}}{K} + \delta$. As indicated in the introduction, we index jobs by their proportionate rate of growth. x is this index, and it runs from zero to unity. $x = 1$ indicates the job which yields the maximum feasible rate of growth. It makes sense to call this "going to school", and to specify that $g(1) = 0$, i.e., that all earnings are sacrificed during schooling. Conversely, $x = 0$ is the job where potential earnings are fully realized, so $g(0) = 1$.

The locus is a "frontier" in the usual sense: points (like p) in the interior are feasible, but never would be optimal; points (like q) beyond the frontier would be preferred by the worker, but are not attainable. We know, of course, that $g'(x)<0$, but can we say anything about its convexity? The answer is that we can rule out convexity (to the origin), but either linear or concave frontiers

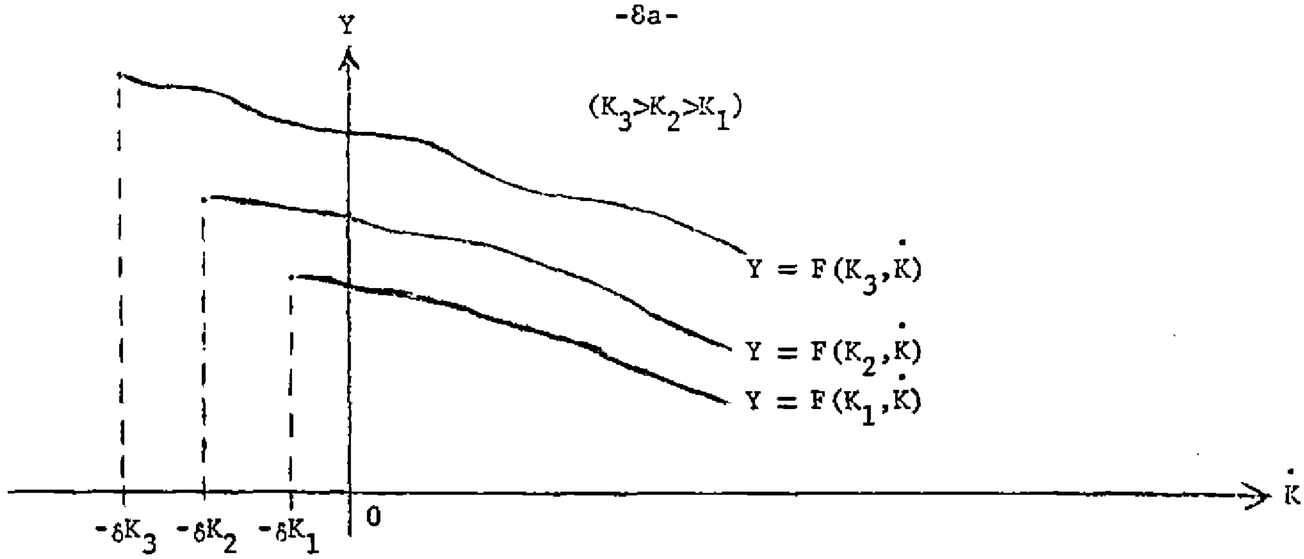

Figure 1
General earnings-investment frontiers

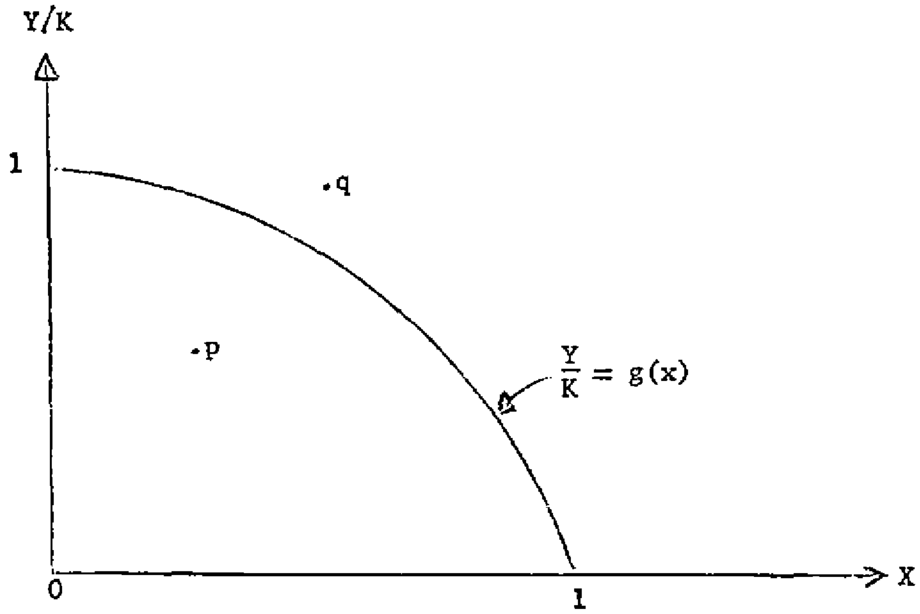

Figure 2
The specific earnings-investment frontier

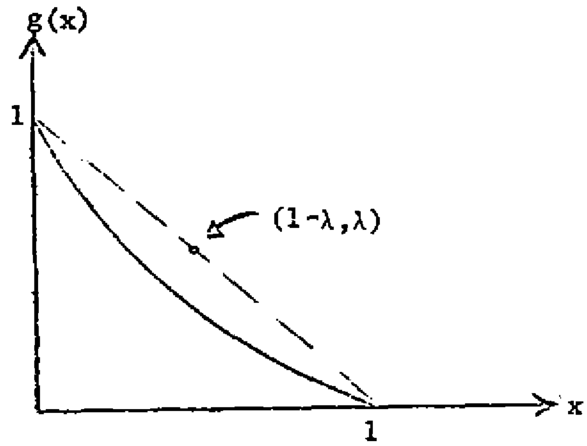

Figure 3
Proof that g(x) cannot be convex

seem possible. The argument is simple. If $g(x)$ were in fact convex, as in Figure 3, and arbitrage was permitted, no worker would accept a job with $0 < x < 1$. This is because he could always spend a fraction λ of his work day at $x = 0$ and a fraction $1-\lambda$ in school ($x = 1$), thereby creating for himself an opportunity set portrayed by the straight line $g(x) = 1-x$ (shown in Figure 3 as a dotted line).[1/] Since this line dominates his market opportunities he would never deal with the market. Employers offering jobs with $0 < x < 1$ would have to raise wages until <u>every</u> job was at least as attractive as arbitrage.

Can a similar argument be used to rule out concave $g(x)$ functions? We think not. Workers clearly prefer jobs above the $g(x) = 1-x$ line to arbitrage. Thus, to rule them out, one would have to show that firms offering such jobs would suffer losses in a competitive market. As this involves analysis of the production functions of the firms for both goods and human capital, such an argument would rest on specific assumptions about these functions. A full analysis of the behavior of firms in the human capital market is beyond the scope of this essay. We refer the interested reader to Rosen [1972], who argues that,due to diminishing returns, the provision of training by firms requires increasing marginal sacrifice of other outputs so that $g(x)$ must be concave.

[1/] The argument clearly assumes no frictions in moving between the job and the school.

While we consider both cases to be open possibilities, the bulk of our analysis assumes $g''(x) < 0$. As we indicate later, a model almost identical to the present one can be constructed using a linear frontier, $g(x) = 1-x$. There is only one important difference. A concave frontier means that combining work with education on the job is strictly better than blending attendance at school with a no-training job, while the linear frontier means they are equally good. If there is some positive interaction between earning and learning, the worker can capture some "pure profit" by on-the-job training. Therefore, the "full wage" (including the value of training), which is equated to the marginal rate of substitution (MRS) between consumption and leisure for optimality, exceeds his potential wage by the amount of pure profit. Only if $g(x) = 1-x$ will there be no pure profits, so that the MRS is equated to the potential wage, a conclusion which Becker and Ghez reach in their model (and claim to be very important).

The differences between $g''(x) < 0$ and $g''(x) = 0$ are at least partly observable. In the latter case, workers should be found "buying" training in schools just about as frequently as on the job. In the former case, training should be bought in schools only during the period of specialization; part-time education should take place predominantly on the job.

3. Statement of the Problem

The individual is assumed to derive utility from three sources: the stream of lifetime real consumption, $c(t)$; the time profile of leisure, $\ell(t)$; and the bequest, or terminal value of real (nonhuman) assets, $A(T)$.[1/] Here t denotes the individual's age, and runs from zero to T, the length of life which is assumed known and exogenous. Specifically, lifetime utility is assumed to be additively separable with a constant rate of time discounting:

$$(3) \qquad \int_0^T U(c,\ell)e^{-\rho t}dt \ , \ B(A(T)),$$

where $U(c,\ell)$ and $B(A(T))$ are assumed to be twice-differentiable strictly concave functions of their arguments, and ρ is what we call the rate of impatience.[2/] In order to rule out optimal paths with segments of zero consumption or zero leisure, we further assume:

$$\lim_{c\to 0} U_c(c,\ell) = \infty \quad \text{for all } \ell$$
$$\lim_{\ell\to 0} U_\ell(c,\ell) = \infty \quad \text{for all } c.$$

Finally, to allow retirement as an endogenous decision, we exclude from our analysis functions which preclude 100% leisure. Letting $\ell(t)$ denote the fraction of available time devoted to leisure, this amounts to assuming that $U_\ell(c,\ell) > 0$ for any c.

[1/] It will be assumed that a person's human wealth dies with him, i.e. cannot be bequeathed.

[2/] A variable rate of impatience could be accomodated without too much difficulty. But it would make the notation more cumbersome, and qualitative results would depend on the time profile of $\rho(t)$.

What are the constraints on this maximization problem? To begin with, there are two constraints on the time budget. Letting h(t) denote the fraction of time devoted to market activity (including work and education), we have:

$$h(t) + \ell(t) = 1 \quad (4)$$

$$h(t) \geq 0 \, . \quad (5)$$

As noted above, the occupational index is bounded between zero and unity:

$$0 \leq x(t) \leq 1 \, . \quad (6)$$

And, finally, the differential equations governing the change in the stocks of human and nonhuman capital are respectively:

$$\dot{K} = \phi(x,h,K) \quad (7')$$

$$\dot{A} = rA + Y(x,h,K) - c \quad (8')$$

where r is the real rate of interest,[1/] the function $\phi(\cdot)$ is the production function for human capital, and $Y(\cdot)$ gives earnings as a function of the job, the number of hours worked, and the level of human capital. This is actually considerably more general than the problem we solve. We begin this way in order to show explicitly how our simplifying assumptions reduce the problem to manageable proportions. There are two initial conditions, corresponding to

---

[1/] Letting r denote the real rate of interest implicitly incorporates changes in the price of consumer goods. To see this, note that if P were the price of goods and i were the nominal rate of interest, the change in nominal financial assets would be given by:

$$\frac{d}{dt}(AP) = i(AP) + PY - cP.$$

But $\frac{d}{dt}(AP) = \dot{A}P + \dot{P}A$, so dividing both sides of the equation by P leads to: $\dot{A} = (i - \frac{\dot{P}}{P})A + Y - c$, which is (8').

the individual's endowments of financial and human wealth,[1]

$$A(0) = A_0 \geq 0$$

$$K(0) = K_0 > 0 \quad ,$$

and both terminal stocks are to be chosen optimally.

To set up this problem in a form suitable for the application of Pontryagin's maximum principle, we substitute (4) into (3), define shadow princes $p(t)e^{-\rho t}$ for human capital and $\mu(t)e^{-\rho t}$ for financial wealth, and write the Hamiltonian function:

$$H(h,x,c,K,A,p,\mu) = e^{-\rho t}[U(c,1-h) + p\phi(x,h,K) + \mu(rA + Y(x,h,K) - c)].$$

First order necessary conditions for a maximum are:[2]

(i) at each instant, $x(t)$, $h(t)$ and $c(t)$ are chosen to maximize H, given K, A, p and μ, and subject to the constraints (5) and (6);

(ii) $\dfrac{\partial H}{\partial K} = -\dfrac{d}{dt}(e^{-\rho t}p)$ for all t;

(iii) $\dfrac{\partial H}{\partial A} = -\dfrac{d}{dt}(e^{-\rho t}\mu)$ for all t;

(iv) $K(T)p(T) = 0$

(v) $\mu(T) = B'(A_T)$.

Since optimal consumption is always positive, the condition for optimal c(t) is easily stated:

(9) $U_c(c,\ell) = \mu.$

---

[1] In interpreting these, it should be noted that K_0 includes both the initial endowment of education and "ability", or whatever else it is that determines productivity.

[2] We have been unable to provide a proof of sufficiency.

And, from (iii)

(10) $$\dot{\mu} = (\rho - r)\mu \ ,$$

so the marginal utility of consumption grows or declines at a steady exponential rate. The consumption-bequest plan has been analyzed extensively elsewhere, and we shall pay little attention to it hereafter.[1/]

Consider next the first-order conditions for optimal x and h, assuming that an interior solution obtains for each (we shall worry about corners presently). They are:

(11") $$U_\ell(c,\ell) = p\phi_h(x,h,K) + \mu Y_h(x,h,K)$$

(12") $$p\phi_x(x,h,K) + \mu Y_x(x,h,K) = 0 \ .$$

We now introduce the following assumption, mentioned in the introduction:

> A1: The functions $\phi(x,h,K)$ and $Y(x,h,K)$ are both homogenous of degree one in K. Specifically,
>
> $\phi(x,h,K) = \Phi(x,h)K$ and $Y(x,h,K) = y(x,h)K$.

This reduces (11") and (12") to:

(11') $$U_\ell(c,\ell) = pK\Phi_h(x,h) + \mu K y_h(x,h)$$

(12') $$pK\Phi_x(x,h) + \mu K y_x(x,h) = 0.$$

So A1 transforms a problem with three state variables (p, μ and K) to one with only two state variables: pK, which is the shadow value (in utils) of the stock of human capital; and K, which is the potential wage rate converted into utils.

[1/] See Atkinson [1971] or Blinder [1974, Ch. 2]. For a treatment which ignores bequests, see Yaari [1964] or Weiss [1972].

The next simplification comes from defining the occupational index, x, in the way suggested in Section 2. Namely, x indexes jobs in terms of equivalent schooling time, so that working h hours at job x is equivalent (in the production of human capital) to spending xh hours in school. Note that this is not a restrictive assumption, but a convenient measurement convention which enables us to write:

$$\phi(x,h) = \phi(xh). \tag{13}$$

While the analysis could be carried out more generally, it will keep things clear if we assume:

> A2: The hourly wage rate does not depend on hours of work. In symbols: $y(x,h)K = g(x)hK$.

This prevents the issues we wish to focus on -- especially the interplay between human investment and work effort -- from becoming clouded with other issues such as overtime pay or penalties for part-time work. Using equation (13) and A2, (11') and (12') reduce to:

$$U_{\ell}(c,\ell) = f'(xh)pKx + \mu K g(x) \tag{11}$$

$$h[f'(xh)pK + \mu K g'(x)] = 0 \quad . \tag{12}$$

Most of our broad results, including the general characterization of the life cycle plan, can be obtained from these two equations (and the corresponding inequalities for corner solutions). However, to get certain more severe (and empirically testable) restrictions on the optimal path, we need to make one of the following linearity assumptions:

A3: The function $f(xh)$ is linear, i.e. $f'(xh) = a$, for all x, h.

or

A3': The function $g(x)$ is linear, i.e. $g'(x) = -1$ for all x.

The meaning of A3' is discussed in Section 2. What is the economic interpretation of A3? Mincer [1974] has argued, both theoretically and empirically, that the functional relation between potential earnings and accumulated years of schooling (denoted by the symbol S) takes the following specific form:

$$K = \gamma\, e^{aS} e^{-\delta t}$$

where δ is the rate of depreciation of earning capacity (possibly zero). Differentiating this logarithmically yields:

$$\frac{\dot{K}}{K} = a\,\dot{S} - \delta .$$

But, by the way we defined x, it is definitionally true that $\dot{S} = xh$,[1/] so we have:

$$\frac{\dot{K}}{K} = axh - \delta. \qquad (7)$$

As the reader will recognize, this is just our general human capital production function (7'), modified as per A1, A3, and (13). To wit, A3 is essentially equivalent to assuming that the Mincer wage function holds.

1/ The reader familiar with Weiszacker [1967] or Sheshinski [1968] will recognize this as an extension of their formulation to allow for variable labor supply.

We present here the results of the model using A3, rather than A3'. However, we have also worked out the model with A3',[1/] and can report that the two models give identical results, with only one exception. The exception was noted in our discussion of the earnings-investment frontier: when the frontier is linear, there is no distinction between the potential wage and the "full wage" (to be defined below). Since neither is observable, this difference is probably subtle enough to be ignored. Mathematically, nonlinearity in $g(x)$ turns out to be almost a perfect substitute for nonlinearity in $f(xh)$, and equation (12) shows why this is so. Clearly, either $f'' < 0$, or $g'' < 0$, but not both, is required for a regular interior maximum for x. Making both functions concave is a kind of "overkill" which complicates the mathematics considerably (and, we would argue, needlessly). While we conjecture that all of our results go through to this more general case, we have been unable to prove it.

To summarize, then, our model is the maximization of (3), subject to the constraints (4)-(6), and the differential equations (7) and:

(8) $\dot{A} = rA + hKg(x) - c$.

The first order conditions for optimal h and x, including now

---

[1/] If $f(xh)$ takes the special constant-elasticity form, this model is very close to Heckman's [1975].

the Kühn- Tucker corner conditions, are:

(14) $$U_{\ell}(c,1-h) \geq apKx + \mu Kg(x)$$

$$\text{with} > \rightarrow h = 0$$

$$h[apK + \mu Kg'(0)] < 0 \rightarrow x=0$$

(15) $$h[apK + \mu Kg'(1)] > 0 \rightarrow x=1$$

otherwise, $h[apK + \mu Kg'(x)] = 0$.

4. Interpretation of the Optimality Conditions

In the next section, we develop a phase diagram to assist us in understanding when the individual elects one of the three possible corner solutions ($x = 0$, $x = 1$, or $h = 0$) rather than the interior solution. To facilitate this, we will need one further assumption:

> A4: Instantaneous utility is separable in consumption and leisure, i.e., $U(c,\ell) = u(c) + v(\ell)$.

However, before doing this, greater understanding of the problem can be achieved by recasting and interpreting the first-order conditions.

Recall that $p(t)$ and $\mu(t)$ are, respectively, the shadow prices (in current utils) of human and nonhuman capital,[1/] so the ratio, $p(t)/\mu(t) = \theta(t)$, denotes the money price of human capital. Thus we can define $V(t) = \theta(t)K(t)$ as the money value of the human capital stock, and trace of behavior of $V(t)$ over time. To do this, use condition (ii) to write:

$$\frac{\dot{p}}{p} = (\rho+\delta) - axh - \frac{g(x)hK}{V}. \quad (16)$$

Then by (7) and (10):

$$\dot{V} = rV - hg(x)K. \quad (17)$$

Finally, since human capital cannot be cashed in at the end of life ($K(T) > 0$), the transversality condition (iv) implies that $V(T) = 0$.

[1/] Consumption goods are the numeraire, hence there is no distinction between a dollar of money and a unit of consumer goods.

This terminal condition allows us to solve differential equation (17) explicitly viz:

$$(18) \qquad V(t) = \int_t^T e^{-r(\tau-t)} h(\tau) g(x(\tau)) K(\tau) d\tau.$$

That is, V(t) is the present value, in time t dollars, of future earnings -- a very natural interpretation of the value of the human capital stock. Equation (17) expresses the fact that the optimal paths of the shadow prices, p(t) and $\mu(t)$, will render the individual incapable of making his human wealth grow faster than the rate of interest, and he can achieve this limiting growth rate only by staying in school ($x = 1$). Looking at human wealth in terms of time $t = 0$ dollars, (17) says that human wealth stays constant during schooling, and thereafter declines _pari passu_ with earnings. Earnings, hg(x)K, are like withdrawals from a "human bank account" where interest is compounded continuously at rate r.

Using the variable V(t), we can rewrite the optimality condition (15) in a form which is readily interpretable:

$$(15') \qquad \begin{array}{llll} -g'(x)K & = & aV & \text{if} \quad 0 < x < 1 \\ -g'(0)K & \geq & aV & \text{if} \quad x = 0 \\ -g'(1)K & \leq & aV & \text{if} \quad x = 1 \end{array}$$

Here $-g'(x)K$ is the marginal cost in terms of foregone earnings per hour of raising x, and it can be shown that aV measures the marginal benefits in higher future earnings.[1/]

Condition (15') can also be given the usual rate-of-return interpretation: Let π be the rate of interest that equates the discounted benefits to the costs; that is, π is implicitly defined by:

$$-g'(x)K = a \int_t^T e^{-\pi(\tau-t)}h(\tau)g(x)K(\tau)d\tau$$

[1/] More precisely, let x(t), h(t) be the optimal policy, and consider a deviation from this policy as follows:

$$\hat{x}(t) = x(t) \text{ for } t \geq t_1 + \tau \text{ or } t \leq t_1$$
$$= x(t) + \varepsilon \text{ for } t_1 < t < t_1 + \tau$$

while h(t) is unchanged. Assuming separability of $U(c,\ell)$ (which is not required, but which simplifies the proof) a necessary condition for the optimum is that such a shift leaves lifetime income unchanged. Since nothing earlier than t_1 can be affected, we might as well measure income from t_1 forward. The change in income is therefore:

$$\Delta J = \int_{t_1}^T e^{-r(t-t_1)}h(t)g(x)[\hat{K}(t)-K(t)]dt + \int_{t_1}^{t_1+\tau} e^{-r(t-t_1)}h(t)K(t)[g(x+\varepsilon)-g(x)]dt$$
$$+ \int_{t_1}^{t_1+\tau} e^{-r(t-t_1)}h(t)[\hat{K}(t)-K(t)][g(x+\varepsilon)-g(x)]dt,$$

where $\hat{K}(t)$ is the capital stock associated with the alternative program. Specifically,

$$\log \hat{K}(t) = \log K(t) + \int_{t_1}^{t_1+\tau} ah(t)\varepsilon dt, \ t > t_1 + \tau.$$

If we take $\lim_{\substack{\tau\to 0\\ \varepsilon\to 0}} \frac{\Delta J}{\varepsilon\cdot\tau}$, the first term (which clearly measures the benefits) goes to $ah(t_1)V(t_1)$, the second term (which clearly measures costs) goes to $-h(t_1)K(t_1)g'(x(t_1))$, and the third term vanishes. The equality in (15') results from setting $\lim_{\substack{\tau\to 0\\ \varepsilon\to 0}} \frac{\Delta J}{\varepsilon\cdot\tau} = 0$.

as the marginal internal rate of return. Then an equivalent statement of (15') is:

$$\pi = r \quad \text{if} \quad 0 < x < 1$$
$$\pi \leq r \quad \text{if} \quad x = 0$$
$$\pi \geq r \quad \text{if} \quad x = 1 \quad .$$

Thus during the schooling period ($x = 1$), the internal rate of return on human investment is strictly greater than the interest rate, despite the fact that hours-of-study are variable.[1/] During the post-school investment period, the marginal rate of return is always equal to the rate of interest. It is worth noting that these results hold for *any* arbitrary leisure profile, and are thus independent of tastes. Therefore, one can verify *ex post* whether individuals have behaved optimally by computing the appropriate rates of return. Of course, when labor supply is a choice variable, the rate of return is ill-defined as an *ex ante* concept.

Let us know turn to the condition for optimal work effort, (14). When there is no investment ($x = 0$), wages are exogenous to the individual,[2/] and, in view of (9), (14) simply states that the marginal rate of substitution is equal to the (potential and actual) wage. Analysis of the age-hours profile is exactly as in our previous

---

[1/] Equality holds at the instant of leaving school. For analogous conditions in an income-maximization model, see Weiss [1971b].

[2/] Of course, part of the problem is to determine when the individual will opt for leaving his wage exogenous.

models of dynamic labor supply with exogenous wages.[1/] Becker and Ghez [1972] have claimed that this marginal condition also holds when wages are endogenous, and that therefore for the analysis of time-versus-goods substitution in consumption we need not worry about why wages change. Two remarks can be made about this finding. First, when there is investment, the observed wage and the potential wage are very different. Nor is their behavior over time identical; we show later that during OJT the ratio of observed wage/potential wage is monotonically rising. Second, as indicated earlier, the MRS is equated to the potential wage only in the special case of a linear earnings-investment frontier. To see, divide the left-hand side of (14) by $U_c(c, \ell)$ and the righthand side by μ (they are equal by (9)) to get:

$$\frac{U_\ell}{Uc} = a \frac{p}{\mu} Kx + Kg(x).$$

Then substitute for p/μ from (15) (assuming an interior solution) to get:

$$\text{(19)} \qquad \text{MRS} \equiv \frac{U_\ell}{Uc} = K[g(x) - xg'(x)].$$

It is easy to verify that the function in square brackets is identically unity only in the case $g(x) = 1 - x$. Otherwise, it is always <u>strictly greater</u> than unity (as long as $0 < x < 1$). We call the righthand side of (19) the "full wage". Clearly the first component, $g(x)K$, measures the <u>current</u> benefits per hour of work. That the second component, $-xg'(x)K$, measures the <u>future</u>

[1/] See Weiss [1972], Blinder [1974, Ch. 3].

benefits per hour of today's work follows from the preceding calculation. That is, the second component of the "full wage" prices x at its marginal (not average) value in computing the future benefits. The full wage exceeds the potential wage by what we called the "pure profits" from OJT in Section 2.

5. <u>The Temporal Succession of Life-Cycle Phases</u>

Due to the possibility of corner solutions, four qualitatively distinct phases might occur in an individual's life cycle. Whether each phase occurs, how long each phase lasts, and the temporal order of the phases form the subject of this section. We give the four phases descriptive names, and numbers which seem to indicate a sensible ordering, as follows:[1/]

Phase I, <u>Schooling</u>: $x = 1$, $h > 0$

Phase II, <u>OJT</u>: $0 < x < 1$, $h > 0$

Phase III, <u>Work</u>: $x = 0$, $h > 0$

Phase IV, <u>Retirement</u>: $h = 0$

Note that when $h = 0$ the value of x is arbitrary as is clear from (15). Economically, this says that a retired person can be considered as holding any job at all (and working zero hours).

To facilitate the exposition, we define two new state variables:

$\lambda(t) = ap(t)K(t)$ = the value (in utils) of human capital, multiplied by a

$\psi(t) = \mu(t)K(t)$ = the potential wage, converted to utils

and adopt assumption A4 (separable utility). Then the four phases are characterized by the following equations:

<u>Phase I (Schooling)</u>

(20.1) $$v'(1-h) = \lambda$$

(21.1) $$\frac{\lambda}{\psi} > -g'(1)$$

[1/] Of course, we have to prove that Phase I comes before Phase II, Phase II comes before Phase III, and so on.

Phase II (OJT)

(20.2) $$v'(1-h) = \lambda x + \psi g(x)$$

(21.2) $$\frac{\lambda}{\psi} = -g'(x)$$

Phase III (Work)

(20.3) $$v'(1-h) = \psi$$

(21.3) $$\frac{\lambda}{\psi} < -g'(0)$$

Phase IV (Retirement)

(20.4) $v'(1) > \lambda x + \psi g(x)$ for all $x \in [0,1]$.

It is clear from (21.1) and (21.3) that the precise shape of g(x) strongly colors the likelihood and length of each phase. For example, suppose $-g'(1) = \infty$, as in Figure 4a. Then (21.1) makes it clear that the individual will never specialize in schooling. The reason is that there are jobs (x's very near unity) which are nearly as effective as schooling in imparting skills, but which pay a positive wage. So finiteness of $-g'(1)$ is a minimal requirement for schooling to be possible. We assume this is true.

As another extreme, suppose $g'(0) = 0$, as in Figure 4b. Then (21.3) shows that Phase III (work without training) is impossible. This is because there would be a job with a wage arbitrarily close to the no-training job, but which gave a finite amount of training.

To portray our model in (λ,ψ) space we must locate the boundaries of each region. Assume first that $h > 0$, so we need only worry about the pre-retirement phases. Since g(x) is continuous from the left at $x = 1$, (21.1) defines the boundary of the schooling region as the ray, $\lambda = -g'(1)\psi$, in Figure 5. Similarly, since g(x) is continuous from the right at $x = 0$, (21.3) defines

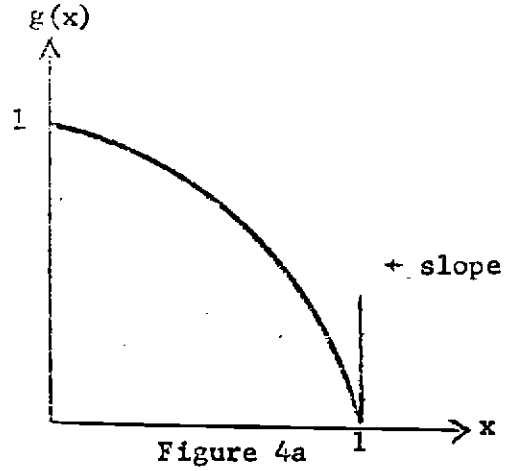

Figure 4a

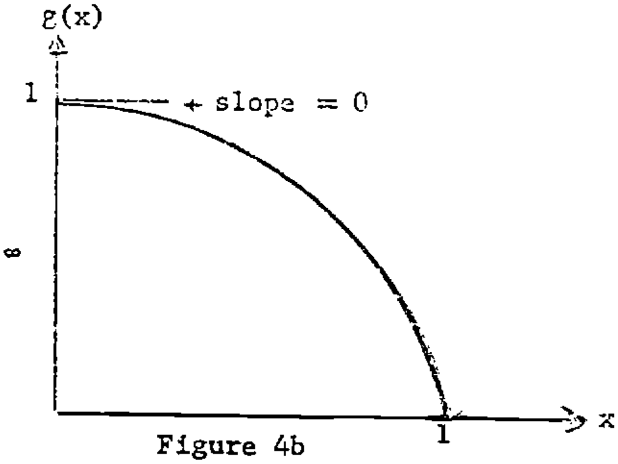

Figure 4b

Alternative g(x) functions which preclude Phase I (Figure 4a) or Phase II (Figure 4b)

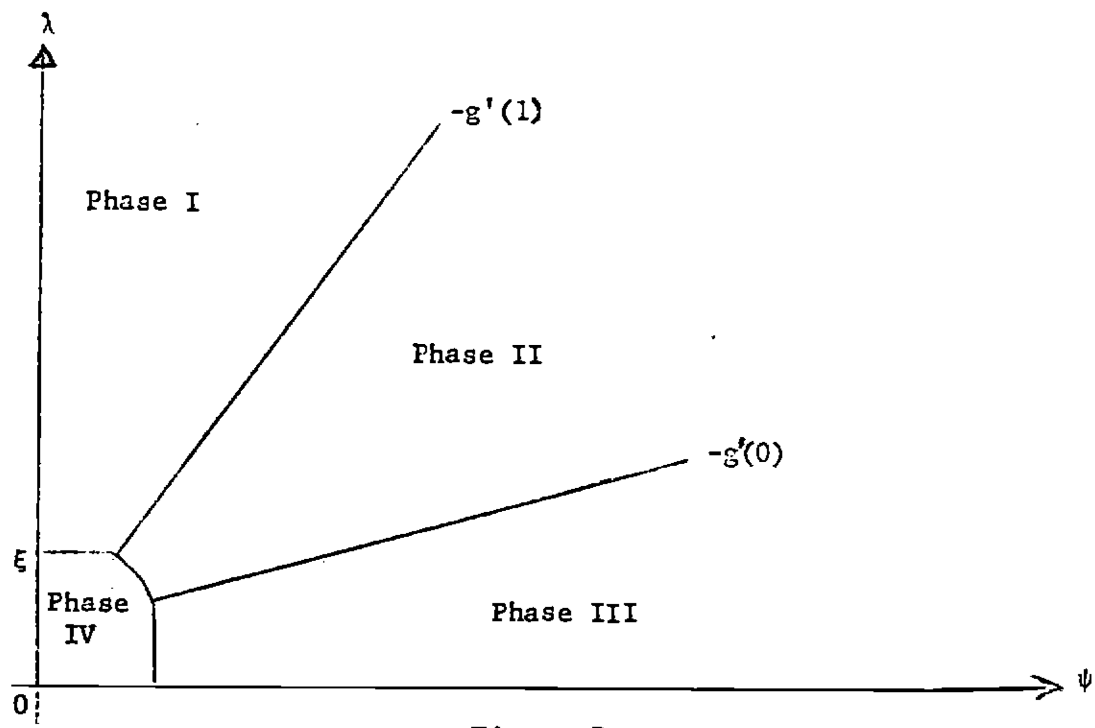

Figure 5
The four phases in (λ,ψ)-space

the boundary of the x = 0 region as the ray, $\lambda = -g'(0)\psi$, in Figure 5. Finally, the boundary of the h = 0 ("retirement") region is given by the solution of:

$$\xi \equiv v'(1) = \max_{0 \leq x \leq 1} \{\lambda x + \psi g(x)\}$$

This last boundary defines a convex[1/] set such that for any pair (λ, ψ) in the interior, the individual will allocate all of his time to leisure.

In this space, the optimal trajectory can, in principle, begin anywhere. But, since the transversality condition is p(T) = 0, it must terminate on the horizontal axis. This trivial observation already establishes that anyone who ever works will indeed have a Phase III, i.e. a finite period of work without OJT. In other words, so long as we do not rule Phase III out of court by

---

1/ Define the function

$$G(\lambda,\psi) \equiv \max_{0 \leq x \leq 1} \{\lambda x + \psi g(x)\}.$$

It is obvious that, regardless of the shape of g(x):

$$\gamma[\max_{0 \leq x \leq 1} \{\lambda_1 x + \psi_1 g(x)\}] + (1-\gamma)[\max_{0 \leq x \leq 1} \{\lambda_2 x + \psi_2 g(x)\}] \geq$$
$$\max_{0 \leq x \leq 1} \{\gamma[\lambda_1 x + \psi_1 g(x)] + (1-\gamma)[\lambda_2 x + \psi_2 g(x)]\}$$

for any $0 \leq \gamma \leq 1$. This says $G(\lambda,\psi)$ is a convex function, and the equation $G(\lambda,\psi) = \xi$ is our desired boundary. Note that when the maximum is found at x = 1, $G(\lambda,\psi) = \lambda$; and when x = 0 is the maximum, $G(\lambda,\psi) = \psi$. Thus above the $-g'(1)$ ray, the boundary is $\lambda=\xi$; and below the $-g'(0)$ ray, the boundary is $\psi=\xi$.

assuming $-g'(0) = 0$, Phase III must be last phase before retirement (or death) in any optimal plan. We believe this makes sense, and is an argument against using the Ben-Porath [1967] specification.

One other interesting observation can be made already. If ψ, the potential wage, is less than ξ, the marginal utility of complete leisure, then the work activity is unattractive. Similarly, if $\lambda < \xi$, the schooling activity is unattractive. However, Figure 5 shows that there is a region where both of these inequalities hold, and yet the individual chooses not to retire. This means that, for some values of ψ and ξ, the option of combining work and schooling by OJT keeps the individual on the job when he would otherwise retire. It can be shown that this is a direct implication of the concavity of the earning-investment frontier.[1/]

To determine the nature of optimal trajectories in Figure 5, it is necessary to know how λ and ψ change over time. Using the definitions of λ and ψ, and equations (7), (10), and (16), it follows that:

$$\dot{\lambda} = \rho\lambda - a\psi g(x)h \quad (22)$$

$$\dot{\psi} = (ahx-\gamma)\psi \ , \quad (23)$$

where $\gamma \equiv r - \rho + \delta$. The sign of γ turns out to be quite important.

During Phase I, since $g(1) = 0$, λ rises at the exponential rate ρ, but the sign of $\dot{\psi}$ depends on the level of h if $\gamma > 0$.

[1/] In the case $g(x) = 1 - x$, the retirement regions expands to fill the entire square.

Specifically, the $\dot{\psi} = 0$ locus for Phase I (drawn in Figure 6) is simply $h = \gamma/a$, where h is a function of λ by (20.1). If $\gamma \leq 0$, ψ is rising throughout Phase I (see Figure 7).

In Phase III, ψ falls or rises at the rate γ, but the behavior of λ is not so simple. When $x = 0$, equation (22) becomes:

$$\text{(22')} \qquad \dot{\lambda} = \rho\lambda - a\psi h \ ,$$

where h is implicitly a function of ψ by (20.3). It can be shown that the $\dot{\lambda} = 0$ locus (depicted in Figures 6 and 7) begins at the point $\lambda = 0$, $\psi = \xi$, slopes upwards, and intersects the $-g'(0)$ ray at a value of h corresponding to $h_A = -g'(0)\frac{\rho}{a}$.[1/]

Phase IV is the simplest. Here λ rises at the exponential rate ρ, while ψ falls or rises at the exponential rate γ. (See Figures 6 and 7).

Behavior of the trajectory in Phase II is governed by the pair of differential equations (22)-(23), with x and h implicitly defined as functions of λ and ψ by (20.2) and (21.2). These can be used to locate and analyze the two stationary loci: $\dot{\lambda} = 0$ and $\dot{\psi} = 0$. The task is rather arduous and uninteresting, and hence

---

[1/] <u>Proof</u>: The locus is defined by $\lambda = \frac{a}{\rho} \Psi h$. By (20.3), as $\psi \to \xi$ from above, $h \to 0$. Therefore $\lambda \to 0$. Next suppose $\lambda = -g'(0)\psi$ for some point on the locus. Then clearly $h = -g'(0)\frac{\rho}{a}$. This establishes the two end points. The positive slope follows immediately once it is noticed that (20.3) implies $\frac{dh}{d\psi} = \frac{1}{v''(\ell)} > 0$. Q.E.D.

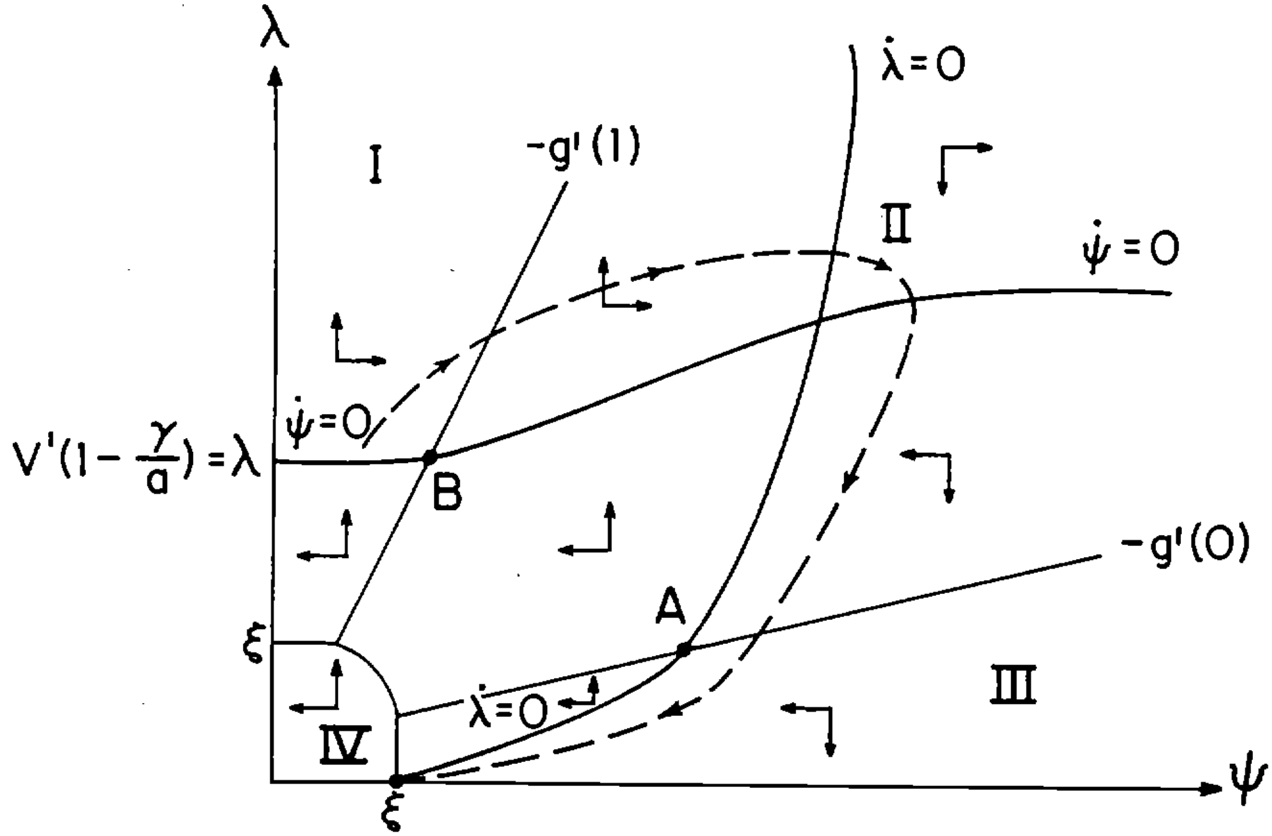

Figure 6
Phase diagram when $\gamma > 0$

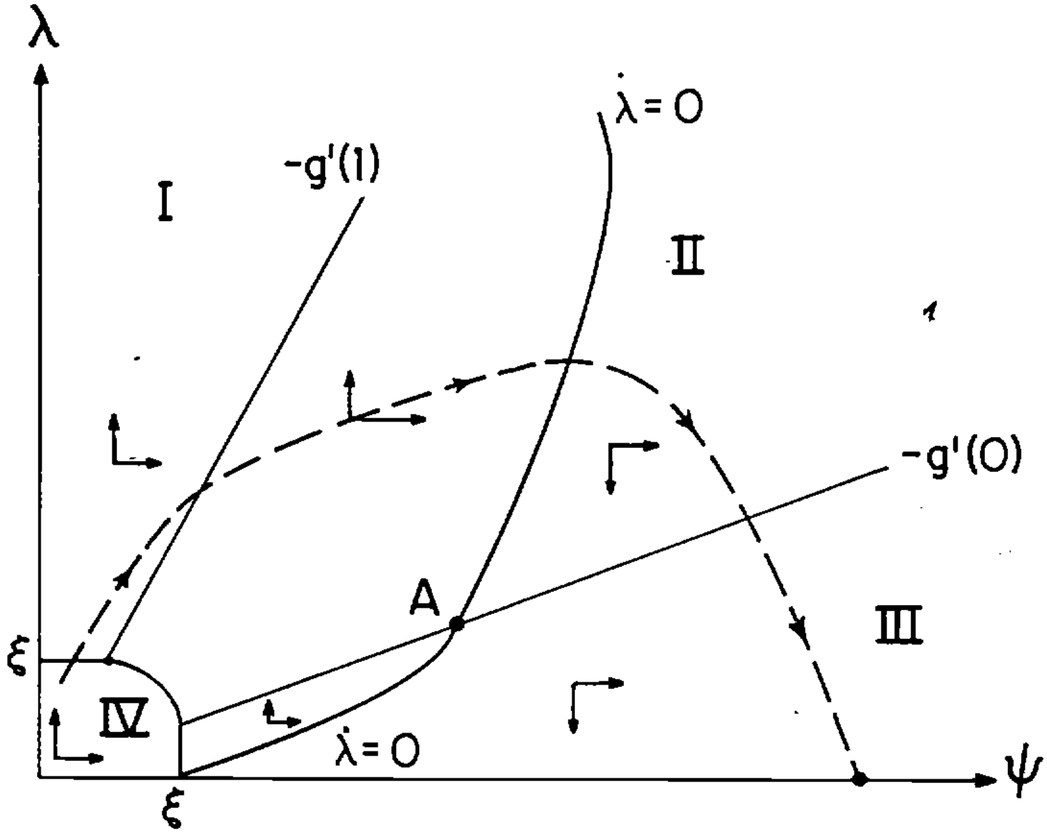

Figure 7

Phase diagram when $\gamma < 0$

is relegated to the appendix, but a few simple properties can be seen immediately:

(a) When $\gamma \leq 0$, ψ is always rising in Phase II, as indicated in Figure 7.

(b) When $\gamma > 0$, there is a $\dot{\psi} = 0$ locus which intersects the $-g'(1)$ ray at the point B in Figure 6 (When $x = 1$, $h = \gamma/a$ on the locus, which defines point B.), and which never touches the $-g'(0)$ ray. (Because $ahx = \gamma$ cannot be satisfied when $x = 0$.)

(c) The $\dot{\lambda} = 0$ locus is the same regardless of the sign of γ. It intersects the $-g'(0)$ ray at point A (If $\dot{\lambda} = 0$ while $\lambda = -g'(0)\psi$, then $h = h_A$.), and does not touch the $x = 1$ ray. (When $x = 1$, $g(x) = 0$, so no positive λ satisfies $\dot{\lambda} = 0$.)

We have drawn both loci with positive slopes. In fact, while they both must start this way (at points A and B), either slope or both might turn negative. This is not important. What is important, and what can be proven (see appendix) is that they always have a unique intersection which is in Phase II, and that the $\dot{\lambda} = 0$ locus cuts the $\dot{\psi} = 0$ locus from below.

It is clear from Figures 6 and 7 that the temporal succession of life-cycle phases will be quite different in the two cases. We believe the $\gamma > 0$ case is more "realistic", and hence will concentrate on it. Our reason is that people with $r \geq \rho$ (and, *a fortiori*, $\gamma > 0$ as long as $\delta > 0$) have an optimal consumption path which is everywhere nondecreasing, while persons with $\rho > r$ have a declining optimal consumption path.[1/] Unfortunately, this case is the more

---

[1/] This simple identification of rising or falling consumption paths with $r-\rho$ is true only under separable utility. For the more general case, see Weiss [1972]. Our point is that this specific model behaves most like the real world when $r \geq \rho$.

difficult. The reason is a phenomemon we call "cycling", that is, the recurrence of a specific phase more than once in an optimal life cycle. For example, inspection of Figure 6 raises the possibility that it might be optimal to attend school, then take OJT, and then return to school again. The reader will doubtless note many other possibilities for cycling when $\gamma > 0$, but there is little room for cycling when $\gamma \leq 0$.

Let us, therefore, deal first with people with "high impatience", i.e. $\rho > r + \delta$.[1/] The following propositions, which rather sharply delimit the kinds of life cycles that could ever be optimal for such a person, follow by inspection of Figure 7.

(a) The only possibility of cycling is that an optimal path might include two disjoint periods of work, separated by a period of OJT.

(b) If there is a period of schooling, it comes either at the beginning of life or immediately after retirement. A period of OJT follows schooling.

(c) If there is a period of retirement, it comes at the beginning of life.[2/]

(d) If there is a period of OJT (and there will be one unless the optimal path includes no training whatever), it is followed by a period of work.

(e) The last years of life are a period of work.

Thus the "normal" life cycle for persons with high impatience (depicted by the dotted path in Figure 7), assuming most people find it optimal to take some schooling and some retirement, is roughly as

---

[1/] Note that if the rate of growth of wages in the absence of investment, $-\delta$, is <u>positive</u> (due to economy-wide productivity changes) and exceeds the real rate of interest, then all persons are in this class.

[2/] The same phenomenon occurs in a dynamic labor supply model with exogenous wages. See Weiss [1972], Blinder [1974, Ch. 3].

follows: an initial period of retirement is followed by schooling, then by OJT, and then by pure work until death.

Of course, life-cycles with early retirement are rarely observed in practice. But this is not because such programs are irrational. Individuals with very high impatience (or very high positive exogenous wage growth) will want to bunch their leisure early in life. To do do, since consumption depends on lifetime discounted earnings, they will have to work very hard when they are old. We may surmise that it is the absence of perfect capital markets that precludes all but inheritors of large fortunes from pursuing such a program.

When impatience is lower, so that $\gamma = r + \delta - \rho > 0$, cycles seem to be possible. Section 7 is devoted to this issue. For the moment, we simply assume that there are no cycles.

6. Labor Supply and Human Investment in a Normal Life Cycle

It can be seen from Figure 6 that, if the life plan includes schooling and has no cycles, the only possibility is: schooling comes first, followed by OJT, work, and then retirement.[1/] We therefore call this the "normal" life cycle, and this section is devoted to examining its properties.

6.1 Phase I: Schooling

Since λ is rising through time during Phase I, and

$$v'(1-h) = \lambda \ , \tag{20.1}$$

it is clear that $\dot{h}(t) > 0$. That is, the amount of time devoted to schooling rises steadily as education progresses -- a prediction to which all former graduate students will doubtless attest. The intuition behind this is simply that the value (in utils) of the human capital stock $\lambda(t)$ rises (even if K itself falls), making leisure more expensive.[2/]

We may also make some rough judgments on the concavity of the hours profile, h(t). Differentiating (20.1) logarithmically gives:

$$\left[\frac{-v''(\ell)}{v'(\ell)}\right]\dot{h} = \frac{\dot{\lambda}}{\lambda} = \rho \ .$$

[1/] This is actually not quite so obvious, since the diagram makes it look as though retirement might come first. However, we show in Section 7 that the same conditions that preclude cycling also preclude early retirement.

[2/] Marginal calculations during this phase involve only leisure and schooling. The work activity is dominated.

Denote the expression in brackets by $R(h) > 0$, and take the time derivative again to get:

$$R'(h)(\dot{h})^2 + R(h)\ddot{h} = 0 \quad (24)$$

It is clear from (24) that $\ddot{h}$ and $R'(h)$ have opposite signs. What is the likely sign of $R'(h)$? Noting that the time budget makes $dh/d\ell = -1$ everywhere, we can write $R'(h)$ as:

$$R'(h) = \frac{d}{d\ell} \frac{v''(\ell)}{v'(\ell)} .$$

In the case of utility functions for choices involving risk, the ratio $v''(\ell)/v'(\ell)$ is called the degree of "absolute risk aversion" (Pratt [1964]), and is generally thought to be an increasing function of ℓ (i.e., to fall in absolute value as ℓ rises). While the present model does not discuss risk, it seems reasonable to suppose that our $v(\ell)$ function also has this property. In that case, (24) implies that $\ddot{h}(t) < 0$, i.e. that h is a concave function of time during schooling.[1/] We take this to be the leading case. And since in Phase I $\frac{\dot{K}}{K} = ah - \delta$, the rate of growth of potential wages would also be an increasing and concave function of time.

Note that all this is independent of γ, and hence applies equally well to individuals with high impatience.

[1/] If, instead $\frac{d}{d\ell}\left(\frac{v''(\ell)}{v'(\ell)}\right) < 0$, h will be a convex function of time.

6.2 Phase III: Work

Since there are some formal similarities between Phase I and Phase III, it is convenient to take up next the case of work with no training.

Here labor supply is governed by the usual condition: that the marginal rate of substitution between leisure and consumption goods be equated to the real wage:

$$v'(1-h) = \psi \equiv u'(c)K. \tag{20.3}$$

Since, by (23), $\dot{\psi}/\psi = -\gamma$ in Phase III, logarithmic differentiation of (20.3) yields:

$$\frac{-v''(\ell)}{v'(\ell)}\dot{h} = -\gamma = \rho-r-\delta, \tag{25}$$

so people with high impatience offer increasing amounts of labor to the market, despite wages which are falling if $\delta > 0$. Contrariwise, persons with more "normal" time discount rates have diminishing labor supply in Phase III. It cannot be stressed too much that these contrasting behavior patterns have absolutely nothing to do with competing "income" and "substitution" effects, although cross-sectional studies of labor supply might possibly confound the two phenomena.

From (25) it is clear that the concavity issue is precisely as it was in Phase I. The sign of $\ddot{h}$ depends only on the behavior of "absolute risk aversion" as ℓ rises, and the more attractive utility functions imply $\ddot{h} < 0$ (regardless of the sign of γ).

The behavior of wage rates (actual and potential coincide) and earnings in Phase III is also of interest. Letting W be the observed wage, we have:

$$\frac{\dot{W}}{W} = \frac{\dot{K}}{K} = -\delta, \text{ a constant.}$$

So <u>logarithmic</u> age-wage profiles should be straight lines in Phase III, which are falling, flat, or rising according as δ is greater than, equal to, or less than zero.

Finally, consider earnings, $Y(t) = h(t)W(t)$. We have:

$$\frac{\dot{Y}}{Y} = \frac{\dot{h}}{h} + \frac{\dot{W}}{W} = \frac{\dot{h}}{h} - \delta .$$

If $\delta > 0$, i.e., if depreciation outweighs economy-wide wage growth, earnings will surely decline. However, if δ is sufficiently negative, they may not. But regardless of the slope of the logarithmic earnings profile, it will certainly be concave if h(t) is:

$$\frac{d}{dt}\left(\frac{\dot{Y}}{Y}\right) = \frac{d}{dt}\left(\frac{\dot{h}}{h}\right) = \frac{\ddot{h}h - (\dot{h})^2}{h^2} < 0 \quad \text{if } \ddot{h} < 0.$$

<u>6.3 Phase II: On-The-Job-Training</u>

To analyze the behavior of hours of work and investment in Phase II, it is convenient to transform the differential equations so that x and h, rather than λ and ψ, are the variables. In (x,h)-space, Phase II is the <u>open</u> unit square, Phase I is the vertical line $x = 1$, Phase III is the vertical line $x = 0$, and Phase IV is the horizontal axis. (See Figure 8.) In a life-cycle without cycling, we enter from Phase I and exit to Phase III.

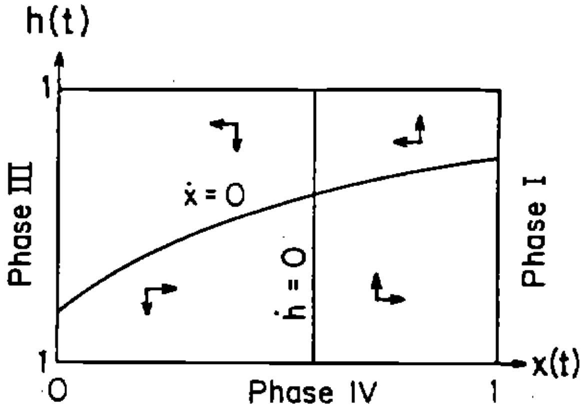

Figure 8
Alternative phase diagram in (x, h)-space

The <u>first-order</u> <u>conditions</u> for an interior maximum are

(14') $$\lambda + g'(x)\psi = 0$$

(15') $$v'(1-h) = \lambda x + g(x)\psi \ .$$

Since the Hamiltonian is <u>strictly</u> concave in (x,h) for given (λ,ψ) we know that (14')-(15') define x and h as <u>unique</u> functions of λ and ψ.[1/] Given these, the two differential equations in $\dot{\lambda}$ and $\dot{\psi}$ can be transformed into a pair of differential equations in h and x. The results are:[2/]

(26) $$\frac{-v''(\ell)}{v'(\ell)}\dot{h} = \rho - \frac{r+\delta}{1+\eta(x)}$$

(27) $$\frac{g''(x)}{g'(x)}\dot{x} = r + \delta - ahx\frac{1+\eta(x)}{\eta(x)}$$

where

$$\eta(x) \equiv \frac{-x\ g'(x)}{g(x)}$$

is the (sign-corrected) elasticity of the g(x) function. This important function has the following properties (as the reader may verify by direct computation):

$$\eta(0) = \eta'(0) = 0$$

(28) $$\eta(1) = \eta'(1) = +\infty$$

$$\eta(x) > 0,\ \eta'(x) > 0,\ \text{for } x > 0.$$

Using (26) and (27) it is easy to partititon Figure 8 into four regions by the $\dot{x} = 0$ and $\dot{h} = 0$ loci. An $\dot{h}=0$ locus only exists for $\gamma>0$ and there it has the simple form:

(29) $$\eta(x) = \gamma/\rho \ ,$$

1/ Uniqueness of λ and ψ for given x and h is obvious since (14') and (15') are linear in λ and ψ.

2/ These equations are derived in the appendix.

which, by (28), defines a unique value of x. Call this $\hat{x}$. Using (27), the end points of the $\dot{x} = 0$ locus,

$$(30) \qquad r + \delta = ahx\left(\frac{1 + \eta(x)}{\eta(x)}\right) ,$$

can easily be determined. When x = 1, h = (r+δ)/a, and when x = 0, $h = -g'(0)\, \frac{r+\delta}{a} < \frac{r+\delta}{a}$. It can also be shown that the locus is upward sloping.[1]

It is clear from Figure 8 then, that h can be no less than (r+δ)/a at the start of Phase II.[2] But does the path proceed smoothly from x = 1 to x = 0 with $\dot{x} < 0$ everywhere, as in Figure 9a, or can "mini cycles" arise, as in Figure 9b? We can prove that trajectories like Figure 9b can never be part of an optimal path because (26) and (27) define $\dot{x}$ and $\dot{h}$ as functions of x and h only, i.e., because $\dot{x}$ and $\dot{h}$ do not depend directly on either the state variables or the costate variables. It therefore follows that $(\dot{x},\dot{h})$ must be the same whenever (x,h) are. But a path like that in Figure 9b would have to cross itself -- include two distinct points in time with the same (x,h) and different $(\dot{x},\dot{h})$ -- which is impossible. This is an important result, since it shows that

---

[1]Proof: Using the definition of η(x), and differentiating (30) yields

$$0 = ax\,\frac{1+\eta(x)}{\eta(x)}\,\frac{dh}{dx} + ah\,\frac{-g(x)^2 + g(x)g''(x)}{g'(x)^2}.$$

Since the term in square brackets is negative, we have

$$\frac{dh}{dx}\Big|_{\dot{x}=0} > 0.$$

[2]We must assume $a > r + \delta$ if there is to be any training at all.

<u>within</u> <u>Phase</u> <u>II</u>, x(t) is monotonically declining -- a property which holds in the simple Ben-Pofath [1967] model, and which is vital if human capital theory is to account for the gross facts. It also implies that, <u>within</u> <u>Phase</u> <u>II</u>, labor supply rises to a single peak and then declines -- a pattern which Stafford and Stephan [1973] took great pains to establish the <u>possibility</u> of. In our model, it is the <u>only</u> pattern which could ever be optimal (so long as $\gamma > 0$).

We can also use equation (26) to determine the concavity of h(t). Letting $R(h) \equiv -v''(\ell)/v'(\ell)$ as before, its time derivative is:

$$(31) \qquad R(h)\ddot{h} = \frac{(r+\delta)\eta'(x)}{\{1+\eta(x)\}^2}\dot{x} - (\dot{h})^2 R'(h).$$

Since the first term on the righthand side is negative, $\ddot{h} < 0$ for utility functions with $R'(h) \gtreqq 0$.[1/]

The behavior of potential wages, K(t), can be displayed conveniently on the (x,h) diagram. Since $\dot{K}/K = ahx - \delta$, the $\dot{K} = 0$ locus (which exists only if $\delta > 0$) is the rectangular hyberbola, $hx = \delta/a$, shown in Figure 10. By showing that the intersection of the $\dot{x} = 0$ and $\dot{h} = 0$ loci (labelled point A in Figure 10) lies above this hyperbola, we will establish that the peak in labor supply precedes the peak in human capital. Point A is defined by $\hat{x}$ and the $\hat{h}$ which satisfies:

$$r + \delta = a\hat{h}\hat{x}\left(\frac{1 + \eta(\hat{x})}{\eta(\hat{x})}\right) .$$

[1/] When $R'(h) < 0$, (31) is ambiguous <u>a priori</u>. However, we already know that there is a unique maximum (where $\dot{h} = 0$), and it is clear that $\ddot{h} < 0$ in the neighborhood of this maximum.

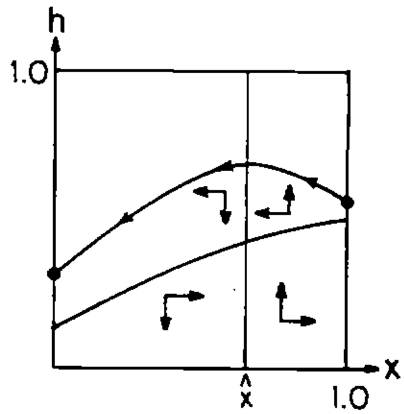

Figure 9a
A normal Phase II

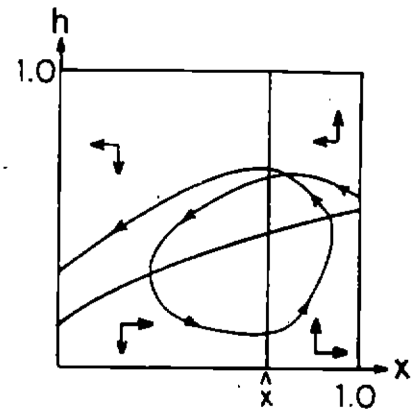

Figure 9b
"Mini-cycle" in Phase II

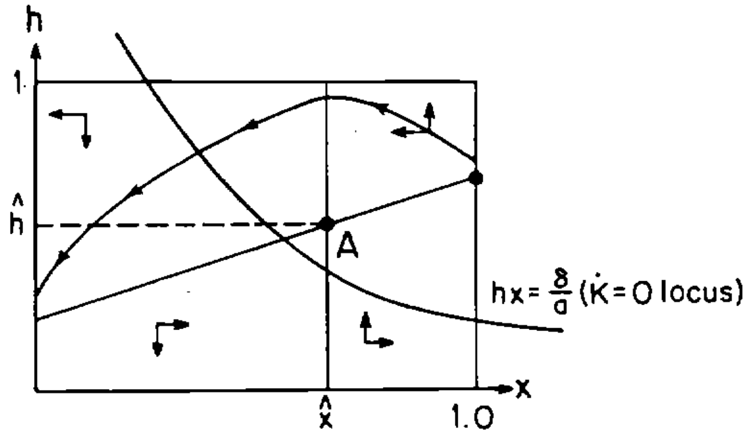

Figure 10
Behavior of potential wages

But $\eta(x) = \gamma/\rho$, so:

$$\hat{h}\,\hat{x} = \frac{\gamma}{a} = \frac{r-\rho}{a} + \frac{\delta}{a} > \frac{\delta}{a} \text{ when } r > \rho.$$

So $r \geq \rho$ is sufficient, though certainly not necessary, for the peak in labor supply to precede the peak in human capital.

What about the peak in observed wages, $W(t) = K(t)g(x(t))$? Since:

$$\dot{W}(t) = Kg'(x)\dot{x} + g(x)\dot{K} ,$$

and since $g'(x)$ and $\dot{x}$ are negative while $g(x)$ is positive, the peak in observed wages (if any) must follow the peak in potential wages (if any). Of course, if $\delta = 0$, both actual and potential wages reach their peak at the Phase II-Phase III switch point, and are level thereafter. Figure 11 below depicts the time profiles of labor supply, human capital, and observed wage rates in Phase II for $\delta > 0$. The peak in observed earnings, $h(t)W(t)$, must come between the peak in h and the peak in W, though its relation to the peak in K is unclear. If the $\delta = 0$ case is a good benchmark, earnings peak earlier than human capital.

Figure 11 shows wages and hours of work rising together at first (when wages are low), then moving in opposite directions (when wages are higher), and finally falling together (when wages are again low). Once again we stress that this has nothing whatsoever to do with income and substitution effects -- although cross-sectional studies over diverse age cohorts might mistakenly identify this phenomenon as a backward bending labor supply function.

One final observation. The rate of growth of human capital,

$$\frac{\dot{K}}{K} = ah(t)x(t) - \delta ,$$

might rise at first in Phase II, but must fall for most of Phase II. It will start out rising if the slope of the trajectory, $\dot{h}/\dot{x}$, exceeds the slope of the rectangular hyberbola $hx = h_1$, where h_1 is the labor supply at the start of Phase II.[1/]

6.4 The Complete Life Cycle

Now that we have analyzed each individual phase of the optimal life cycle, we are in a position to put the results together and describe the complete age profile of investment, work effort, earning capacity, wage rates and earnings.

6.4.1 Investment in Human Capital

We have seen that $x(t)$ is constant at unity during Phase I, declines monotonically during Phase II, and is constant at zero during Phase III. (It is not a meaningful concept in Phase IV.) To show that $x(t)$ looks as portrayed in Figure 12, we need only establish that it is continuous. Continuity within Phase II follows immediately from the fact that both λ (t) and ψ (t) are continuous functions of time and that ψ (t) is strictly positive, since:

$$(15') \qquad -g(x) = \frac{\lambda}{\psi} ,$$

[1/]The precise mathematical condition is a complex one involving the first and second derivatives of both the $v(\ell)$ and $g(x)$ functions, and does not seem worth writing down. Note that the product xh is what other writers have called "investment time". Investment time thus defined may rise at first in Phase II, but must fall thereafter.

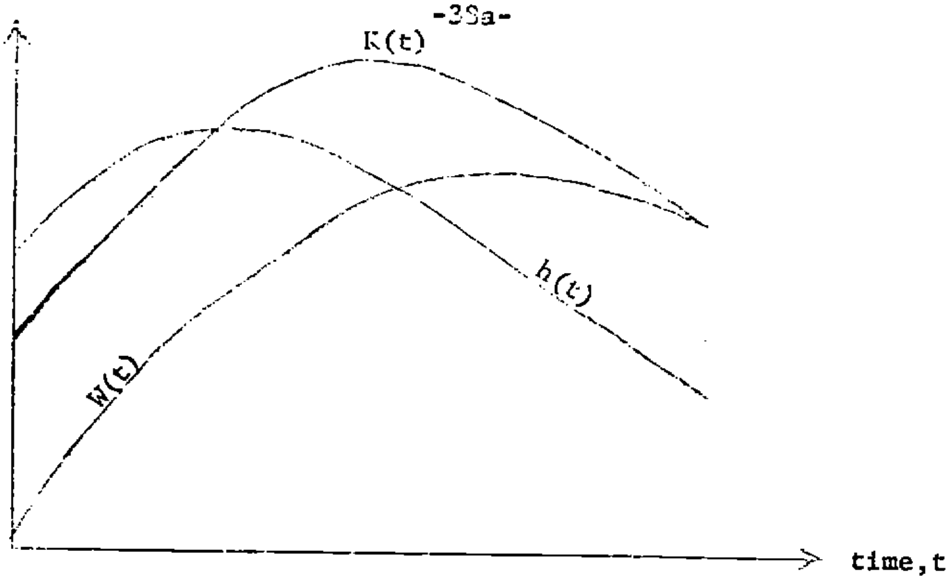

Figure 11
Time profiles within Phase II

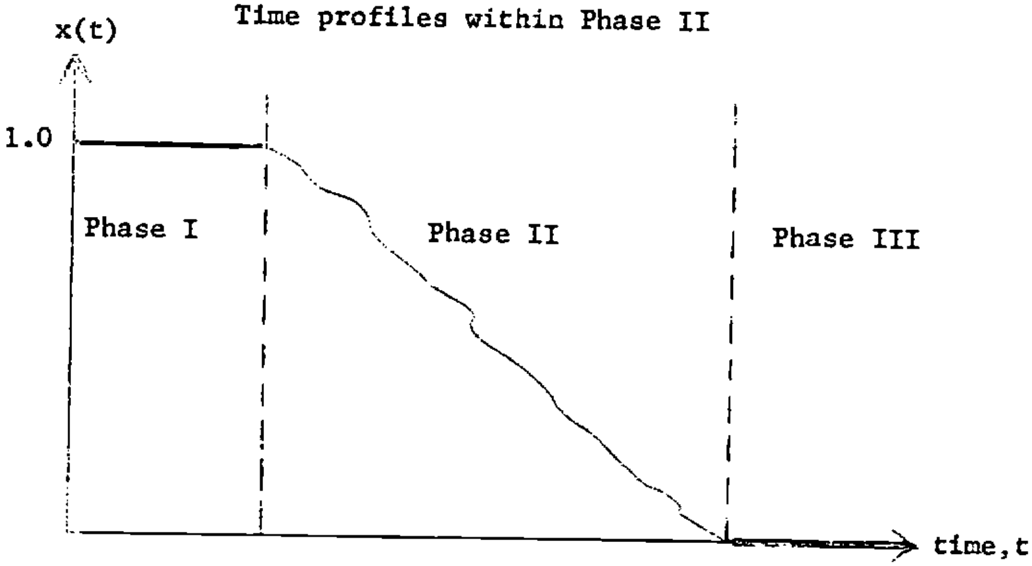

Figure 12
Time profile of human investment

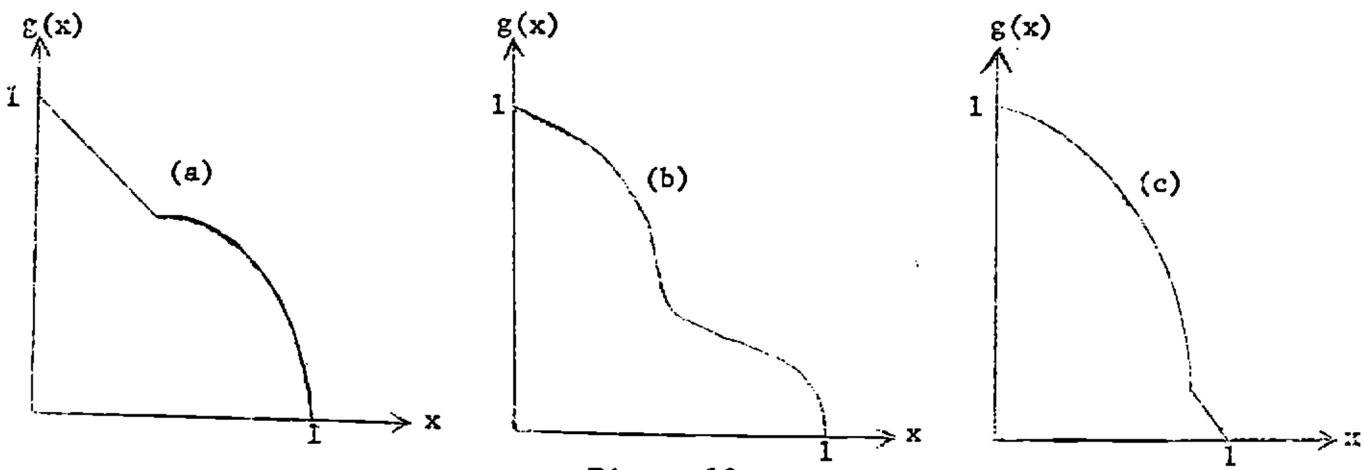

Figure 13
Alternative g(x) functions giving discontinuities in $\gamma(t)$ at II-III switch point (a), within phase II (b), or at I-II switch point (c)

and $-g'(x)$ is monotonically increasing. Continuity across the I-II and II-III switch points follows from:

$$\lim_{\frac{\lambda}{\psi} \to -g'(1)} x(t) = 1, \quad \lim_{\frac{\lambda}{\psi} \to -g'(0)} x(t) = 0,$$

which are immediate implications of the continuity and monotonicity of $g'(x)$. Examples of $g(x)$ functions which give discontinuous $x(t)$ either within Phase II or at a switch point are provided in Figures 13.[1/] Nothing can be said in general about the concavity of $x(t)$ within Phase II.

6.4.2 Labor Supply

So far we have shown that $h(t)$ is rising in Phase I, continues rising to a peak and then declines in Phase II, and declines throughout Phase III. It remains to show that $h(t)$ and $\dot{h}(t)$ are both continuous functions of time.

Continuity within each phase is trivial to establish. In Phase I, $v'(1-h) = \lambda(t)$, λ and v' are continuous, and $v'' < 0$. In Phase II, $v'(1-h) = \psi(t)$, and the exact same argument works. Finally, in Phase II, we have:

$$v'(1-h) = \lambda x + \psi g(x) \quad ,$$

and given our assumptions about $g(x)$, both x and $g(x)$ are continuous functions of time.[2/]

[1/] Mincer's empirical work, which finds a discontinuity in $x(t)$ at the end of schooling, suggests that Figure 13c is closest to the truth.

[2/] This argument makes it appear that continuity of $h(t)$ requires continuity of $x(t)$. This is not so. x and $g(x)$ will display discontinuous behavior at the same points in time. We only require that $\lambda x + \psi g(x)$ be continuous. For example, optimal $x(t)$ is a discontinuous function of time when $g(x)=1-x$, but in this case $\lambda x + \psi g(x) = \psi + x(\lambda - \psi)$, which is continuous.

To establish continuity at the switch points, we must prove:

(i) $\lim_{x \to 1} \lambda x + \psi g(x) = \lambda$ (I-II switch)

(ii) $\lim_{x \to 0} \lambda x + \psi g(x) = \psi$ (II-III switch)

(iii) $\lim_{t \to R} h(t) = 0$, where R is the age of retirement (III-IV) switch)

Conditions (i) and (ii) follow by inspection given only that x and g(x) are continuous with $g(0) = 1$ and $g(1) = 0$.[1/] To prove (iii), we use the condition:

$$\psi(R) = v'(1 - h(R))$$

to rewrite it:

$$\lim_{t \to R} \psi(R) = \xi. \quad \text{(Since } h(R) = 1\text{)}.$$

Referring back to the phase diagram in Figure 6, it is apparent that this must be true since $\psi(t)$ is continuous.

What about the continuity of $\dot{h}(t)$? Since $v'(\ell)$, $v''(\ell)$ and $\eta(x)$ are well-behaved functions, and since we have just seen that x(t) is continuous, (26) immediately implies that $\dot{h}(t)$ is continuous within Phase II. Proceeding analogously, to prove continuity across switch points, we must show that:

(iv) $\lim_{x \to 1} \frac{r+\delta}{1+\eta(x)} = \cdot$ [2/] (I-II switch)

(v) $\lim_{x \to 0} \{\rho - \frac{r+\delta}{1+\eta(x)}\} = -\gamma$ [3/] (II-III switch)

Conditions (iv) and (v) follow immediately by noting that $\lim_{x \to 1} \eta(x) = +\infty$ and $\lim_{x \to 0} \eta(x) = 0$. However, $\dot{h}(t)$ is discontinuous across the III-IV switch since $\lim_{t \to R} \dot{h}(t) = \frac{-\gamma\xi}{v''(1)} < 0$.

[1/] Again, we do not believe that continuity of x is needed. It is just that a different proof would be required.

[2/] This condition comes from comparing (26) which holds in Phase II, with $(-v''/v')\dot{h} = \rho$, which holds in Phase I.

[3/] This condition comes from comparing (26) with (25), which holds in Phase III.

Collecting all these results, Figure 14 depicts the life-cycle labor supply profile in the leading case: $\gamma \equiv r-\rho+\delta > 0$, $R'(h) > 0$. In a word, the profile is smooth (except at retirement), single-peaked, and has no inflection points.

6.4.3 Human Capital

During Phase I, K(t) rises at the increasing rate $\frac{\dot{K}(t)}{K(t)} =$ $ah(t) - \delta$. It may continue to rise at an increasing rate at the start of Phase II where:

$$\frac{\dot{K}(t)}{K(t)} = ah(t)x(t) - \delta \quad ,$$

but soon will rise at a decreasing rate, and eventually will decline if and only if $\delta > 0$. Finally, K(t) falls at the steady rate δ in Phase III -- the only phase where it is observable. (See Figure 15.)

6.4.4 Actual Wage Rates

Observed wages, of course, are zero until the completion of schooling and correspond to K(t) in Phase III. During Phase II we have

$$\log W(t) = \log K(t) + \log g(x)$$

so that the slope of the logarithmic wage profile is:

$$\frac{\dot{W}(t)}{W(t)} = \frac{\dot{K}(t)}{K(t)} - \eta(x)\frac{\dot{x}}{x} \quad .$$

This will always exceed $\dot{K}/K$. Further, since as $x \to 0$ both $\dot{x}$ and $\frac{\eta(x)}{x}$ approach finite limits, the time path of log W(t) will exhibit a kink at the Phase II-Phase III switch point. (See Figure 16.) Since nothing is known about $\ddot{x}$, nothing is known about its concavity.

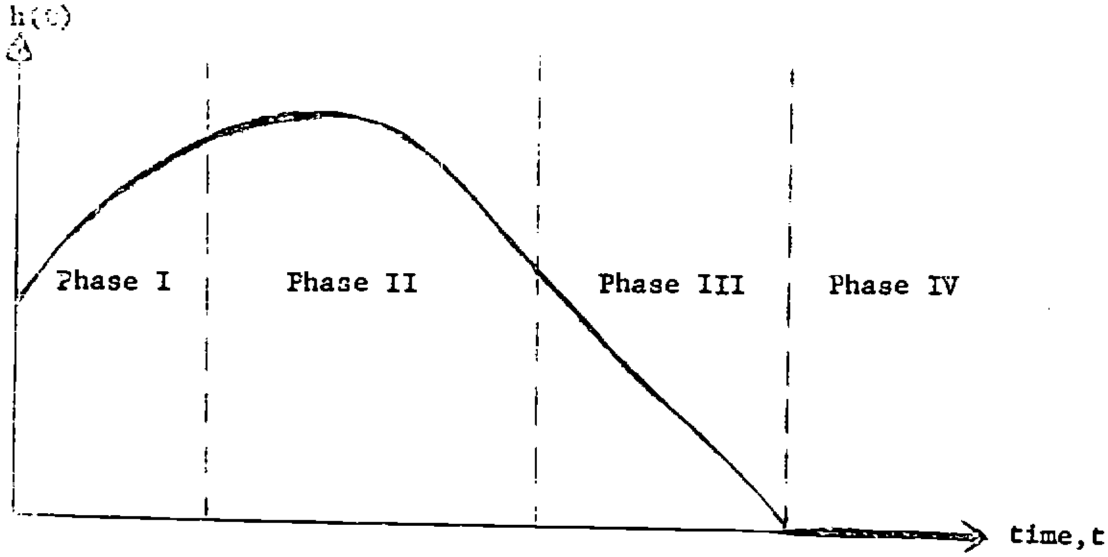

Figure 14
Labor supply over the life cycle

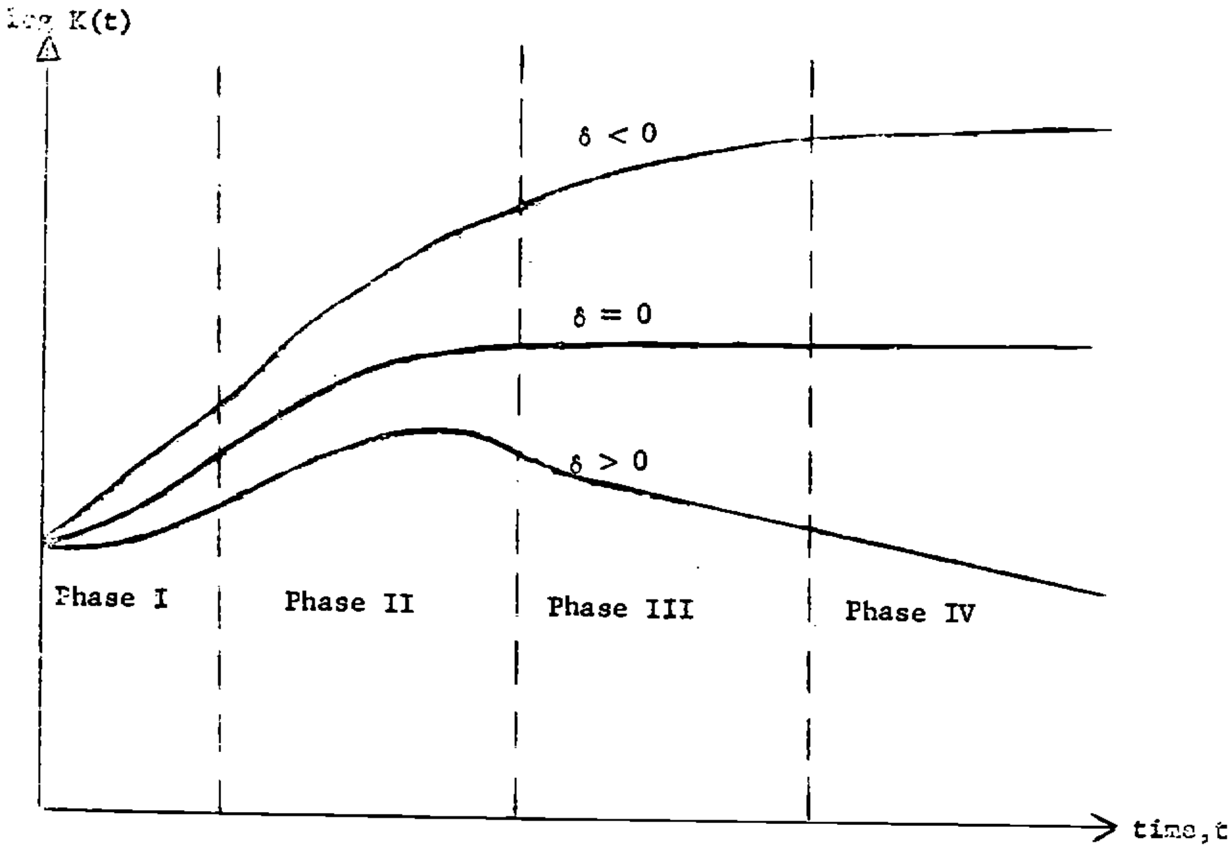

Figure 15
Human capital over the life cycle

6.4.5 Earnings

The log of earnings is given by:

$$\log Y(t) = \log h(t) + \log h(t).$$

so:

$$\frac{\dot{Y}}{Y} = \frac{\dot{h}}{h} + \frac{\dot{K}}{K} - \eta(x)\frac{\dot{x}}{x}.$$

During Phase I, of course, $Y(t) = 0$. At the start of Phase II, all three terms are positive, so $Y(t)$ is certainly growing. Then $h(t)$ peaks and begins to decline. If (and only if) $\delta > 0$, $K(t)$ also peaks in Phase II. But $-\eta(x)\frac{\dot{x}}{x} > 0$ throughout Phase II, so there is no particular reason to think that $Y(t)$ ever declines in Phase II (though it might). Of course, in Phase III, $\dot{h} < 0$ and $\frac{\dot{K}}{K} = -\delta$, so Y falls unless δ is substantially negative. Figure 17 depicts several possible logarithmic earnings profiles.

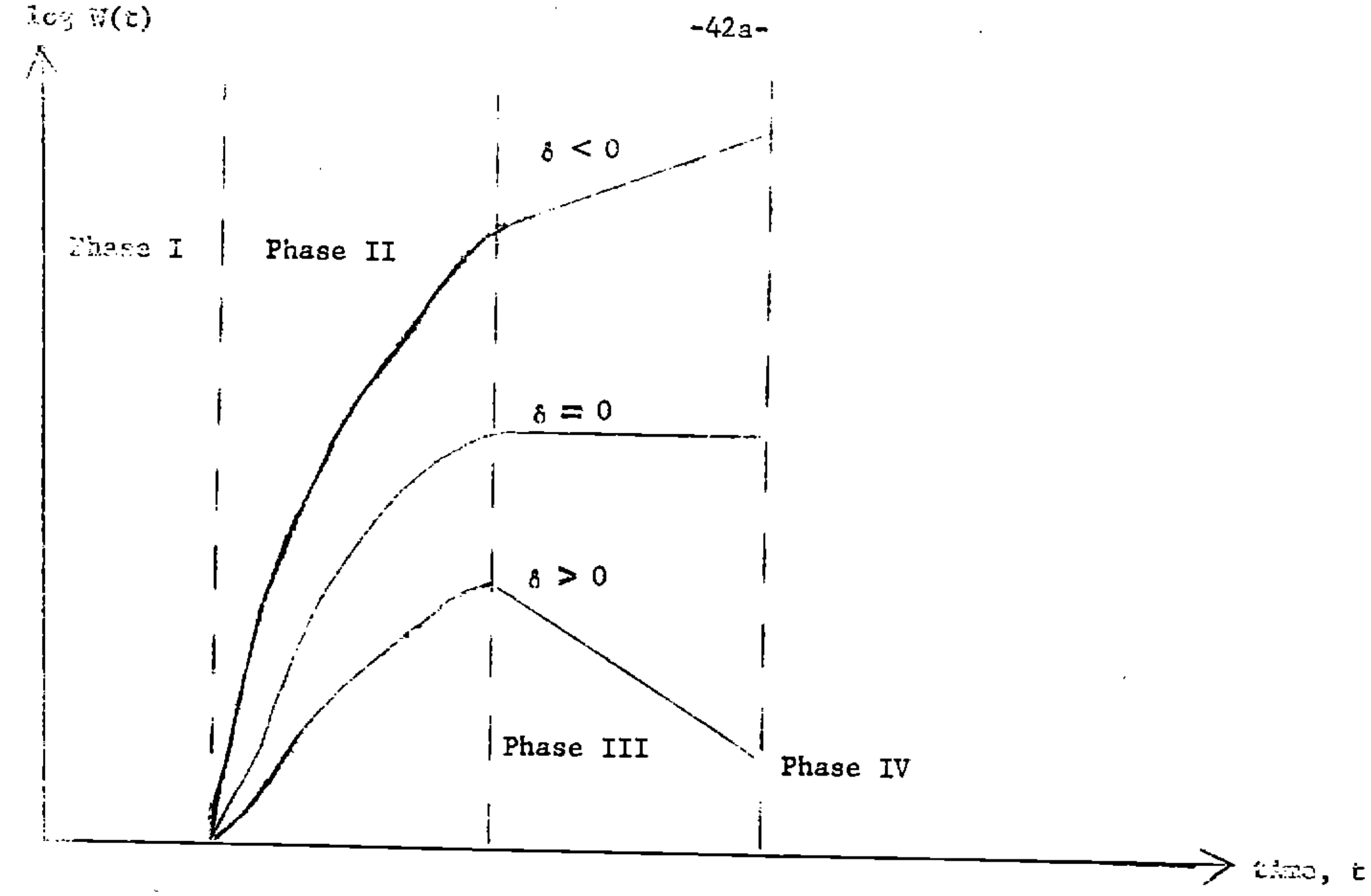

Figure 16
Wage rates over the life cycle

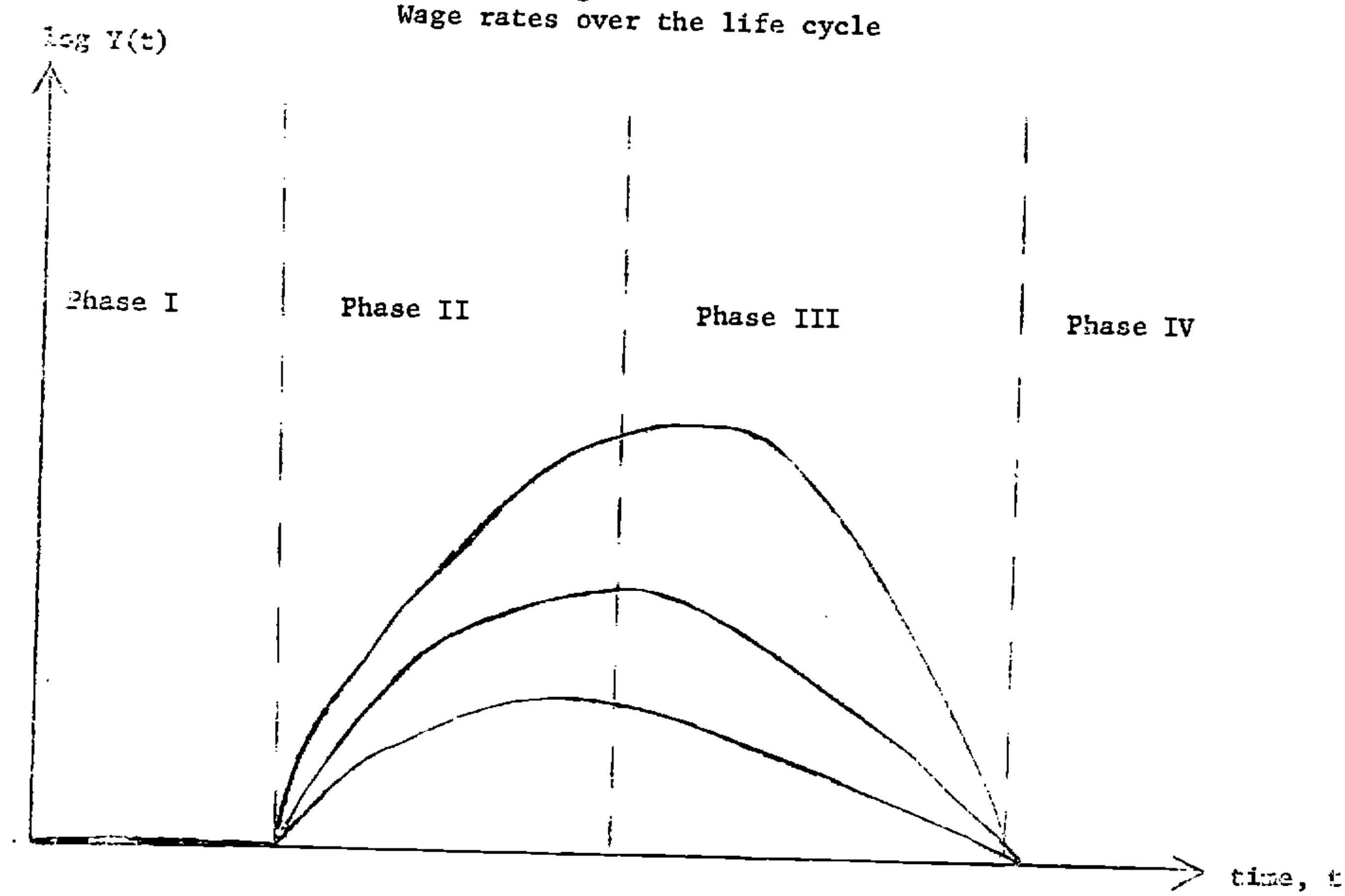

Figure 17
Earnings over the life cycle

VII. The Question of Cycling

We now address ourselves to two important questions which have thus far been avoided:

(a) When cycling arises, what broad characterization of the optimal path can be made?

(b) What meaningful conditions can be derived which exclude the possibility of cycling?

In Section 6, we proved that the optimal trajectory can never cross itself.[1/] This result enables us to answer the first question, since it implies that cycles must be "expanding" as in Figure 18, rather than "contracting". Thus a cycling path can be broken down into several "quasi-life-cycles", each beginning with a period of schooling (with the possible exception of the first). Since the contours of constant work effort look like blow-ups of the border of the retirement region (which is the special case $h = 0$), with higher contours connoting higher h, we see that the individual works harder during his second "quasi cycle" than in his first. Also, he spends more time in Phase II. So, in a program with cycling, the time profiles of $x(t)$ and $h(t)$ might look something like Figures 19 and 20 respectively.

We can develop strong conditions which rule out cycles by considering the minimum length of time required to complete a "quasi cycle". If this time exceeds the available life, T, then cycles are impossible. In particular, we focus on the last

---

1/ Actually, we only proved that it could never cross itself in Phase II. It is trivial to prove this for the other phases as well.

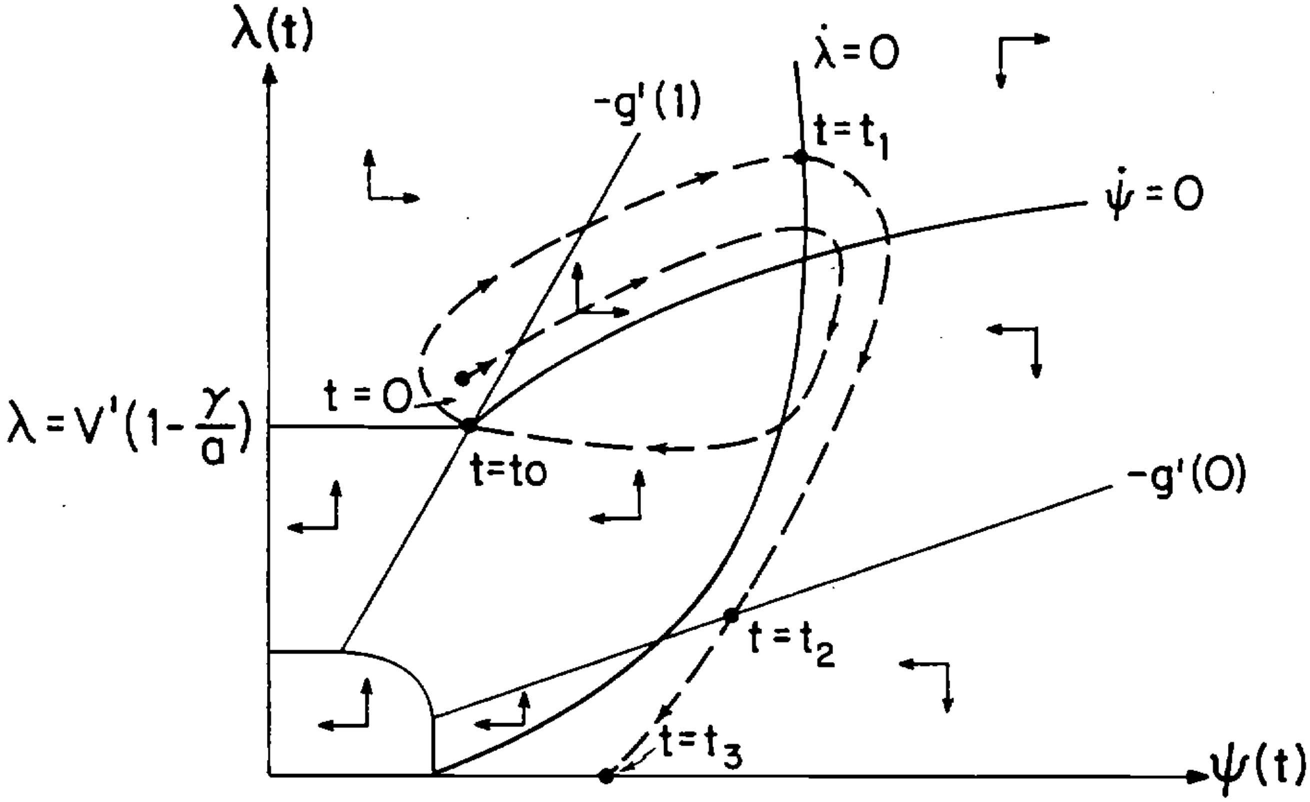

Figure 18

A possible cycle when $\gamma > 0$

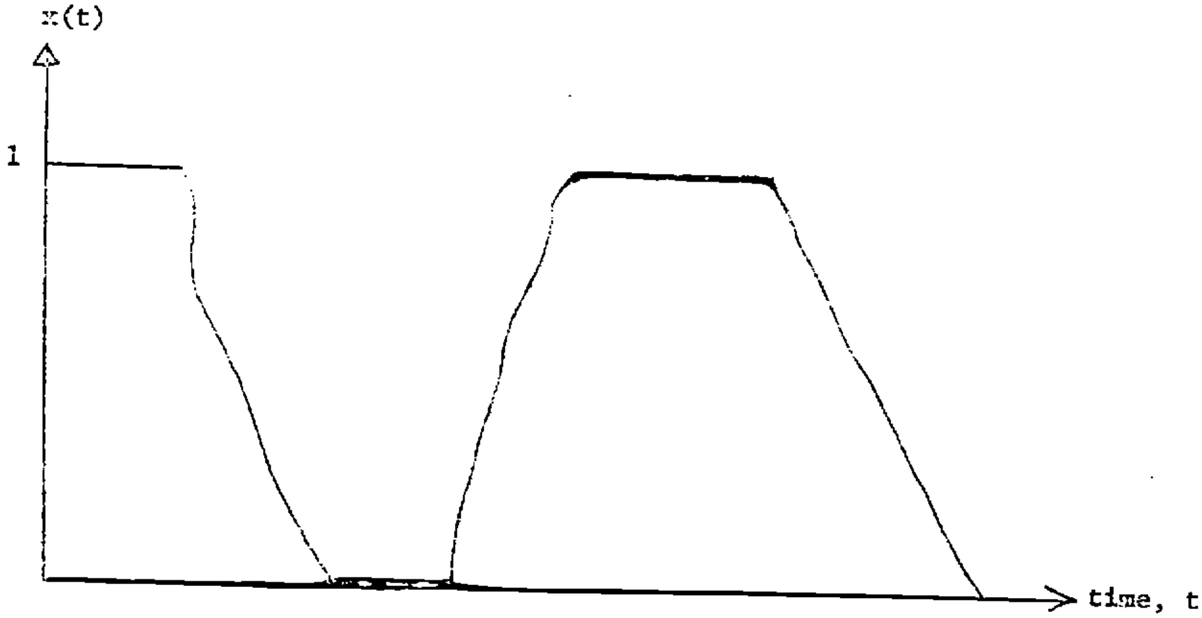

Figure 19
Human investment with cycling

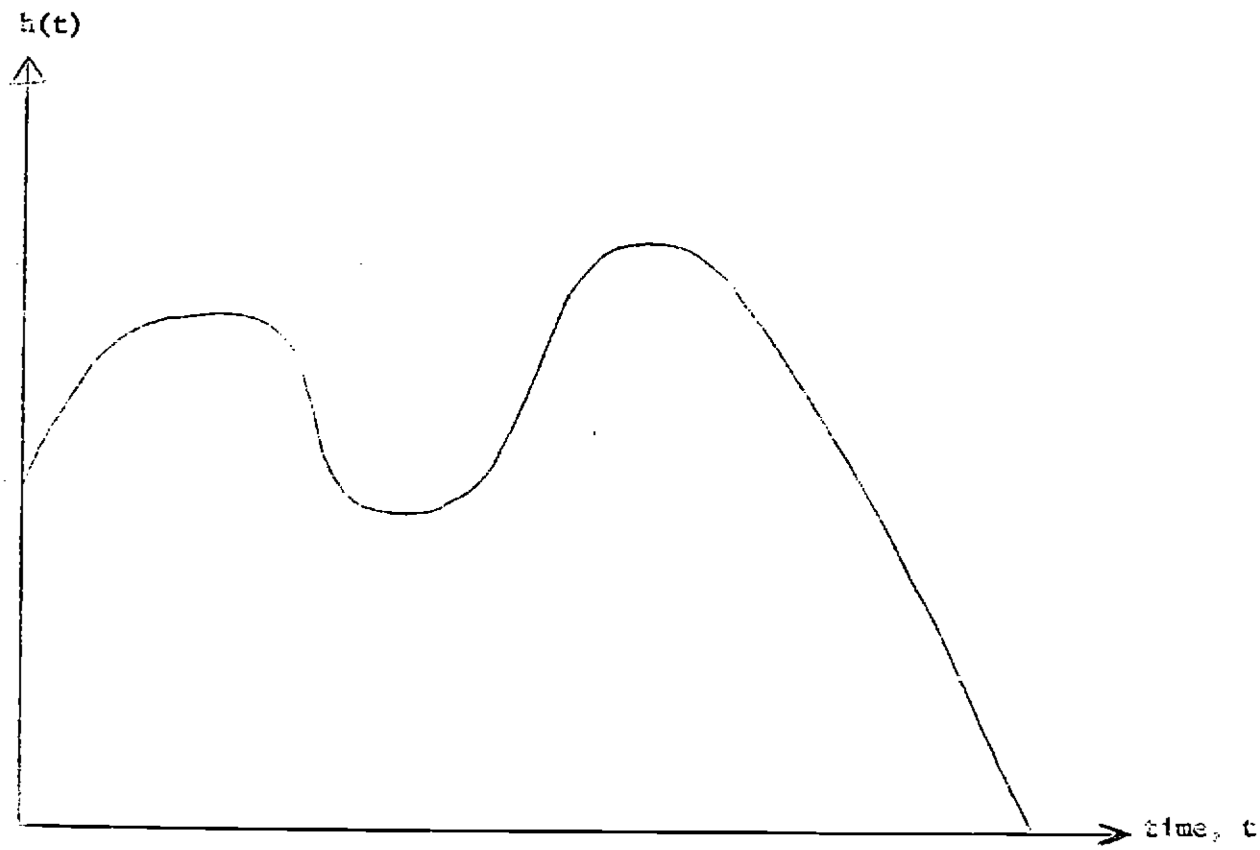

Figure 20
Labor supply with cycling

quasi cycle, and let:

t_0 = age of starting last schooling period

t_1 = age of starting last OJT period

t_2 = age of starting last work period

t_3 = age of ending last work period (possibly $t_3 = T$).

These four points on the optimal trajectory are indicated in Figure 18.

Consider first the length of the schooling period, $t_1 - t_0$. Since we are considering a cycling path, we know that:[1/]

$$(32) \qquad h(t_0) \leq \gamma/a = \frac{r+\delta-\rho}{a} .$$

(See Figure 18.) Furthermore, we know from Figure 8 that:

$$(33) \qquad h(t_1) \geq (r+\delta)/a .$$

Finally, we know that optimal behavior implies that $v'(1-h)$ grows at the exponential rate ρ in Phase I (see equation (20.1)). Thus we can derive a minimal time required to pass through Phase I which will depend on the rate of impatience, the elasticity of the marginal utility function, and the parameters a, r and δ. Specifically, since

$$\lambda(t_0) \leq v'(1 - \frac{\gamma}{a}) \qquad \text{by (32)},$$

and

$$\lambda(t_1) \geq v'(1 - \frac{r+\delta}{a}) \qquad \text{by (33)},$$

we have:

$$(34') \qquad e^{\rho(t_1 - t_0)} = \frac{v'(\frac{a-r-\delta}{a})}{v'(\frac{a+\rho-r-\delta}{a})} .$$

[1/] Note that no such statement can be made about the schooling period in a trajectory without cycling. Thus the bound to be derived will not apply to "normal" life cycles.

Equation (34') places a lower bound on the interval t_1-t_0 which increases as ρ falls,[1/] specifically:

$$S^* = \frac{1}{\rho} \log \frac{v'(\frac{a-r-\delta}{a})}{v'(\frac{a+\rho-r-\delta}{a})} .$$

It is also intuitively clear that there must be some minimal period of work during which the benefits from schooling are realized. This is the idea of our lower bounds on the lengths of Phases II and III, which depend basically on the tradeoffs embodied in the earnings-investment frontier. To derive these bounds, recall our earlier notation (p.): $\theta(t) = p(t)/\mu(t)$.

[1/] This last statement holds under the assumption, introduced in Section 6, that $\frac{d}{d\ell}\frac{v''(\ell)}{v'(\ell)} > 0$. To prove it, let S* be the root of (34'), viz:

$$e^{\rho S^*} v'(1 - \frac{r+\delta}{a} + \frac{\rho}{a}) = v'(1 - \frac{r+\delta}{a}).$$

Since the righthand side of this expression is independent of ρ, $\frac{dS^*}{d\rho} < 0$ follows if we can show:

$$\frac{\partial}{\partial\rho}\Big|_{S^*\text{constant}} \{e^{\rho S^*} v'(1 - \frac{r+\delta}{a} + \frac{\rho}{a})\} > 0 .$$

Now, adopting the shorthand notation $\beta \equiv 1 - (r+\delta)/a$, this partial is:

$$S^* \, e^{\rho S^*} v'(\beta + \frac{\rho}{a}) + \frac{1}{a} e^{\rho S^*} v''(\beta + \frac{\rho}{a}) > 0$$

or

$$S^* > \frac{1}{a} \frac{-v''(\beta + \frac{\rho}{a})}{v'(\beta + \frac{\rho}{a})}.$$

But by (34) this says:

$$\log \frac{v'(\beta)}{v'(\beta + \frac{\rho}{a})} > - \frac{\rho}{a} \frac{v''(\beta + \frac{\rho}{a})}{v'(\beta + \frac{\rho}{a})} .$$

To prove this, define the function: $f(\ell) \equiv \log v'(\ell)$. Then $f'(\ell) = \frac{v''(\ell)}{v'(\ell)} < 0$ and $f''(\ell) > 0$ by our assumption. Thus $f(\ell)$ is <u>greater than</u> its first-order Taylor approximation around $\ell = \hat{\ell}$ viz.:

$$f(\ell) > f(\hat{\ell}) + f'(\hat{\ell})(\ell - \hat{\ell}).$$

Letting $\hat{\ell} = \beta + \frac{\rho}{a}$, $\ell = \beta$ as a special case proves the desired inequality. Q.E.D.

In this notation, t_1, t_2 and t_3 are defined by (see Figure 18):

(35) $$a\theta(t_1) = -g'(1)$$

(36) $$a\theta(t_2) = -g'(0)$$

(37) $$\theta(t_3) = 0 \quad .$$

It is easy to show that:

(38) $$\dot{\theta} = (r+\delta)\theta - h[g(x) - xg'(x)] \text{ in Phase II}$$

(39) $$\dot{\theta} = (r+\delta)\theta - h \text{ in Phase III} .$$

First consider Phase II, where the solution to (38) is:

$$\theta(t)e^{-(r+\delta)(t-t_1)} = \theta(t_1) - \int_{t_1}^{t} e^{-(r+\delta)(\tau-t_1)} h[g(x)-xg'(x)]d\tau .$$

Setting $t = t_2$ and using (35) and (36), we have:

$$-g'(0)e^{-(r+\delta)(t_2-t_1)} = -g'(1) - a \int_{t_1}^{t_2} e^{-(r+\delta)(\tau-t_1)} h[g(x)-xg'(x)]d\tau .$$

Now, replacing h(t) in the integral by 1.0 and replacing g(x)-xg'(x) by -g'(1) makes the integral strictly larger, so we have the inequality:

$$-g'(1)a \int_{t_1}^{t_2} e^{-(r+\delta)(\tau-t_1)} > -g'(1) + g'(0)e^{-(r+\delta)(t_2-t_1)} ,$$

or

$$a \, \frac{1-e^{-(r+\delta)(t_2-t_1)}}{r+\delta} > 1 - \frac{g'(0)}{g'(1)} e^{-(r+\delta)(t_2-t_1)} .$$

Our lower bound on the length of Phase II, call it J^*, therefore satisfies:[1/]

$$J^* = \frac{1}{r+\delta} \log \frac{a - \frac{g'(0)}{g'(1)}(r+\delta)}{a-r-\delta} .$$

The bound depends on the parameters a, r and δ, and also on the degree of concavity of the earnings-investment frontier.[2/] Tastes

---

[1/] It should be clear that the bound applies equally well to the OJT period of a normal life cycle, since the only things assumed were that Phase II is preceded by Phase I and precedes Phase III.

[2/] That is, on how much smaller -g'(0) is than -g'(1).

are not involved since we obtained the bound by assuming the individual works as hard as he can.

Almost the same reasoning can be used to place a lower bound, W^*, on the length of the last Phase III. Solving (39) explicitly gives:

$$\theta(t)e^{-(r+\delta)(t-t_2)} = \theta(t_2) - \int_{t_2}^{t} e^{-(r+\delta)(\tau-t_2)}h(\tau)d\tau.$$

Letting $t = t_3$ and using (36) and (37), we obtain:

$$-g'(0) = a\int_{t_2}^{t_3} e^{-(r+\delta)(\tau-t_2)}h(\tau)d\tau < a\int_{t_2}^{t_3} e^{-(r+\delta)(\tau-t_2)},$$

since $h(\tau) < 1$. Thus:

$$-g'(0) < \frac{a}{r+\delta}\left[1-e^{-(r+\delta)(t_3-t_2)}\right],$$

or

$$W^* = \frac{1}{r+\delta} \log \frac{a}{a + g'(0)(r+\delta)}.$$

Again the bound depends on a, r, δ, and the $g(x)$ function.

Putting these results together, we find that there are bounds on the minimum lengths of Phases II and III which are the <u>same</u> for every individual, and there is a bound for the minimal schooling period which depends on ρ. Let ρ^* be the value of ρ satisfying:

$$S^*(\rho^*) = T - J^* - W^*.$$

Then cycling is certainly impossible for persons with $\rho \leq \rho^*$. Since we know that cycling is only a problem when $\rho < r + \delta$, a sufficient (but far from necessary) condition to rule out cycling is: $r + \delta \leq \rho^*$. The reader should note that the condition $\rho \leq \rho^*$ rules out early retirement in the $\gamma > 0$ case as well, since the bounds $S^*(\rho^*)$, J^*, and W^* apply here as well. In a word, we will never get cycles if

<u>either</u> ρ is "large" (i.e. $\rho \geq r + \delta$) or ρ is "small" (i.e., $\rho \leq \rho^*$). If $\rho^* < r + \delta$, there will be an intermediate range in which cycling is possible.

It may be interesting to work out a numerical example of these bounds. Suppose a, which would be the rate of return on human capital in the case of an infinite lifetime, is 8%, while $r + \delta$ is 6%. A simple form for the earnings-investment frontier is the quadratic: $g(x) = 1 - \frac{1}{2}x - \frac{1}{2}x^2$. Given these choices, (40) and (41) can be used to compute: $J^* = 18.3$, $W^* = 7.8$. From (34) it is clear that a facile choice of utility function is:

$$v'(\ell) = e^{-\ell}.$$

For a rate of impatience of $\rho = .02$, (34) gives $S^* = 12.5$ as the minimal schooling period. Adding these, we find that a new cycle cannot start unless there are more than 38.6 years remaining.[1/]

[1/] The reader is reminded that these are all weak lower bounds. The actual amount of time taken by the last quasi cycle is thus strictly greater (and probably considerably greater) than 38.6 years.

VIII. In Conclusion

We have presented a life-cycle model of the behavior of a utility-maximizing individual free to allocate his daily time budget among leisure, work and education. The main substantive conclusions are listed in Section 1, and depicted in Figures 12-17. They need not be rehashed here.

Several generalizations of the analysis immediately come to mind. More general forms of the human capital production function could be tried, as could a non-separable utility function. Indeed, the list of arguments of the utility function could be expanded to include human capital, or, indeed, nonhuman capital. More important than all these, we imagine, would be allowance for the capital-market imperfections that severely constrain the choices of those poorly endowed with financial wealth.

Still, so long as we interpret the results as a benchmark around which there will surely be deviations (some systematic, some random), there are a number of interesting uses for the model.

First, it may be possible to do the usual kinds of comparative-dynamic exercises. How would an increase in initial financial wealth alter the optimal plan? What about an increase in initial human wealth? How sensitive is the optimal plan to the rate of impatience -- a taste parameter that presumably differs across people.[1/] The reader can no doubt think of many similar questions.

[1/] This question is addressed, in the context of a far simpler model, by Beach, Maital and Maital [1973].

While we suspect that most of these can only be answered under specific functional forms for $v(\ell)$ and $g(x)$, some weak results may hold in greater generality.

Second, the model can provide the micro-foundation for a simulation model of the income distribution, along the lines of Blinder [1974]. The chief conclusion of that work is that the distribution of wage rates, taken there to be exogenous, is the principal contributor to income inequality. The present model can generate that distribution endogenously, given assumed distributions of tastes and endowments, and thus can fill what is perhaps the major gap in the positive theory of income distribution.

Third, the long-run incidence of various taxes in a world with human-capital accumulation is virtually unexplored territory. It seems feasible to incorporate some simple taxes -- such as a linear income or wage tax -- into the model and examine, either analytically or through simulation, the effects of these programs of the acquisition of human capital. It could be that human-investment responds to taxation more substantially than do hours of work.

In a word, there are a host of interesting and important questions which simply cannot be addressed by a life-cycle model which considers <u>either</u> labor-leisure choices <u>or</u> labor-education choices, but not both.[1] By demonstrating the feasibility of

---

[1] For example, if income maximization is the posited goal, a proportional tax (subsidy) on wages cannot possibly alter behavior.

handling both decisions together, we hope to have hastened the day when the powerful tools of life-cycle economic theory will be brought to bear on these issues.

Mathematical Appendix

The purpose of this appendix is to establish the properties of the $\dot{\psi} = 0$ and $\dot{\lambda} = 0$ lines in the region in which both x and h attain an interior solution. We wish to show that:

1. there exists a unique intersection point.
2. At the intersection the $\dot{\lambda}=0$ line cuts the $\dot{\psi} = 0$ line from below.

Uniqueness

Notice first that the necessary conditions (14)-(15) define a unique (λ,ψ) pair for any (x,h) pair. (The converse is also true.) It is therefore sufficient to show that the equations (see eqs. (22) (23) in the text),

$$\text{(A.1)} \quad \rho\lambda - a\psi g(x)h = 0$$

$$\text{(A.2)} \quad (ahx-\gamma)\psi = 0$$

provide a unique solution for h and x. Using the first order condition (15), (A.1) can be written:

$$\text{(A.3)} \quad \rho + ah\frac{g(x)}{g'(x)} = 0.$$

And, of course, since $\psi > 0$, (A.2) implies,

$$\text{(A.4)} \quad ahx - \gamma = 0.$$

Since, under our assumptions, $\frac{g(x)}{g'(x)}$ is monotonically increasing in x, a solution to (A.3)-(A.4) exists and it must be unique.

It remains to verify that it is an admissable solution, i.e. that $0<x^o<1$, $0<h^o<1$. That $x^o, h^o > 0$ is obvious. Also $x^o < 1$ since $\frac{-g'(1)}{g(1)} = \infty$. It remains to show that $h^o = \frac{\gamma}{ax^o} = \frac{\rho}{a}\frac{-g'(x^o)}{g(x^o)} < 1$.

Let us define $F(x) \equiv \frac{-xg'(x)}{g(x)}$. Clearly $F'(x) > 0$. We wish to show that $x^o > \frac{\gamma}{a}$. Since $F(x)$ is monotonically increasing, this is equivalent to the requirement $F(x^o) > F(\frac{\gamma}{a})$. But $F(x^o) = \frac{\gamma}{\rho}$ by definition. It is therefore necessary to show that:

$$\frac{\gamma}{\rho} > -\frac{\gamma}{a} \frac{g'(\frac{\gamma}{a})}{g(\frac{\gamma}{a})} \quad \text{or} \quad g(\tfrac{\gamma}{a}) > -\frac{\rho}{a} g'(\tfrac{\gamma}{a}).$$

Note, however, that $\frac{\rho}{a} = \frac{r+\delta}{a} - \frac{\gamma}{a} < 1 - \frac{\gamma}{a}$. It is therefore sufficient to show that

$$g(\tfrac{\gamma}{a}) > (1 - \tfrac{\gamma}{a})(-g'(\tfrac{\gamma}{a})).$$

But this last inequality follows directly from the assumed concavity of $g(x)$. (It states that $g(1) = 0$ is greater than the first-order Taylor expansion around $x = \gamma/a$.)

Shape of the Stationary Loci

Due to the concavity of the Hamiltonian in the controls, the two necessary conditions (14)-(15) define a unique pair (h,x) for any (λ,ψ). Let us denote these solutions by:

$$x = X(\lambda,\psi)$$

$$h = H(\lambda,\psi)$$

where

$$H_\lambda = \frac{x}{v''(1-h)} > 0 \qquad X_\lambda = -\frac{1}{\psi g''(x)} > 0$$

$$H_\psi = \frac{g(x)}{v''(1-h)} > 0 \qquad X_\psi = -\frac{g'(x)}{\psi g''(x)} < 0.$$

Thus, considered as a pair of equations in λ and ψ, (A.3)-(A.4) become:

$$0 = F^1(\lambda,\psi) \equiv \rho\lambda - a\psi g[X(\lambda,\psi)]H(\lambda,\psi)$$

$$0 = F^2(\lambda,\psi) \equiv aH(\lambda,\psi)X(\lambda,\psi) - \gamma.$$

We wish to show the Jacobian of this system evaluated at λ^o,ψ^o is positive. This implies that the $\dot{\lambda} = 0$ line cuts the $\dot{\psi} = 0$ line from below. (It also guarantees the uniqueness of the solution, if it exists.)

$$J = \begin{vmatrix} F^1_\lambda & F^1_\psi \\ F^2_\lambda & F^2_\psi \end{vmatrix} = \begin{vmatrix} \rho-a\psi(hg'(x)X_\lambda+g(x)H_\lambda) & -ahg(x)-a\psi(hg'(x)X_\psi+g(x)H_\psi) \\ a(xH_\lambda+hX_\lambda) & a(xH_\psi+hX_\psi) \end{vmatrix}$$

$$= a\rho(xH_\psi+hX_\psi) + a^2hg(x)(xH_\lambda+hX_\lambda) + a^2h\psi[g(x)-xg'(x)][H_\psi X_\lambda-H_\lambda X_\psi].$$

Owing to the signs of X_λ, X_ψ, H_λ, H_ψ, the only negative term in this sum is: $a\rho hX_\psi$. However, since $\rho = ah\frac{g(x)}{-g'(x)}$ in Phase II, we have:

$$a\rho hX_\psi = ah\,\frac{ahg(x)}{-g'(x)}\,\frac{-g'(x)}{\psi g''(x)} = \frac{(ah)^2g(x)}{\psi g''(x)} < 0.$$

But this cancels out with the term:

$$a^2h^2g(x)X_\lambda = \frac{(ah)^2g(x)}{-\psi g''(x)} > 0\ ,$$

so $J > 0$. Q.E.D.

We now prove two further properties of the loci:

(a) the $\dot{\psi}=0$ and $\dot{\lambda}=0$ loci have positive slopes at x=1 and x=0 respectively;

(b) along the $\dot{\psi}=0$ locus, $d(\lambda/\psi)/d\psi < 0$ while along the $\dot{\lambda}=0$ locus, $d(\lambda/\psi)|d\lambda > 0$.

We have

$$\frac{d\lambda}{d\psi}\Big|_{\dot{\lambda}=0} = -\frac{F^1_\psi}{F^1_\lambda} = \frac{ag(x)h+ahg'(x)X\psi+a\psi g(x)H\psi}{\rho-ahg'(x)X_\lambda-a\psi g(x)H_\lambda}$$

At the border between regions II and III x = 0, which implies $H_\lambda = 0$. Given the signs of X_ψ, H_ψ, and X_λ, it follows that at this point

$\frac{d\lambda}{d\psi}\Big|_{\dot{\lambda}=0} > 0.$

Also

$$\frac{d\lambda}{d\psi}\Big|_{\dot{\psi}=0} = -\frac{F^2_{\psi}}{F^2_{\lambda}} = -\frac{xH_{\psi} + hX_{\psi}}{xH_{\lambda} + hX_{\lambda}} .$$

At the border between regions I and II x=1, which implies $H_{\psi} = 0$. Again, given the signs of X_{ψ}, H_{ψ} and H_{λ}, it follows that $\frac{d\lambda}{d\psi}\Big|_{\dot{\psi}=0} > 0$ at this point. In order to determine the concavity properties of the $\dot{\psi}=0$ and $\dot{\lambda}=0$ lines, it is convenient to write the solution of the first order conditions (14)-(15) as function of the level of λ and the ratio $\frac{\lambda}{\psi}$. Thus:

$$x = \tilde{X}(\lambda, \tfrac{\lambda}{\psi}) \quad \text{where} \quad \tilde{X}_{\lambda} = 0 \quad \tilde{X}_{\frac{\lambda}{\psi}} > 0$$

$$h = \tilde{H}(\lambda, \tfrac{\lambda}{\psi}) \qquad\qquad \tilde{H}_{\lambda} > 0 \quad \tilde{H}_{\frac{\lambda}{\psi}} < 0.$$

Consider now the condition $\dot{\lambda} = 0$ or $\rho + \frac{ag(x)}{g'(x)}h = 0$. Since an increase in $\frac{\lambda}{\psi}$ raises x, which decreases the quantity $\frac{ag(x)}{g'(x)}$ in absolute value, an increase in λ (which will increase h) is necessary to restore equality. Thus along the $\dot{\lambda} = 0$ line $\frac{d(\frac{\lambda}{\psi})}{d\lambda} > 0$. A similar argument, based on writing h as $h=\tilde{H}(\psi, \frac{\lambda}{\psi})$, shows that along the $\dot{\psi} = 0$ line we must have $\frac{d(\frac{\lambda}{\psi})}{d\psi} < 0$.

REFERENCES

A. B. Atkinson, "Capital Taxes, the Redistribution of Wealth and Individual Savings". Review of Economic Studies, 1971, 209-228.

G. Becker, Human Capital. National Bureau of Economic Research, New York 1964.

G. Becker and G. Ghez, " The Allocation of Time and Goods Over the Life Cycle". National Bureau of Economic Research, January 1974.

Y. Ben-Porath, "The Production of Human Capital Over Time". Journal of Political Economy, August 1967.

A. S. Blinder, Toward an Economic Theory of Income Distribution, Cambridge: MIT Press, 1974.

R. Eisner and R. Strotz, "The Determinants of Business Investment", in D. B. Suits et. al. Impacts of Monetary Policy, Englewood Cliffs, 1963, pp. 60-338.

T. Ishikawa, "A Simple Jevonian Model of Educational Investment Revisited", Harvard University Discussion Paper, No. 289, April 1973.

J. Heckman, "Estimates of Human Capital Production Functions Embedded in a Life Cycle Model of Labor Supply" in N. Terleckyj (ed.), Household Production and Consumption, Studies in Income and Wealth Vol. 139, forthcoming in 1975.

J. Hirschleifer, "On the Theory of Optimal Investment Decision", Journal of Political Economy, August 1958.

D. Jorgenson, "Capital Theory and Investment Behavior", American Economic Review, May 1963, pp. 137-176.

M. Landsberger and U. Passy, "Human Capital, Its Shadow Price and Labor Supply", Technion Mimeograph Series No. 138, August 1973.

J Mincer, Schooling Experience and Earnings, National Bureau of Economic Research, New York 1974.

S. Rosen, "Income Generating Functions and Capital Accumulation", Harvard University discussion paper No. 306, June 1973.

S. Rosen, "Learning and Experience In the Labor Market", Journal of Human Resources, Summer 1972.

E. Sheshinski, "On the Individual's Life Time Allocation Between Education and Work", Metroeconomica, April-June 1968.

F. Stafford and P. Stephan, "Labor, Leisure and Training Over the Life Cycle", University of Michigan Working Paper No. 7374-14, November 1973.

Y. Weiss, "Learning By Doing and Occupational Specialization", Journal of Economic Theory, June 1971a, pp. 189-198.

Y. Weiss, "Investment in Graduate Education", American Economic Review, December 1971b, pp. 883-852.

Y. Weiss, "On the Optimal Pattern of Labor Supply", Economic Journal, December 1972, pp. 1293-1315.

Y. Weiss, "Notes on Income Generating Functions", unpublished mimeograph, Princeton 1974.

C. C. von Weiszacker, "Training Policies Under Conditions of Technical Progress: A Theoretical Treatment", in Mathematical Models in Educational Planning, Organization for European Co-operation and Development, Paris 1967.

M. Yaari, "On the Consumer's Lifetime Allocation Process", International Economic Review, 1964, pp. 304-317.