Volume Title: Education, Income, and Human Behavior

Volume Author/Editor: F. Thomas Juster, ed.

Volume Publisher: NBER

Volume ISBN: 0-07-010068-3

Volume URL: http://www.nber.org/books/just75-1

Publication Date: 1975

Chapter Title: Measuring the Obsolescence of Knowledge

Chapter Author: Sherwin Rosen

Chapter URL: http://www.nber.org/chapters/c3698

Chapter pages in book: (p. 199 - 232)

8. Measuring the Obsolescence of Knowledge

by Sherwin Rosen

INTRODUCTION

In this chapter a method is outlined for determining rates of obsolescence and depreciation of knowledge and skills, and preliminary estimates are presented of identifiable parameters for white male high school and college graduates in 1959. The conceptual framework of the study rests on the by now well-known view that knowledge embedded in human agents of production can be treated as a kind of capital (Becker, 1964; Bowman et al., 1968). Learning is the embodiment of a portion of existing knowledge in oneself and represents the acquisition of a capital good or investment. Since embodied knowledge is not directly observable, estimation requires some prior economic analysis. Therefore, a model of optimum accumulation of knowledge is developed here, based on the hypothesis that individuals learn from their working experiences.

PRELIMINARIES

Generally speaking, several dimensions of capital deterioration must be distinguished. First there is the concept of *obsolescence,* defined as negative changes in capital values that are solely a function of chronological time. Obsolescence occurs because stocks of knowledge available to society change from time to time. Different generations of graduates acquire knowledge from schools at various points in time, and obsolescence is obviously related to some concept of "vintage." Knowledge available to be learned systematically changes as research and innovation push out the frontiers of various subjects. Sometimes new knowledge proves

NOTE: I am indebted to Edward Zabel, Richard Rosett, and G. S. Maddala for advice on a number of points. Yoram Barzel and F. Thomas Juster contributed useful comments on an initial draft. Financial support from the Carnegie Commission on Higher Education and the National Institute of Education is gratefully acknowledged, but they do not bear any responsibility for the views presented in this chapter.

received knowledge to be incorrect or at least less general than was supposed at an earlier time. Similarly, production innovations often render useless skills associated with prior methods. In both cases, capital losses are imposed on those embodying the earlier knowledge and skills. However, this need not necessarily be the case. New discoveries can augment previously available knowledge in an essentially orthogonal manner. Moreover, both the process of simplification—which renders existing knowledge more accessible to students—and innovations in teaching methods themselves make standard exposures to learning environments more productive. Such changes increase value added from given resource inputs and reduce private and social costs of learning. In these cases no absolute capital losses are involved, but there is a real sense in which patterns of relative capital losses emerge. Finally, gross output of educational institutions can change from time to time if there are corresponding changes in "raw material." It is often argued, for instance, that the process of evolution implies increasing ability of successive generations. As will be clear from the discussion below, there are few possibilities for distinguishing among these dimensions in my work, and usually all are combined into a single rate of obsolescence.

The second concept that must be identified is *depreciation,* defined as negative changes in capital values which depend on the age of persons possessing knowledge and skills, and which are more or less independent of chronological time and generational differences. Depreciation arises because the ability of individuals to apply acquired skills and knowledge to income-producing opportunities systematically changes with age. Some have maintained that appreciation characterizes early phases of working life, analogous to the effects of storage on the quality of wine. However, depreciation finally occurs as a result of increasing probabilities of death and morbidity as well as general deterioration of mental and physical capacities associated with aging. Further, it will be argued below that learning is not wholly confined to schools, but occurs for very long periods after formal schooling ends. Capacity to learn and adapt to new situations may decrease with age.

Two general methods are available for examining questions concerning obsolescence and depreciation of capital goods. Direct observation is one possibility. In the case of physical capital, engineering studies of the useful life of machines and case studies of particular innovations may be valuable. In the case of human skills, there are opportunities to engage in psychological and physiological testing of individuals at different ages. For scientific knowledge, citation studies determining half-lives of publications and related methods are often suggestive (Lovell, 1973). The problem may also be cast into the standard framework of technical change by studying educational production functions, relating learning measures to a variety of educational inputs (Coleman et al., 1966), and computing various productivity measures. An alternative and complementary methodology may be derived from an economist's perspective by recognizing that all problems of obsolescence and depreciation ultimately relate to the theory of capital value. Observations on changes in capital values, rather than on their physical counterparts, provide a great deal of information on deterioration rates (Hall, 1968). Though not widely known, value methods have been applied with great success to certain types of physical capital, such as transportation equipment (Cagan, 1971; Hall, 1971). A similar approach is taken here. It is based on a natural application of the theory of capital and is useful for organizing the measurement problem. Furthermore, valuation methods have some practical advantage in that knowledge or skill as capital, though a useful construct for many problems, is not yet capable of direct measurement. Only the consequences of learning are observable through effects on income and other behavioral variables.

A few more clarifying comments are necessary before turning to specifics. First, in all these problems the major difference between knowledge capital and physical capital is the former's absence of observable market valuation. However, no real difficulty arises on that score, for services of knowledge and skills embodied in people are traded on well-developed rental markets—namely, labor markets—and rental values contain the same information as capital values. The theory of capital is just as well carried out in terms of flow or rental prices as in terms of stock or asset prices as long as the accounting is done correctly. Second, it is necessary to keep in mind that all models of the sort put forth below, whether relating to knowledge or to machines, are no better than the valuation hypothesis applied to them. It will be assumed here that individuals are systematically paid in proportion to the services of their knowledge. Knowledge undoubtedly has many characteristics of a common property resource, since acquisition of a portion of the existing stock by one person in no sense diminishes the quantity available for others to acquire. But learning activities require personal expenditures of resources in terms of outlays of money, time, and effort. Since acquisition of learning is not free, there is no reason to suppose that implicit market rental prices of existing skills do not systematically reflect social as well as private productivity.

If, in addition to embodying existing knowledge in people, learning creates new knowledge available to society at large, private marginal product may differ from social product. That is, private incentives for learning and investment can affect realized patterns of obsolescence. By concentrating on individual behavior, I am justified in ignoring feedbacks between embodied learning and the creation of new knowledge whose values are not captured by their innovators. Thus patterns of obsolescence rates are treated as exogenous; the analysis refers to new knowledge as it has actually evolved, on the basis of previous private incentives for accumulation and not on the basis of how new knowledge *should* have evolved in the presence of appropriate subsidies to inventive activities that would confer external benefits to society as a whole.

The remainder of the chapter is organized as follows. First, the nature of valuation methods is illustrated by means of a simple example, and a basic analytical difficulty is examined: Since individuals can partially avoid the consequences of obsolescence by "retooling" and learning new techniques as part of their working experience, observed market incomes (rentals) reflect both knowledge acquired in school and knowledge acquired in conjunction with work activity. Following this example, a theory of learning by experience is sketched that is based on the principles of capital accumulation. The model is made operational by demonstrating an explicit mechanism in the labor market whereby optimum learning can be achieved and by showing that life-cycle earnings can be approximated by a nonlinear function of age or work experience. Estimated parameters are sufficient to identify depreciation-obsolescence rates, certain "vintage" effects, and some "ability" factors. Finally, preliminary estimates of the model using 1960 census of population data are presented and interpreted, and tentative conclusions are made. Data limitations preclude identification of all parameters at this stage of the investigation, although the feasibility of the method is indicated.

AN EXAMPLE

To what extent is it possible to estimate depreciation and obsolescence rates from observations on earnings of individuals possessing various amounts of education obtained at different points in time? Clearly, some kind of vintage model (Solow, 1960) is required by analogy with technical change and depreciation of physical capital. Let us take automobiles as an example. Ideally, information exists on last year of school completed, major subject ("make" and "model number"), year of graduation (vintage), and age (depreciation) and can be related to annual rentals or income (Griliches, 1967). Consider the following elementary model.

Let X_t denote earnings of persons t years of age in a cross section, all of whom have completed the same level of formal education; h_t denotes an index of knowledge and skill possessed by a person age t, and R is the implicit market rental price per unit of knowledge during the year of observation. Suppose individuals receive their education at age "zero," where t represents years of work experience, and working life is N years in length. Estimation requires further specification.

1 Assume that the conditions for existence of a capital aggregate are fulfilled (Fisher, 1965), meaning that knowledge of various ages and vintages can be (conceptually) measured in "equivalent units." For example, the stock of knowledge acquired from one vintage is a fixed percentage more or less than the stock of another vintage, and similarly for skills possessed by people of different ages. It is clear that assumptions of this sort are necessary for estimation to proceed at all. If knowledge and skill depreciate at rate δ_j in the year of life j, the current stock of skill of a person age t is his initial stock multiplied by a factor $\prod_{j=0}^{t-1}(1 - \delta_j)$. In addition, successive generations enter the labor market with increasingly larger initial stocks of knowledge. Let γ_i represent an annual improvement factor in knowledge obtained from school in calendar year i, taking the origin as N years ago. γ_i is related to the relative rate of obsolescence in year i. It follows that

$$h_t = h_0 \prod_{i=0}^{N-t}(1 + \gamma_i) \prod_{j=0}^{t-1}(1 - \delta_j) \tag{8-1}$$

where h_0 is initial knowledge of a person N years of age in the cross section. Then Eq. (8-1) describes the evolution of knowledge over successive generations. Obsolescence and depreciation factors are not separately identified in this formulation, since age and vintage are linearly related. However, this fact has little bearing on the conclusions to be derived from this example.

2 A natural assumption concerning valuation is that earnings are proportional to the services of knowledge rented at each age. That is,

$$X_t = Rh_t \tag{8-2}$$

Equation (8-2) is based on the assumption that there is competition in the labor market and that persons possessing greater skills earn correspondingly greater amounts. It is important to note that Eqs. (8-1) and (8-2) constitute a "theory" of income determination for individuals—a quite elementary theory to be sure, but a theory nevertheless. Manipulation of these equations reveals that

$$X_t / X_{t-1} = (1 - \delta_{t-1}) / (1 + \gamma_{N-t+1}) \tag{8-3}$$

Comparing earnings of persons one year apart in age (or vintage) provides estimates of combined obsolescence-depreciation factors for each year.

Are such estimates reliable? The provisional answer must be "no." If the δ and γ terms are all positive, Eq. (8-3) implies that age-earnings profiles in the cross section are monotonically decreasing: older persons possess less skill because they received less from their education (acquired at an earlier date) and also because what they did learn has depreciated over a longer period of time. The conclusion is unaltered if some of the initial depreciation terms are negative (indicating "appreciation"), as long as they are not too large in absolute value. In any event, the major prediction of Eqs. (8-1) and (8-2) is clearly rejected by observation. Earnings rise with age for at least 15 years after graduation at every level of schooling (Hanoch, 1967), indicating that the model is at best seriously incomplete.

Evidently, knowledge is not produced only in schools, and learning does not cease after formal schooling ends. Instead, after some period of full-time school activity, knowledge is most efficiently acquired by shifting its source of production to the labor market and allowing people to *learn from their working experiences.* It can be argued that formal schooling equips students to learn new skills more effectively on their own, but whatever the role of formal schooling, it is certain that if individuals learn from work experience, the stock of knowledge at each age consists of several vintages. Individuals have strong incentives to acquire new skills as they become available in order to maintain their capital intact, and these incentives must be incorporated in the model. Owners of used cars seldom undertake expenditures necessary to make them indistinguishable from new cars, but the same is not true of skills.[1]

The result of this logic suggests that in addition to depreciation and obsolescence effects, net learning or investment terms should be incorporated into the function describing the evolution of embodied knowledge over a person's lifetime. Moreover, current gross accumulation costs must be subtracted from gross rentals (earnings capacity) to arrive at the age-earnings function (Becker, 1964). This, however, causes a serious conceptual difficulty to arise. Knowledge embodied in a person is not directly observable, and it is necessary to estimate capital accumulation (in value terms) at each age as well as obsolescence and depreciation rates. If working life is N years, N observations on income are available, one for each age. Yet it is necessary to estimate more than N variables—gross capital accumulation or learning at each age, as well as terms in γ and δ. Hence the problem cannot be solved without imposing a priori restrictions on some of the unknown variables to reduce their number. The solution adopted here is to posit a particular relationship between working experience and learning, based on a model of optimum learning in the labor market. My restrictions stem from a particular learning model and must stand or fall on the basis of that particular construction. It should be borne in mind that basic observational limitations preclude a straightforward accounting approach to the problem. The example above shows that in principle, rates of obsolescence and depreciation cannot be estimated as a "pure" problem in measurement and in the absence of a model.

The model I constructed is discussed in some detail in the following section. I have tried to make the arguments as accessible as space limitations permit and to spell out all assumptions underlying the estimates that will follow.

THE MODEL

First, I shall sketch the general economic framework of the model and then analytically state the problem and its solution, and finally I shall derive age-earnings profiles implicit in the model and suitable for estimation.

[1] Skills are more like residential structures than consumer durables in this respect. The entire issue usually is ignored in durable goods studies: Expenditure after initial purchase is treated as "normal maintenance," but so-called maintenance expenditures are really investments and change the economic life of the goods.

Markets for Learning Opportunities

Economists have long recognized that labor market activities involve simultaneous purchases and sales, or tie-in contracts, between workers and their employers, and that approach is pursued here. Learning is a *joint product* of work activity. A learning environment is implicit in every assignment of work routine and each job is associated with a definite amount of learning opportunity and work activity. Suppose that knowledge is completely vested in the person acquiring it and has general market value, not specific to any firm. Workers sell the services of their knowledge, but at the same time they purchase opportunities to learn something, depending on the type of job chosen. By the same token, firms purchase the services of their employees' knowledge and also sell them opportunities to learn, depending on the type of job provided. For a given job and implied work-learning combination, individuals apply their existing knowledge and skill both to produce marketable output for employers and to embody additional knowledge in themselves. In making employment applications, workers are faced with a great variety of choices among various jobs, each offering different opportunities to learn. It is choice among jobs—each associated with a fixed learning potential and work assignment, but with the ratio between the two varying from job to job—that offers a margin of choice and the possibility for constructing operational models of optimum accumulation of knowledge in the labor market.

Markets for learning opportunities are cleared through the market for jobs. Market equilibrium is characterized by a set of implicit prices of learning options revealed to workers and employers in the form of equalizing wage differences between jobs. Ordinarily, jobs yielding larger learning possibilities sell for higher unit prices, and an individual working at one of them earns less income than he could if he worked at a job with a lesser possibility for learning. Earnings forgone is the price paid for learning in the labor market. Workers demand jobs with learning content and are willing to pay that price to increase their future earning prospects. Learning options are supplied because the learning content of work is not fixed once and for all, but can be altered by reallocating resources from production of physical output to teaching. Firms engage in multiple production and in a sense also are in the "education business." Given market prices for learning, employers choose the optimum combination of work-learning activities offered by *designing* jobs in the appropriate manner. The costs of providing greater learning opportunities are the additional physical output lost as a result of devoting greater proportions of input time to teaching and learning rather than to current production. Rising supply price results from increasing marginal rates of transformation between marketable output forgone and learning activity, and a competitive market insures that learning opportunities are supplied at marginal production cost.[2]

As will be seen below, maximization of lifetime wealth by workers implies optimum choices of jobs over the life cycle. A corresponding progression through a sequence of work activities is implied and constitutes a theory of occupational mobility. Moreover, such choices generate observable lifetime earnings patterns.[3] Thus the model yields an age-earnings generating function, whose parameters depend on variables relevant for making the best choices. Age-earnings profiles are determined by obsolescence and depreciation rates, initial stocks of knowledge, and a few other parameters, thus providing a basis for estimation.

Learning by Experience

A complete statement of the problem and its solution requires more precise specification. Attention is focused on one human factor of production, a particular kind of skill and knowledge. Let h_t denote the stock of knowledge embodied in a person at the beginning of period t. z_t represents gross learning between periods t and $t + 1$, defined as the gross change in stock between those dates. Assume that depreciation-obsolescence occurs at a constant geometric rate δ over the time spanned by the data.[4] Then gross learning equals the net change in stock plus depreciation:

$$z_t = (h_{t+1} - h_t) + \delta h_t = h_{t+1} - (1 - \delta) h_t \qquad (8\text{-}4)$$

[2] The reader is referred to Rosen (1972) for details and some wider implications of markets for learning options. A related model has been developed by Ben-Porath (1967). The novelty of the present model lies in the joint-product-learning-market construction, in a learning-by-experience context.

[3] The theory can be viewed in terms of supply and demand for *lifetime* incomes, in that workers choose an optimum progression through a hierarchy of work-learning activities. Thus current labor market contracts involve implicit forward contracts for future income. To the extent that work-connected learning is firm-specific and workers share returns, lifetime earnings patterns must be the same as in cases where knowledge has general market value, as long as there is competition in the market for lifetime earnings.

[4] Since there is no possibility of distinguishing between age (depreciation) and work experience (obsolescence) in census data, δ must be treated as a combined deterioration rate. No analytical difficulties arise if δ is not constant over working life, but empirical implementation is more difficult.

Of course, z_t must be nonnegative. Each work activity or job is associated with a given value of an index I, measuring the size of the learning option connected with it. I is an index of gross learning potential on each job and represents the amount of "space" and "time" devoted to learning rather than to current production. For concreteness, the reader might think of this index as the labor market analogue of the teacher-student ratio relevant to formal schooling. For example, the value of I associated with management trainees exceeds that for executive vice-presidents; the value for carpenter's apprentices is greater than that for journeymen carpenters; etc.

As noted above, labor market equilibrium establishes a functional relationship between implicit prices and learning attributes of jobs. Let the function $P(I)$ represent the market equilibrium (shadow) price of jobs offering a learning-potential index of I. $P(I)$ is implicit learning expenditure incurred by the worker on option I. On the assumption of a rising supply price of options, $P'(I)$ and $P''(I)$ are positive: the marginal cost of learning opportunities is positive and increasing. Further $P(0) = 0$, for no expenditures need be undertaken if learning is zero. A person's earnings equal the value of services he has to sell *minus* the cost of the learning option he buys. If R_t is the implicit market rental price on the services of embodied knowledge h in period t and if y_t is observed earnings during that period, then

$$y_t = R_t h_t - P(I_t) \tag{8-5}$$

where $R_t h_t$ is earning capacity, or the value of services rented during the period, and $P(I_t)$ is expenditure on the learning option purchased at age t. Thus $P(I)$ is market-determined in such a way to "equalize" wages across work activities with alternative learning values. Given current knowledge, Eq. (8-5) shows that the worker is confronted by a market-determined trade-off between current earnings (y) and learning opportunity (I). Current income is sacrificed if positive I is chosen, but future earnings prospects are enhanced through increased future values of h. The assumptions on $P(I)$ ensure concavity of the transformation function.

Obviously, the next step requires specifying a relationship between learning options and actual learning. Notice my continual use of the terms *option* and *opportunity* in discussing learning possibilities. The reason for this is to allow for differences among individuals in the amounts of real knowledge obtained from the same work activity. Workers differ with respect to ability and other requisites for learning. To account for these facts, it is helpful to postulate a production function relating gross learning to the nature of the job I and to embodied knowledge h. The amount a person knows clearly affects his capacity to learn:

$$z = \alpha f(I, h) \qquad \text{with } z \geq 0 \text{ and } f(0, h) = 0 \tag{8-6}$$

where α is a generalized ability parameter that may vary from person to person.[5] Assumed properties of the production are as follows: (1) $f_I > 0$, $f_h > 0$—jobs with greater learning content increase realized gross learning, and additional knowledge increases real learning capacity; (2) f_{II} and f_{hh} are negative—marginal products of knowledge and options in producing learning are diminishing; (3) $c \geq f_{Ih} \geq 0$, where c is a constant—more knowledge can increase the real investment capacity of a marginal option, but only to a limited extent; and (4) $f(I,h)$ is concave—learning in the labor market is not subject to increasing returns.

Wealth (discounted lifetime earnings) at age of entry into the labor force is

$$\sum_{t=0}^{N} y_t/(1 + r)^t \tag{8-7}$$

where r is a fixed rate of discount. The problem is to choose a sequence $\{I_t\}$ over working life that maximizes lifetime earnings [Eq. (8-7)], subject to the restrictions (8-4), (8-5), (8-6), and h_0, an initial endowment of knowledge at the time of entry into the market. Optimum values of I_t and starting stock h_0 imply corresponding values for z_t in each period, by Eq. (8-6). Hence the sequence $\{I_t\}$ and h_0 imply a corresponding sequence $\{z_t\}$ describing learning patterns over working life. Moreover, $\{z_t\}$ and h_0, along with Eq. (8-4), describe the evolution of knowledge $\{h_t\}$ over the life cycle. Finally, $\{h_t\}$ and $\{I_t\}$ can be substituted into Eq. (8-5) to generate observable age-earnings patterns.

In this chapter interest is centered on income profiles resulting from optimum accumulation rather than on the occupational mobility function I_t. Therefore, it is convenient to transform the problem by directly substituting the constraints (8-4) and (8-6)

[5] More generally, Eq. (8-6) might include t as an argument to allow for life-cycle changes in learning capacity. That possibility is ignored here.

into the definition of income [Eq. (8-5)] at the outset. Define total cost of realized learning as $F(z,h)$. Then if $I = g(z,h)$ is an inverse function of f, $F(z,h) = P[g(z,h)]$. The assumptions on P and f imply the following: (1) F_z and $F_{zz} > 0$—marginal cost of learning is positive and increasing; (2) $F_h < 0$ and $F_{hh} > 0$—greater knowledge can decrease the total costs of learning, but at a decreasing rate; (3) $F_{zh} \leq 0$, but exceeds some negative amount—greater knowledge can decrease the marginal costs of learning, but only to a limited extent; and (4) $F_{zz}F_{hh} - F_{zh}^2 > 0$—total cost is a strictly convex function of learning and knowledge. Learning is subject to increasing cost because larger options are available only at increasing unit price, and learning is not subject to increasing returns to scale.

The problem can now be stated simply as follows. Maximize

$$\sum_{t=0}^{N} \{R_t h_t - F[h_{t+1} - (1-\delta)h_t, h_t]\}/(1+r)^t \qquad (8\text{-}8)$$

with respect to a sequence of values $\{h_t\}$, subject to h_0 and $h_{t+1} \geq (1-\delta)h_t$. At a given level of skill, the term $[Rh - F(z,h)]$ in Eq. (8-8) defines a trade-off between actual learning and current earnings. It also describes the effects of greater knowledge on these terms of trade and on future trade-offs. Choice of $\{I_t\}$ is suppressed, but is an automatic consequence of choice of $\{h_t\}$, from Eqs. (8-4) and (8-6).

Maximization of Eq. (8-8), subject to the initial endowment and the restriction that gross learning cannot be negative, is a dynamic programming problem. Only the major features of the solution are discussed here, and formal proofs are omitted.[6] The following two properties are essential:

1 Optimum learning patterns consist of two segments. Define a *critical age* T, where T is at most $N - 1$ (age at retirement minus one year). Then optimum gross learning is positive at all ages less than T and is set equal to zero for all ages greater than, or equal to, T: $z_t > 0$ for $t < T$, and $z_t = 0$ for $t \geq T$. Gross knowledge is accumulated up to the critical age T, after which no investment options are purchased. From time T onward, gross learning is zero, and embodied capital is allowed to deteriorate at the depreciation-obsolescence rate δ. Specialization and nonmarginal behavior

[6] The problem is formulated in terms of dynamic programming in Rosen (1971), an early version of this chapter. Derivation and rigorous proof of the optimum policy are also found there.

beyond age T follow from the fact that embodied knowledge has no value after working life ends, but accumulation is always costly. Certainly no investment option is chosen in the last year of working life, since only zero returns can be obtained on it. The critical age T is less than $N-1$ if certain limiting properties of $P(I)$ and $\alpha f(I,h)$ obtain. When $F_z(0,h) > 0$, marginal cost of learning can exceed marginal return at ages less than $N-1$, so that $T < (N-1)$.

2 A necessary condition for optimality during the "investment period" $t < T$ is

$$(1 + r)F_z(z_{t-1}, h_{t-1}) = R_t - F_h(z_t, h_t) + (1 - \delta)F_z(z_t, h_t) \qquad t < T \tag{8-9}$$

Using the notation $F_{it} = F_i(z_t, h_t)$ to avoid writing the arguments of the functions each time, Eq. (8-9) can be rearranged to read:

$$F_{zt-1}\left[r + \frac{\delta - (F_{zt} - F_{zt-1})}{F_{zt-1}}\right] = R_t - F_{ht}$$

The meaning is clear. F_z is the marginal cost of learning in terms of current income forgone. The term on the left converts stock costs into periodic flows through amortization by a factor reflecting interest expense, depreciation-obsolescence, and capital revaluation next period. The term on the right is marginal revenue in flow terms, reflecting next period's rental value and the marginal value of knowledge for increasing future learning capacity. An equivalent expression in terms of stocks can be obtained by iteration of Eq. (8-9):

$$F_{zt} = \frac{R_{t+1} - F_{ht+1}}{(1 + r)} + \frac{[R_{t+2} - F_{ht+2}](1 - \delta)}{(1 + r)^2} + \cdots + \frac{[R_{T-1} - F_{hT-1}](1 - \delta)^{T-t-1}}{(1 + r)^{T-t}} + \frac{R_T(1 - \delta)^{T-t}}{(1 + r)^{T-t+1}} + \cdots + \frac{R_N(1 - \delta)^{N-t-1}}{(1 + r)^{N-t}}$$

This expression states the familiar criterion that marginal cost equals discounted marginal revenue. Finally, if R_t is sufficiently regular, the assumptions on F_{ij} for ensuring sufficient conditions for a maximum are satisfied, and the solution is unique.

Age-Earnings Profiles

Explicit solutions for h and y as functions of t can be found by choosing a functional form for F and applying condition (8-9). A slightly more general procedure is adopted here. Condition (8-9) is linearized by use of Taylor's series approximations in the neighborhood of some point. The technique is exact if $F(z,h)$ is quadratic.

For the time being, consider a case where $R = R_t$ for all values of t. Equation (8-9) is approximated near a value $\bar{h}$, implicitly defined by

$$(r+\delta)F_z(\delta\bar{h},\bar{h}) = R - F_h(\delta\bar{h},\bar{h}) \tag{8-10}$$

$\bar{h}$ is a stationary point, assumed to exist, at which net learning is zero; $\bar{h}$ is maintained indefinitely. In fact Eq. (8-10) is a condition of optimality that would hold at a stationary state if a person had an indefinitely long lifetime. Define a new variable $\tilde{h}_t = h_t - \bar{h}$. The arguments of Eq. (8-9) are h_t, h_{t+1}, and h_{t+2}, so that linearization yields a second-order linear difference equation in h_t. The homogeneous part is

$$\tilde{h}_{t+1} - B\tilde{h}_t + (1+r)\tilde{h}_{t-1} = 0 \tag{8-11}$$

with

$$B = (1-\delta) + \frac{(1+r)F_{zz} + F_{hh} - (1-\delta)F_{zh}}{(1-\delta)F_{zz} - F_{zh}} \tag{8-12}$$

The derivatives in Eq. (8-12) are understood to be evaluated at $(\delta\bar{h},\bar{h})$. The general solution of Eq. (8-11) as an explicit function of t is

$$\tilde{h}_t = C_1\lambda_1{}^t + C_2\lambda_2{}^t \tag{8-13}$$

where λ_1 and λ_2 are the roots of Eq. (8-11) given by

$$\lambda = \frac{B \pm \sqrt{B^2 - 4(1+r)}}{2}$$

It can be shown that λ_1 and λ_2 are real numbers, precluding certain cycles in the generation of $\tilde{h}_t$. C_1 and C_2 are constants determined by initial and terminal conditions h_0 and $z_T = 0$. Therefore, C_1 and C_2 are functions of h_0 and $\bar{h}$. Finally, it follows from the definition of λ that

$$(\lambda_1)(\lambda_2) = (1 + r)$$
$$\lambda_1 + \lambda_2 = B \qquad (8\text{-}14)$$

The functional form for lifetime income for which estimates can be made is obtained by substituting the knowledge-generating function (8-13) into the definition of earnings:

$$y_t = Rh_t - F(z_t, h_t)$$

Choosing a first-order Taylor's approximation for F yields a double geometric function of age or experience (t):

$$y_t = k_0 + k_1\lambda_1{}^t + k_2\lambda_2{}^t \qquad t < T$$

where the k's are constants.[7] Furthermore, it has been established that $z_t = 0$, or $h_t = (1-\delta)h_{t-1}$, for the phase $t \geq T$. It is also true that $y_t = Rh_t$ for that phase, since no investment costs are incurred then; i.e., $F(0,h) = 0$. Therefore, the complete earnings-generating function over working life is given by

$$y_t = k_0 + k_1\lambda_1{}^t + k_2\lambda_2{}^t \qquad 0 \leq t < T \qquad (8\text{-}15a)$$

$$y_t = (1 - \delta)y_{t-1} \qquad T \leq t \leq N \qquad (8\text{-}15b)$$

To establish the claim that the model captures all essential features of observed age-earnings profiles, consider Eq. (8-15*a* and *b*) more closely. Suppose that λ_1 is less than unity, that λ_2 is greater than unity—the product of $\lambda_1\lambda_2$ exceeds unity from Eq. (8-14)—and that k_1 and k_2 are negative. Then the first two terms in Eq. (8-15*a*), $k_0 + k_1\lambda_1{}^t$, plot a rising, concave, geometric curve, whereas the third term, $k_2\lambda_2{}^t$, plots falling and increasingly negative values. If k_2 is sufficiently small, the sum of the two curves results in earnings' rising at early working ages, having a relatively flat middle portion, and falling as the critical age T is approached. At

[7] The income-generating function for $t < T$ is equivalent to a second-order linear difference equation $y_t - (\lambda_1 + \lambda_2)y_{t-1} + (\lambda_1\lambda_2)y_{t-2} =$ constant, as can be seen by direct substitution. Also, if a second-order Taylor's approximation for $F(z, h)$ is used, terms in $\lambda_1{}^{2t}$, $\lambda_2{}^{2t}$, and $(\lambda_1\lambda_2)^t$ as well as in $\lambda_1{}^t$ and $\lambda_2{}^t$ are required.

age T, Eq. (8-15*b*) takes over, and earnings fall at rate $(1-\delta)$ until work life ends. Clearly, Eq. (8-15*a*) can duplicate observed age-earnings profiles. As will be seen, the assumption of constant R is not crucial to this characterization.

Model (8-15*a*) contains seven parameters—k_0, k_1, k_2, λ_1, λ_2, δ, and T—and can be estimated by maximum likelihood methods from observed earnings.[8] Once the critical age T is found, δ is estimated from the earnings pattern beyond age T. The k's and λ_1 and λ_2 are estimated nonlinearly from the portion to the left of T and identify $(1+r)$ and B, from Eq. (8-14). Equation (8-12) shows that B is a function of the second derivatives of investment costs F_{ij}. With an additional assumption, the F_{ij} terms can be reduced to one more parameter. Suppose the learning-production function is approximated by $z = aI^{1-\beta}h^{\beta}$, where β is the marginal product of knowledge with respect to the output of learning. Next, approximate the equalizing-difference function by a quadratic, $P(I) = AI^2$. Some algebraic manipulation reveals that

$$\begin{aligned} F_{zh}/F_{zz} &= -[2\beta/(1+\beta)]\delta \\ F_{hh}/F_{zz} &= \beta\delta^2 \end{aligned} \qquad (8\text{-}16)$$

making use of the fact that $\bar{z} = \delta\bar{h}$ at the point of approximation. Substituting Eq. (8-16) into the definition of B [Eq. (8-12)] yields a relation between r, δ, B, and β, allowing identification of β, since r, δ, and B are estimated independently.

Finally, examination of optimality condition (8-9) brings out the importance of future expectations on learning behavior because current choices depend on anticipated future events. No analytical difficulties arise once an actual expectations mechanism is postulated. However, many alternative specifications are possible. At this stage of the investigation, only one possibility is considered, involving an assumption of perfect foresight. Write

$$y_t = (1+\rho)^t[Rh_t - F(z_t, h_t)] \qquad \rho > 0 \qquad (8\text{-}17)$$

[8] Of course, other investigators have recognized the wealth of information contained in age-earnings patterns. For example, see Johnson (1970) and Mincer (1970). In those papers the distinction between gross and net investment is not clear. Moreover, the equations actually estimated are not explicitly derived from a formal model, and some of the parameters cannot be interpreted. A rather different approach is taken by Eckaus and El-Safty (1972).

where ρ reflects the secular rise over time of cross-sectional age-earnings patterns. On this specification, both anticipated and realized rentals per unit of skill, as well as total costs of accumulation, rise at rate ρ over the worker's lifetime. All the previous conclusions above are unaltered because the factor $(1+\rho)^t$ multiplies both sides of condition (8-9) and cancels out. The knowledge-generating function (8-13) is still valid, though a few minor alterations are necessary. Now the stationary point of approximation is defined by $(r + \delta - \rho)F_z(\delta\bar{h},\bar{h}) = R - F_h(\delta\bar{h},\bar{h})$, where the discount factor has been corrected for real growth in the economy. Thus the discount term $(1+r)$ in Eq. (8-14) must be replaced by the "real" rate of interest $(1+r)/(1+\rho)$. Otherwise, Eq. (8-15*a*) remains intact. Finally, $(1-\delta)$ in Eq. (8-15*b*) must be replaced by $(1-\delta)\ (1+\rho)$, since R is growing at rate ρ, though capital is deteriorating at rate δ. Evidently, observed life-cycle earnings beyond age T do not fall if ρ exceeds δ.

ESTIMATION

A model of lifetime earnings patterns of a single worker has just been derived. Although panel, or time-series, data are most appropriate for estimating income-generating functions such as Eq. (8-15*a* and *b*), these data are not available in sufficient detail, and parameters must be estimated from cross sections by age of individuals' earnings in a single year. Here the model is transformed to a cross-sectional basis, and then the data and estimates are presented.

A Vintage Model

An advantage of using cross-sectional data for estimation is that it is reasonable to assume equalization of rental prices per unit of real knowledge among all individuals, regardless of age or vintage. The major problem, common to both cross-sectional and cohort data, is to impute current stocks of knowledge and learning that depend on prior learning patterns and expectations across age groups. The method described below employs an implicit assumption of unbiased expectations. In effect, the procedure corrects cross-sectional observations for two types of exponential growth and allows intergenerational comparisons to be made. If these adjustments are valid, observations on individuals τ years of age are proportional to earnings actually received ξ years ago by individuals currently $(\tau+\xi)$ years old, and imputation of prior learning patterns is possible.

Let v denote an index of vintage or "generation number," with

$v = 0,1, \ldots, N$. Members of the oldest living generation in the cross section are chosen as the origin of v. If $y_t(v)$ is actual income at age t of members of generation v, Eq. (8-4) becomes

$$y_t(v) = (1+p)^{t+v}\{R_0 h_t(v) - F[z_t(v), h_t(v)]\} \quad (8\text{-}18)$$

where R_0 was the rental price of knowledge N years ago. Later generations receive capital gains because their earnings are higher at any given (z,h,t) combination; younger persons can look forward to greater real wealth if economic growth raises per capita incomes over time. Now as already noted, ρ has symmetrical effects on marginal returns and costs of learning for each generation and age. Consequently, the approximation (8-13), describing growth of knowledge during the learning period $0 \leq t < T$, still holds, except for correction of the rate of interest to include real growth, $(1+r)/(1+\rho)$ in Eq. (8-14). Moreover, on the assumptions previously stated regarding $F(z,h)$, the critical age T can be shown to be identical for all generations: $T(v) = T$. However, the constants in Eq. (8-13) are *not* invariant across generations, for they depend on initial endowments, which differ from generation to generation. In other words, C_1 and C_2 are functions of v. They can be written

$$C_i = a_{i0} + a_{i1} h_0(v) \qquad i = 1,2$$

where $h_0(v)$ is initial stock of knowledge at time of entry into the labor force of generation v and the a_{ij}'s are constants, dependent on $\bar{h}$, λ_1, λ_2, and δ.

Assume

$$h_0(v) = (1+\gamma)^v h_0(0)$$

Initial knowledge of successive generations grows at rate γ, all relative to the oldest living generation. γ is an exogenous vintage effect, capturing secular improvements in knowledge obtained from schools and in "basic ability" of individuals receiving diplomas.

A first-order approximation to Eq. (8-18) and some algebraic manipulation yield

$$y_t(v) = [k_{01} + k_{11}\lambda_1{}^t + k_{12}\lambda_2{}^t + k_{13}(1+\gamma)^v \lambda_1{}^t + k_{14}(1+\gamma)^v \lambda_2{}^t](1+\rho)^{t+v} \qquad t < T \quad (8\text{-}19)$$

where the k's are constants. Let X_t denote observed earnings at age t in the cross section. Then $v = (N - t)$ in the cross section, and $X_t = y_t(N-t)$. Substituting for v in Eq. (8-19), an observable function is

$$X_t = b_0 + b_1\lambda_1{}^t + b_2\lambda_2{}^t + b_3[\lambda_1/(1+\gamma)]^t + b_4[\lambda_2/(1+\gamma)]^t \qquad t < T \qquad (8\text{-}20a)$$

where the b's are functions of corresponding k's in Eq. (8-19) and of $(1+\gamma)^N$ and $(1+\rho)^N$.

For the period $T \leq t \leq N$, proceed as follows: First, earnings for those ages are defined by $y_t(v) = (1+\rho)^{t+v}R_0h_t(v)$, since $z_t(v) = 0$. Second, there is nothing in all the above to alter a previous conclusion that $h_t(v) = (1-\delta)h_{t-1}(v)$ for $t \geq T$ within each generation. Finally, $h_t(v+1) \simeq (1+\gamma)h_t(v)$, for $t \geq T$. Then

$$X_t/X_{t-1} = y_t(v)/y_{t-1}(v+1) \simeq [h_t(v)/h_{t-1}(v)][h_{t-1}(v)/h_{t-1}(v+1)] = \frac{1-\delta}{1+\gamma}$$

Therefore,

$$X_t = [(1-\delta)/(1+\gamma)]X_{t-1} = X_{T-1}[(1-\delta)/(1+\gamma)]^{t+1-T} \qquad t \geq T \quad (8\text{-}20b)$$

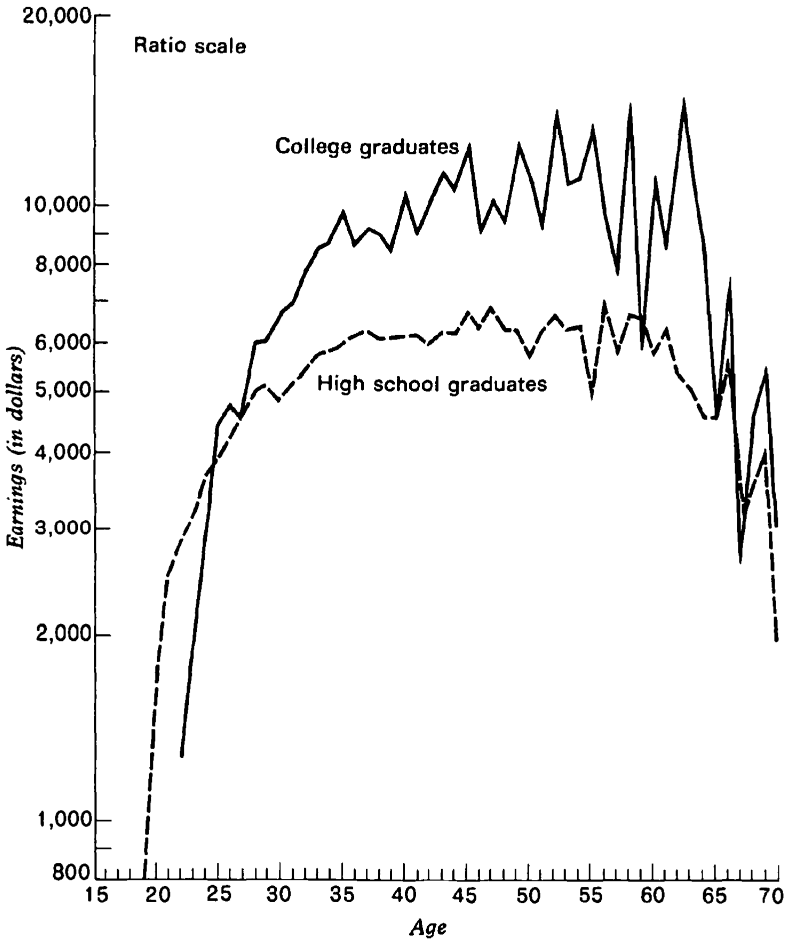
With the addition of stochastic terms, the cross-sectional age-earnings-generating function Eq. (8-20a and b) can be estimated by nonlinear maximum likelihood methods, with many degrees of freedom. Estimated parameters are sufficient to identify T, δ, γ, and $(1+r)/(1+\rho)$ exactly. On the further assumption of a Cobb-Douglas learning production function and quadratic learning-option cost function, the marginal product of knowledge with respect to learning β is also identified. Note that parameters for depreciation-obsolescence, vintage, critical age, and the real interest rate are estimated independently of precise assumptions regarding $F(\bar{z},\bar{h})$ and $P(\bar{I})$.

Data The data source is the 1960 census of population 1/1,000 sample. Records of males 14 years of age or older and not in the Armed

Forces were drawn from the sample and classified by educational attainment, age, race, and employment status. The number of nonwhites at each age was too small to yield reliable estimates of age-income profiles, especially at the college level, and they were omitted from consideration at this time. Age-earnings profiles (excluding nonemployment income) were estimated for white male high school and college graduates from corresponding means at each age; these are presented in Figure 8-1.

A major difficulty concerns the measurement of an earnings concept appropriate to the problem at hand. Most rate-of-return

FIGURE 8-1 ***Age-mean labor income, 1959; white male high school and college graduates***



SOURCE: 1960 census of population, 1/1,000 sample.

studies use annual incomes of members of the labor force; in effect, zero values are assigned to leisure time of members of the work force and to individuals out of the labor force. If nonworking time is valued at the hourly wage rate, the computation should refer to hourly wages multiplied by a "standard" number of hours to arrive at an earnings potential. Neither measure seems appropriate to the present problem. A main component of depreciation is inability to work at all, or at least not at maximum efficiency, because of ill health (broadly interpreted). Many individuals retire from the labor force for that reason and also to escape the consequences of obsolescence. Therefore, the value of their "leisure" cannot be evaluated at the wage of those who are still in the labor force. For this reason, the data in Figure 8-1 were computed over *all* individuals, whether or not they were in the labor force. Thus the measure used corresponds to expected yearly earnings of all survivors in 1959, classified by age. Actual earnings of those unemployed and out of the labor force are counted at zero, tantamount to assigning a zero value to their leisure. The earnings measure actually used undoubtedly overstates the case for including nonworking individuals at imputed wages less than those of the employed. No further adjustment has been made for probability of survival up to each age for that reason. Resulting biases are likely to be about the same for high school and college graduates. Thus, between-group comparisons of the estimates should be valid.

The reader will note considerable variability in earnings patterns of Figure 8-1, in contrast to the rather smooth profiles implied by the model. Most of the jaggedness in Figure 8-1 is due to sampling variation.[9] The number of observations on which means are computed falls very sharply with age. For example, there are fewer than 20 college graduates in the sample at each age past 60. Hence sampling variation increases with age. Moreover, the variance of working compared with nonworking individuals in the 1/1,000 sample is extremely high in the older age groups, and the data also exhibit serial correlation with respect to age. This, too, greatly contributes to the variation apparent in Figure 8-1.

The real problem in obtaining a firmer resolution of typical earnings streams is that the sample size (within age groups) is small. Though experimentation with alternative earnings concepts defi-

[9] Of course the model admits variations among individuals in h_0, r, and certain parameters in $F(z,h)$, implying corresponding earnings variations. For elaboration, see Rosen (1972).

nitely would be worthwhile, it is clearly unwarranted, given the small samples currently available.[10] Sensitivity of the estimates to alternative earnings measures must await larger samples, soon to be published. Therefore, the estimates presented below are not definitive and are to be taken as an indication of feasibility of the method.

Finally, most studies of rates of return to education smooth age-income data to remove the variability evident in Figure 8-1 and reduce the probability of computing multiple internal rates of return. Though estimates of internal rates of return may not be affected much by smoothing, the same cannot be said of estimated parameters in the present model. Essentially, time-series methods are used to estimate model Eq. (8-20*a* and *b*). As is well known, moving averages of random numbers can themselves generate what appears to be systematic cyclical behavior in smoothed series. Thus estimation of Eq. (8-20*a* and *b*) on data smoothed in that manner can result in estimates that are simply artifacts of the smoothing scheme and reflect no real underlying parameter of interest. To maintain consistency with the preliminary nature of the empirical investigation at this stage, I have chosen to estimate the model from the raw data of Figure 8-1, with no further adjustment to remove effects of sampling variation.

Estimation

Maximum likelihood estimates of model Eq. (8-20*a* and *b*) are obtained by an iterative, least squares procedure. For expository convenience, write Eq. (8-20*a*) as $X_t = \phi_1(t)$, where $\phi_1(t)$ is the right-hand side of Eq. (8-20*a*), a nonlinear function of work experience during the period $t < T$. Write Eq. (8-20*b*) as $X_t = \phi_2(t)$, where $\phi_2(t) = X_{T-1}[(1-\delta)/(1+\gamma)]^{t+1-T}$. Define two variables:

$$\tau = t \quad \text{and} \quad D = 0 \quad \text{for } 0 \leq t \leq T-1$$

$$\tau = T \quad \text{and} \quad D = 1 \quad \text{for } T \leq t \leq N$$

Therefore, the function to be estimated can be written

[10] The sampling base was narrowed by omitting current school enrollees and foreign-born persons. Earnings of the more homogeneous group are larger than those shown in Figure 8-1, but the general patterns remain intact. The high sampling variability seen in Figure 8-1 is also present in data drawn from a narrower sampling base. The real problem is small samples at older ages, and further computation is simply not worthwhile.

$$X_t = \phi_1(\tau) + D\phi_2(t - \tau) + u_t \qquad (8\text{-}21)$$

where u_t is a random variable with the usual properties.[11] For any given value of T, both τ and D are defined, and Eq. (8-21) is estimated by nonlinear least squares. Hence estimates of λ_1, λ_2, δ, and γ in Eq. (8-20*a* and *b*) are conditional on the assumed value of T. Unconditional estimates of λ_1, λ_2, δ, γ, and T are obtained by estimating Eq. (8-21) at all possible values of T and choosing the value of T (and associated values of other parameters) that minimizes the sum of squared residuals of X_t.

Obviously, T cannot be estimated with any tolerable degree of accuracy from the data of Figure 8-1, since sampling variation in X_t at older ages is so large. Therefore, Eq. (8-20*b*) cannot be estimated very well. However, it is clear that some parameters can be estimated with reasonable precision from Eq. (8-20*a*), the first portion of the earnings pattern. To investigate that possibility, I chose T in the neighborhood of actual age 66 for both groups and fitted the equation

$$X_t = b_0 + b_1\lambda_1{}^t + b_2\lambda_2{}^t + b_3[\lambda_1/(1+\gamma)]^t + b_4[\lambda_2/(1+\gamma)]^t + u_t \qquad t < T \quad (8\text{-}22)$$

to the right-hand portion of each profile.

It is important to recognize that estimation of Eq. (8-22) does not convey sufficient information to identify δ. However, Eq. (8-22) potentially identifies the real rate of interest, the vintage effect γ, and the parameter B [see Eq. (8-14)] for each education group. Furthermore, B identifies a function relating δ and the second derivatives F_{ij} of F [see Eq. (8-12)]. If $P(\bar{I})$ and $f(\bar{z}, \bar{h})$ are approximated by quadratic and log linear functions, respectively, the F_{ij} terms in Eq. (8-12) can be reduced to one additional parameter β. That is, B and the real rate of interest identify a *function* relating δ and the relative marginal product of knowledge β. Estimates of this function for high school and college graduates provide some interesting and informative between-group comparisons. It is important to emphasize that data limitations preclude estimation of the full

[11] The slightly unusual treatment of dummy variables in Eq. (8-21) is necessary to make the income-generating function continuous. D and τ are so defined that the two nonlinear segments link up at age T.

model [Eq. (8-20*a* and *b*)] at this time. In principle, all parameters are exactly identified, given a large-enough sample.

The income-generating function Eq. (8-22) has been estimated under two alternative specifications. Each is considered in turn.

First, assume that γ is small enough to be ignored: $\gamma \simeq 0$. Then the estimating Eq. (8-22) becomes

$$X_t = c_0 + c_1\lambda_1{}^t + c_2\lambda_2{}^t + u_t \tag{8-23}$$

where the c's and λ's are regression coefficients. The systematic portion of Eq. (8-23) can also be written as a second-order linear difference equation:

$$X_t - (\lambda_1 + \lambda_2)X_{t-1} + (\lambda_1\lambda_2)X_{t-2} = \text{constant}$$

Two methods are available for estimating λ_1 and λ_2 from the data in Figure 8-1: Estimate the linear difference equation equivalent to Eq. (8-23); and estimate Eq. (8-23) directly. The linear functional form is an advantage, but no satisfactory theory is available for estimating nonstationary difference equations. Therefore, I have chosen the second method. The computations are burdensome, but the standard nonlinear regression model applies to Eq. (8-23), and properties of residuals can be tested.

Rather than use packaged nonlinear regression programs, Eq. (8-23) has been estimated by artificially generating variables $\lambda_1{}^t$ and $\lambda_2{}^t$ for various numerical values of λ_1 and λ_2, regressing X on the constructed variables, and choosing the pair of values (λ_1, λ_2) that maximizes the coefficient of determination R^2. This allows examination of the likelihood surface for possible irregularities and problems of convergence to a global maximum. Values of R^2 at various values of λ_1 and λ_2 for high school graduates are shown in Table 8-1, using income data beginning at age 19 and ending at age 65 (i.e., $t =$ actual age minus 19.0, and $T = 47$). Intervals of .05 were chosen on the grid (λ_1, λ_2), and no experimentation was made with finer grids in the neighborhood of the maximum. No attempt was made to compute the information matrix at the maximum, and unconditional standard errors are unavailable. Standard errors, given in the note for the equation at the maximum, are conditional on the maximum likelihood estimate of (λ_1, λ_2). Maximum R^2 occurs at the combination (.85, 1.45) and is close to the maximum likelihood estimate of λ_1 and λ_2. Examination

TABLE 8-1 ***Coefficients of determination (R^2) for earnings regressions of high school graduates (ages 19 to 65)***

λ_1 \ λ_2	1.20	1.15	1.10	1.05	.95	.90	.85	.80	.75	.70
1.05	.5640	.6174	.6880		.9192	.9298	.9022	.8553	.8026	.7507
1.10	.4240	.4861		.6880	.9065	.9380	.9165	.8669	.8076	.7477
1.15	.3292		.4861	.6174	.8925	.9420	.9274	.8784	.8170	.7539
1.20		.3292	.4240	.5640	.8797	.9435	.9349	.8875	.8257	.7616
1.25	.2265	.2855	.3794	.5238	.8686	.9434	.9398	.8942	.8327	.7683
1.30	.1975	.2539	.3462	.4927	8590	.9424	.9429	.8988	.8378	.7735
1.35	.1761	.2302	.3206	.4681	.8505	.9408	.9446	.9020	.8415	.7774
1.40	.1595	.2116	.3002	.4479	.8429	.9389	.9455	.9040	.8442	.7803
1.45	.1463	.1966	.2834	.4310	.8360	.9367	.9457	.9053	.8460	.7824
1.50	.1354	.1841	.2693	.4166	.8298	.9345	.9454	.9061	.8472	.7838
1.55	.1263	.1736	.2573	.4041	.8241	.9321	.9449	.9064	.8480	.7849
1.60	.1186	.1646	.2470	.3932	.8188	.9298	.9441	.9064	.8484	.7855
1.65	.1119	.1567	.2378	.3835	.8139	.9275	.9432	.9062	.8486	.7859
1.70	.1060	.1498	.2298	.3749	.8094	.9253	.9423	.9059	.8486	.7861
1.75	.1009	.1437	.2226	.3671	.8052	.9232	.9412	.9054	.8484	.7862
1.80	.0963	.1383	.2162	.3601	.8013	.9211	.9401	.9048	.8482	.7861
1.85					.7977	.9191	.9391	.9042	.8479	.7860
1.90					.7943	.9172	.9380	.9036	.8475	.7858
1.95					.7911	.9154	.9369	.9030	.8471	.7855
2.00					.7882	.9137	.9359	.9023	.8466	.7852

NOTE: For definition of symbols, see text. At the maximum, the equation is

$$X_t = 6311.59 - \underset{(242.47)}{6692.94}\,(.85)^t - \underset{(.0674)}{.5305}(1.45)^t(10)^{-5}$$

Conditional standard errors are in parentheses. Durbin-Watson statistic = 2.20.

SOURCE: Author's computation.

of residuals at (.85, 1.45) did not reveal the presence of serial correlation with respect to age (Durbin-Watson statistic = 2.20).

As might be anticipated, Eq. (8-23) was much more difficult to estimate for college graduates because of greater variation in mean earnings. Beginning of work life was chosen to be 23 years of age, and the data were cut off at age 64. Age 65 was not chosen in this case because earnings at that age were unusually small as a result of a large increase in the proportion out of the labor force in the sample data. When the above procedure was straightforwardly applied, R^2 statistics did not converge to a maximum, undoubtedly because of ill-behaved residuals. Therefore, a second-order approximation to the income-generating function was tried, adding terms in λ_1^{2t}, λ_2^{2t}, and $(\lambda_1\lambda_2)^t$ to Eq. (8-23). Convergence was achieved, but the Durbin-Watson statistic was estimated at 3.00, revealing the presence of negative serial correlation in the residuals. To account for serially correlated disturbances, X_t and the λ^t terms were transformed by $Q_t^* = Q_t - aQ_{t-1}$, where Q is the original variable and a is an estimate of first-order serial correlation. a was estimated to be -0.5 from the first-stage regression. Finally, transformed data were used to compute Eq. (8-23), and results are shown in Table 8-2. The likelihood surface is smooth, and maximum likelihood estimates of (λ_1, λ_2) are in the neighborhood of (.90, 1.20). The Durbin-Watson statistic indicates no remaining serial correlation.

An attempt was made to estimate Eq. (8-22) without the a priori restriction $\gamma = 0$. The technique is the same as described above, except that computations must be carried out on a three-dimensional grid $(\lambda_1, \lambda_2, \gamma)$.

Since true values of γ are undoubtedly small, Eq. (8-22) in its complete form is extremely difficult to estimate. For example, if γ is as large as .05, h_0 $(N - \tau)$ doubles every 15 years—surely an enormous rate of increase. There is another reason for expecting the estimate of γ to be small. It is commonly argued that educational institutions serve as a filtering device for sorting individuals according to "ability." There is no way of knowing from income data alone whether individuals achieving a specified level of education in successive generations have been drawn from the same percentile of the ability distribution. The "filter content" of any level of school achievement may be getting coarser over time, and the estimate of γ reflects filter effects as well as improvements in knowledge. In any event, for plausible values of γ, collinearity between variables λ_i^t and $[\lambda_i/(1 + \gamma)]^t$ is so high that they cannot be distinguished from one another. For example, examining the grid point (.85, 1.45, .01), zero-order correlation coefficients are .9996 for variables involving λ_1 and .9999 for variables involving λ_2! Orthogonalizing the independent variables will not help either. Therefore, the interesting parameter γ is potentially identified, but unfortunately cannot be estimated in this way. No doubt cross sections for several different years would be helpful

TABLE 8-2 ***Coefficients of determination (R^2) for earnings regressions of college graduates (transformed data; ages 23 to 64)***

λ_2 \ λ_1	1.20	1.15	1.10	1.05	.95	.90	.85	.80	.75	.70
1.05	.6412	.6819	.7317		.8626	.8622	.8407	.8089	.7746	.7415
1.10	.5255	.5800		.7317	.8569	.8636	.8378	.7960	.7500	.7054
1.15	.4350		.5800	.6819	.8501	.8642*	.8365	.7885	.7351	.6833
1.20		.4350	.5255	.6412	.8434	.8642	.8360	.7844	.7266	.6706
1.25	.3246	.3901	.4844	.6093	.8376	.8639	.8359	.7821	.7217	.6633
1.30	.2922	.3574	.4536	.5845	.8327	.8635	.8358	.7808	.7188	.6587
1.35	.2688	.3332	.4302	.5652	.8286	.8631	.8359	.7800	.7169	.6557
1.40	.2515	.3149	.4122	.5501	.8253	.8627	.8359	.7795	.7156	.6357
1.45	.2385	.3009	.3981	.5380	.8226	.8623	.8359	.7791	.7146	.6522
1.50	.2281	.2899	.3868	.5282	.8204	.8621	.8359	.7788	.7139	.6510
1.55	.2201	.2811	.3778	.5203	.8186	.8618	.8359	.7786	.7133	.6501
1.60	.2137	.2740	.3704	.5137	.8171	.8617	.8360	.7784	.7128	.6493
1.65	.2085	.2682	.3643	.5083	.8159	.8615	.8360	.7783	.7124	.6487
1.70	.2042	.2635	.3592	.5037	.8149	.8614	.8360	.7781	.7121	.6481
1.75	.2007	.2595	.3550	.4998	.8140	.8614	.8361	.7780	.7118	.6476
1.80	.1978	.2561	.3513	.4965	.8133	.8613	.8361	.7779	.7115	.6472
1.85	.1953	.2532	.3482	.4937	.8128	.8613	.8361	.7778	.7113	.6468
1.90	.1932	.2508	.3455	.4912	.8123	.8613	.8362	.7778	.7110	.6464
1.95	.1913	.2487	.3432	.4891	.8119	.8614	.8362	.7777	.7108	.6461

* This apparent second maximum is due to rounding.

NOTE For definition of symbols, see text. All variables (i.e., X_t, $\lambda_1{}^t$, $\lambda_2{}^t$) transformed by $Q_t^* = Q_t + .5Q_{t-1}$, where Q_t is the original variable. At the maximum, the equation is

$$X_t = 17209.18 - \underset{(805.99)}{11769.50}(.90)^t - \underset{(.3112)}{.4822}(1.20)^t(10)^{-3}$$

Conditional standard errors are in parentheses. Durbin-Watson statistic = 2.02.

SOURCE: Author's computation.

here. Be that as it may, if γ is small, estimates in Tables 8-1 and 8-2 should be reasonable approximations to their specification in the model.

Though the true annual rate of growth γ may be small, power to discern intergenerational differences must increase for generations further apart in age. For example, if γ is .01, the difference in initial stocks of generations 30 years apart is on the order of one-third larger for the younger group. Therefore, as a second-best procedure, specify a step function

$$h_0(v) = \begin{cases} \eta_1 & \text{for } 0 \le v \le m \\ \eta_2 & \text{for } m < v \le N \end{cases}$$

or, equivalently, $h_0(v) = \eta_2 + (\eta_1 - \eta_2)d$, where d is a dummy variable with value 0 if $v > m$ and value 1 otherwise. Substitution into Eq. (8-20*a* and *b*) and (8-22) yields

$$X_t = b_{10} + b_{11}\lambda_1{}^t + b_{12}\lambda_2{}^t + b_{13}(d\lambda_1)^t + b_{14}(d\lambda_2)^t \quad (8\text{-}24)$$

Equation (8-24) has been estimated for high school graduates, with $m = 15$, $d = 1$ for persons having 15 years of experience or less in the cross section, and zero otherwise, distinguishing individuals who received their diplomas after World War II from those who graduated prior to the end of the war. An additional reason for making the break around 1945 is that opportunities for investment in the labor market were substantially different before and after that date. The hypothesis $\eta_1 > \eta_2$ requires estimates of b_{13} and b_{14} to be positive. The estimated equation for high school graduates at maximum R^2 (.85, 1.40) is

$$\begin{aligned} X_t = 6341.5 &- \underset{(2543.9)}{5655.8}\lambda_1{}^t - \underset{(.353)}{2.7358}\lambda_2{}^t(10^{-5}) \\ &- \underset{(2466)}{1003.4}(d\lambda_1)^t - \underset{(.226)}{.4359}(d\lambda_2)^t(10^{-5}), \quad R^2 = .9509 \end{aligned}$$

(Conditional standard errors are in parentheses.) Since the coefficients b_{13} and b_{14} are negative, the hypothesis of no vintage effects for high school graduates cannot be rejected.

INTERPRETATION OF RESULTS

According to condition (8-14), $\lambda_1\lambda_2$ estimates the net rate of discount $(1 + r)/(1 + \rho)$ for each group. Apply the estimates of Tables 8-1 and 8-2 to obtain

$(1 + r)/(1 + \rho) = (.85)(1.45) = 1.2325$ for high school graduates
$(1 + r)/(1 + \rho) = (.90)(1.20) = 1.08$ for college graduates

or net discount rates of about 23 percent and 8 percent for high school and college graduates, respectively. Real income per capita rose at an annual rate of about 2 percent from 1920 to 1960. Hence ρ cannot be larger than .02, and real gross discount rates no larger than 26 percent and 10 percent. In either case, the estimates indicate lower real rates of interest for college graduates, as would be expected from the theory of human capital.

The most thorough internal rate-of-return estimates from 1959 earnings data are presented by Hanoch (1967). He reports marginal internal rates of return of about 17 percent for high school graduates and 7 percent for college graduates. It is remarkable that the estimates above, which have been derived by entirely different methods, are so close to his. This surely strengthens confidence in the present approach. Hanoch uses the standard discounted comparison of income streams at two different levels of schooling, including earnings during schooling periods. My estimates are derived only from the shape of lifetime earnings patterns *within* groups, and do not rely on earnings during school. The estimates above are "internal" to each group and avoid all questions of comparability regarding ability differences between graduates at different levels of schooling. At face value, comparison of the two sets of estimates suggests that adjustment of estimated internal rates of return for ability differences is not too important, a conclusion consistent with independent investigations of that question (Becker, 1964). The estimates also suggest that rates of return to formal schooling are not very different from rates of return to learning in the labor market.

λ_1 and λ_2 also identify a parameter B in Eq. (8-14), but further interpretation requires additional assumptions. If the learning function can be approximated by $z = \alpha I^{1-\beta}\ h^{\beta}$ and if equalizing wage differences are quadratic, condition (8-16) applies, and λ_1 and λ_2 identify a function relating β, the relative marginal product of knowledge, and δ, depreciation and obsolescence.[12] Denote this function by $G(\delta,\beta) = 0$. Then $G(\delta,\beta)$ is the locus of pairs (δ,β) that all result in the same realized age-earnings pattern (for the phase $t < T$). In a sense, $G(\delta,\beta)$ approximates an "isoquant." However, the entire lifetime earnings pattern, rather than an annual flow, is held constant.

β is an index of efficiency of embodied knowledge in creating new knowledge, relative to other requisites for learning. An additional unit of knowledge enables individuals with higher values of β to

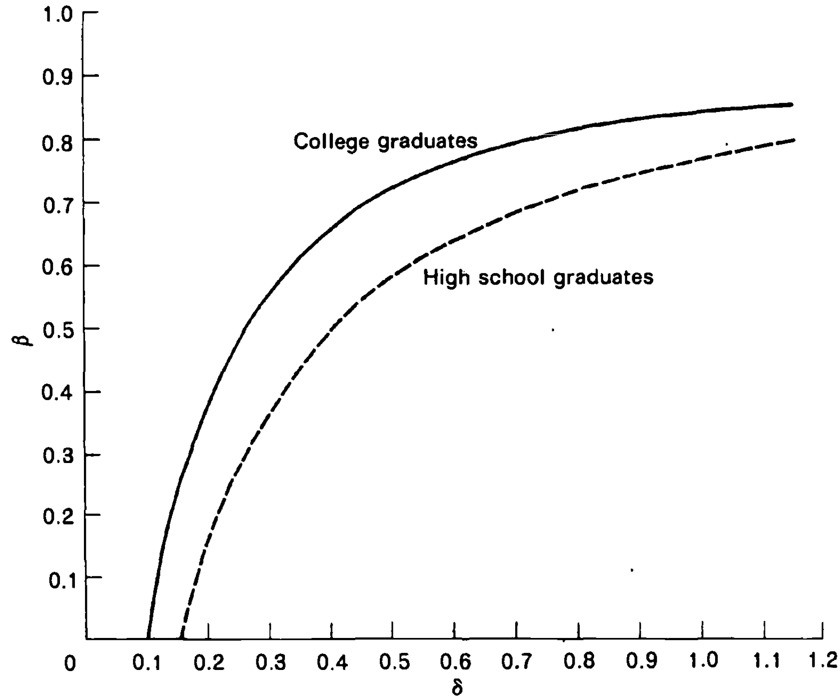
[12] The general "ability" parameter α is not identified and cannot be estimated from income data alone.

choose smaller and less costly learning opportunities to achieve a given amount of learning. Therefore, β is an index of the extent to which prior knowledge affects present learning capacity. Variations in β across different groups of individuals may be inherent either in the individuals themselves or in the nature of their work activities. In the first case, β represents an index of learning capacity. For example, higher levels of formal education may teach an individual how to use his acquired knowledge to learn new tasks more efficiently. In the second case, higher values of β are not in any sense "superior" to lower values. Production functions differ according to the product produced. If high school and college graduates systematically engage in different types of production activities and if their skills are truly different from one another and constitute different factors of production, there is no reason why their learning production functions should not differ.

Values of $G(\delta,\beta)$ implicit in the estimates of Tables 8-1 and 8-2 are shown in Figure 8-2. The functions are positively inclined: larger values of learning capacity must be offset by greater depreciation-obsolescence rates to achieve a constant age-earnings profile. Any personal disadvantage of knowledge deteriorating at a greater rate must be compensated by greater relative learning efficiency of embodied knowledge. Otherwise, earnings profiles could not remain unchanged.

Unfortunately, my inability to estimate the upper tail of the earnings pattern precludes identifying δ. Therefore, it is not possible to state where the true values of δ and β lie within the constraints imposed by Figure 8-2. However, Figure 8-2 suggests lower bounds for depreciation-obsolescence rates in the neighborhood of 10 percent for college graduates and 15 percent for high school graduates. I am unaware of comparable estimates in the literature, but these numbers appear rather high, especially when compared with most types of fixed capital. Based on the 10 and 15 percent bounds, corresponding half-lives of a "unit" of knowledge are at most 6.6 and 4.3 years, respectively, and implausibly small, a priori. Of course it is always possible that the quadratic and geometric mean assumptions on which Figure 8-2 is based are not tenable. Inadequacies of data may bias the levels of the curves in Figure 8-2 even if they are tenable. Intuitively, these biases should affect both groups more or less equally. Thus the finding that $G(\delta,\beta)$ is uniformly higher for college graduates than for high school graduates may have greater validity than inferences regarding intercepts.

FIGURE 8-2 *Implicit functions,* $G(\delta, \beta) = 0$ (based on estimates in Tables 8-1 and 8-2)



The curves could very well cross in principle, and the fact that they do not is an empirical matter. Therefore, if high school and college graduates are truly subject to the same rate of depreciation-obsolescence, then college graduates are more efficient users of prior knowledge in acquiring more knowledge. If the two groups are of equal learning efficiency, high school graduates are subject to greater rates of depreciation-obsolescence.

CONCLUSION In this chapter, a very general class of income-generating functions of the form $y = Rh - F(z,h)$ has been analyzed on the basis of the hypothesis that people learn from work experience. Applying the principles of optimum accumulation yields restrictions on the evolution of embodied knowledge and learning over working lifetimes and permits transformation of unobserved knowledge and learning components into specific functions of working experience. A method for estimating depreciation-obsolescence rates and other interesting parameters has been derived from a nonlinear, reduced-form relationship in the model, relating earnings to years of work experience. A portion of the reduced-form function was fitted to 1960 census earnings data. Since limitations of sample size precluded estimating the entire age-earnings function, not all potentially identifiable parameters in the model could be estimated at this time. However, some parameters were estimated, and the results are sufficiently promising to indicate feasibility of the method.

Real rates of interest implicit in income profiles have been estimated for white male high school and college graduates. Estimated values are 23 percent and 8 percent, respectively, and compare well with internal rates of return described in the literature. This comparison and the fact that they are based on an entirely different methodology are strong evidence in favor of the applicability of the model.

Though the form of the model actually estimated is limited, still it is sufficient to identify a vintage effect measuring intergenerational growth of knowledge obtained from schooling. No evidence of positive vintage effects for high school graduates was found, but this result cannot be taken as an indication that the quality of secondary schools has not improved over the years. Value added for high school education may very well have increased over time, even though gross output did not, because of offsetting changes in the role of high schools as institutes of certification. High school graduates may have been drawn from successively lower percentiles of the ability distribution, high school diplomas exhibiting decreasing "filter-year" content over time. More direct evidence on the point is available in the recent study by Taubman and Wales (1971). The results presented here are consistent with that study, though again they are based on different methods.

More tentative conclusions can be derived if learning functions are approximately of the form $z = \alpha I^{1-\beta} h^{\beta}$, where I is an index of learning opportunities implicit in work activities. In this formulation β measures the relative marginal product of knowledge in producing learning, whereas α is a measure of all-around ability. Depreciation-obsolescence rates (δ) have not been identified, since only a portion of age-earnings functions could be estimated. Measures such as β and δ cannot be distinguished on the basis of truncated earnings functions estimated here, since many combinations of these parameters are consistent with observed earnings. Only implicit functions $G(\beta,\delta)$ giving all possible β-δ combinations consistent with observed income patterns for high school and college graduates are identified, and they suggest lower bounds on δ of

.10 for college and .15 for high school graduates. It must be stressed that this identification problem is due to a limitation of data and not method. Both parameters are potentially identified from the full model.

For further interpretation, consider the following conceptual experiment. Suppose $G(\beta,\delta)$ were the same for both high school and college graduates. Then any real differences in learning capacity (in the sense of β) between them would be exactly offset by opposite differences in depreciation-obsolescence. If it were possible to "give" typical high school graduates the true (β,δ) combination actually possessed by typical college graduates, truncated income patterns of high school graduates would be practically identical to what is observed. The estimates above indicate larger values of β at *every* possible depreciation-obsolescence rate for college graduates. Therefore, the consequences of depreciation-obsolescence are more severe for high school graduates, for they cannot overcome them as readily. No matter what the true values of β and δ, the result suggests that college graduates are more efficient learners *relative* to their depreciation-obsolescence rates than high school graduates. College graduates' greater learning capacity more than offsets whatever true value of δ they face and it is one factor leading to greater lifetime earnings. Lifetime earnings of college graduates are also larger because they face lower real rates of interest. In addition, it is probable that they enter the market with greater initial knowledge (h_0) and possess greater all-around ability in the sense of α. Needless to say, if this characterization is correct, the model cannot distinguish between the differences in earnings resulting from college education as a selection process and as a real producer of embodied knowledge.

It is apparent that firmer conclusions will require data based on larger samples. The sensitivity of the results to other earnings concepts and other expectations hypotheses also remains to be investigated. Research currently under way will provide some better answers.

References

Becker, G. S.: *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education,* National Bureau of Economic Research, New York, 1964.

Ben-Porath, Y.: "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy,* vol. 75, pp. 352–365, August 1967.

Bowman, M. J., et al.: *Readings in the Economics of Education,* UNESCO, New York, 1968.

Cagan, P.: "Measuring Quality Changes and the Purchasing Power of Money: An Exploratory Study of Automobiles," in Zvi Griliches (ed.), *Price Indexes and Quality Change: Studies in New Methods of Measurement,* Harvard University Press, Cambridge, Mass., 1971, pp. 215–239.

Coleman, J. S., et al.: *Equality of Educational Opportunity,* U.S. Office of Education, Washington, 1966.

Eckaus, R. S., and A. El-Safty: "An Approach to the Estimation of Income-Educational Relations Based on Learning Theory," M.I.T., Department of Economics, 1972. (Mimeographed.)

Fisher, F. M.: "Embodied Technical Change and the Existence of an Aggregate Capital Stock," *Review of Economic Studies,* vol. 32, pp. 263–288, October 1965.

Griliches, Z.: "Hedonic Price Indexes Revisited," *Proceedings of the American Statistical Association Business and Economics Section,* Washington, D.C., 1967

Hall, R.: "Technical Change and Capital from the Point of View of the Dual," *Review of Economic Studies,* vol. 35, pp. 35–46, January 1968.

Hall, R.: "The Measurement of Quality Change from Vintage Price Data," in Zvi Griliches (ed.), *Price Indexes and Quality Change,* Harvard University Press, Cambridge, Mass., 1971, pp. 240–271.

Hanoch, G.: "An Economic Analysis of Earnings and Schooling," *Journal of Human Resources,* vol. 2, pp. 310–329, Summer 1967.

Johnson, T.: "Returns from Investment in Human Capital," *American Economic Review,* vol. 60, pp. 546–560, September 1970.

Lovell, M.: "The Production of Economic Knowledge," *Journal of Economic Literature,* vol. 11, pp. 27–55, March 1973.

Mincer, J.: "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach," *Journal of Economic Literature,* vol. 8, pp. 1–26, March 1970.

Rosen, S.: "Knowledge, Obsolescence and Income," abstracted in *Econometrica,* vol. 39, pp. 177–178, July 1971.

Rosen, S.: "Learning and Experience in the Labor Market," *Journal of Human Resources,* vol. 7, pp. 326–342, Summer 1972.

Solow, R. M.: "Investment and Technical Progress," in K. Arrow, S. Karlin, and P. Suppes (eds.), *Mathematical Methods in the Social Sciences, 1959: Proceedings of the First Stanford Symposium,* Stanford University Press, Stanford, Calif., 1960, pp. 89–104.

Taubman, P., and T. Wales: *Economic Returns to Higher Education,* National Bureau of Economic Research, New York, 1971. (Mimeographed.)