On the Concept of Health Capital and the Demand for Health

Michael Grossman

National Bureau of Economic Research

The aim of this study is to construct a model of the demand for the commodity "good health." The central proposition of the model is that health can be viewed as a durable capital stock that produces an output of healthy time. It is assumed that individuals inherit an initial stock of health that depreciates with age and can be increased by investment. In this framework, the "shadow price" of health depends on many other variables besides the price of medical care. It is shown that the shadow price rises with age if the rate of depreciation on the stock of health rises over the life cycle and falls with education if more educated people are more efficient producers of health. Of particular importance is the conclusion that, under certain conditions, an increase in the shadow price may simultaneously reduce the quantity of health demanded and increase the quantity of medical care demanded.

I. Introduction

During the past two decades, the notion that individuals invest in themselves has become widely accepted in economics. At a conceptual level, increases in a person's stock of knowledge or human capital are assumed to raise his productivity in the market sector of the economy, where he produces money earnings, and in the nonmarket or household sector, where he produces commodities that enter his utility function. To realize

This paper is based on part of my Columbia University Ph.D. dissertation, "The Demand for Health: A Theoretical and Empirical Investigation," which will be published by the National Bureau of Economic Research. My research at the Bureau was supported by the Commonwealth Fund and the National Center for Health Services Research and Development (PHS research grant 2 P 01 HS 00451-04). Most of this paper was written while I was at the University of Chicago's Center for Health Administration Studies, with research support from the National Center for Health Services Research and Development (PHS grant HS 00080). A preliminary version of this paper was presented at the Second World Congress of the Econometric Society. I am grateful to Gary S. Becker, V. K. Chetty, Victor R. Fuchs, Gilbert R. Ghez, Robert T. Michael, and Jacob Mincer for their helpful comments on earlier drafts.

potential gains in productivity, individuals have an incentive to invest in formal schooling or in on-the-job training. The costs of these investments include direct outlays on market goods and the opportunity cost of the time that must be withdrawn from competing uses. This framework has been used by Becker (1967) and by Ben-Porath (1967) to develop models that determine the optimal quantity of investment in human capital at any age. In addition, these models show how the optimal quantity varies over the life cycle of an individual and among individuals of the same age.

Although several writers have suggested that health can be viewed as one form of human capital (Mushkin 1962, pp. 129–49; Becker 1964, pp. 33–36; Fuchs 1966, pp. 90–91), no one has constructed a model of the demand for health capital itself. If increases in the stock of health simply increased wage rates, such a task would not be necessary, for one could simply apply Becker's and Ben-Porath's models to study the decision to invest in health. This paper argues, however, that health capital differs from other forms of human capital. In particular, it argues that a person's stock of knowledge affects his market and nonmarket productivity, while his stock of health determines the total amount of time he can spend producing money earnings and commodities. The fundamental difference between the two types of capital is the basic justification for the model of the demand for health that is presented in the paper.

A second justification for the model is that most students of medical economics have long realized that what consumers demand when they purchase medical services are not these services per se but, rather, "good health." Given that the basic demand is for good health, it seems logical to study the demand for medical care by first constructing a model of the demand for health itself. Since, however, traditional demand theory assumes that goods and services purchased in the market enter consumers' utility functions, economists have emphasized the demand for medical care at the expense of the demand for health. Fortunately, a new approach to consumer behavior draws a sharp distinction between fundamental objects of choice—called "commodities"—and market goods (Becker 1965; Lancaster 1966; Muth 1966; Michael 1969; Becker and Michael 1970; Ghez 1970). Thus, it serves as the point of departure for my health model. In this approach, consumers *produce* commodities with inputs of market goods and their own time. For example, they use traveling time and transportation services to produce visits; part of their Sundays and church services to produce "peace of mind"; and their own time, books, and teachers' services to produce additions to knowledge. Since goods and services are inputs into the production of commodities, the demand for these goods and services is a derived demand.

Within the new framework for examining consumer behavior, it is assumed that individuals inherit an initial stock of health that depreciates over time—at an increasing rate, at least after some stage in the life cycle—and can be increased by investment. Death occurs when the stock falls below a certain level, and one of the novel features of the model is that individuals "choose" their length of life. Gross investments in health capital are produced by household production functions whose direct inputs include the own time of the consumer and market goods such as medical care, diet, exercise, recreation, and housing. The production function also depends on certain "environmental variables," the most important of which is the level of education of the producer, that influence the efficiency of the production process.

It should be realized that in this model the level of health of an individual is *not* exogenous but depends, at least in part, on the resources allocated to its production. Health is demanded by consumers for two reasons. As a consumption commodity, it directly enters their preference functions, or, put differently, sick days are a source of disutility. As an investment commodity, it determines the total amount of time available for market and nonmarket activities. In other words, an increase in the stock of health reduces the time lost from these activities, and the monetary value of this reduction is an index of the return to an investment in health.

Since the most fundamental law in economics is the law of the downward-sloping demand curve, the quantity of health demanded should be negatively correlated with its shadow price. The analysis in this paper stresses that the shadow price of health depends on many other variables besides the price of medical care. Shifts in these variables alter the optimal amount of health and also alter the derived demand for gross investment, measured, say, by medical expenditures. It is shown that the shadow price rises with age if the rate of depreciation on the stock of health rises over the life cycle and falls with education if more educated people are more efficient producers of health. Of particular importance is the conclusion that, under certain conditions, an increase in the shadow price may simultaneously reduce the quantity of health demanded and increase the quantity of medical care demanded.

II. A Stock Approach to the Demand for Health

A. The Model

Let the intertemporal utility function of a typical consumer be

$$U = U(\phi_0 H_0, \ldots, \phi_n H_n, Z_0, \ldots, Z_n), \tag{1}$$

where H_0 is the inherited stock of health, H_i is the stock of health in the ith time period, ϕ_i is the service flow per unit stock, $h_i = \phi_i H_i$ is total consumption of "health services," and Z_i is total consumption of another commodity in the *ith* period.[1] Note that, whereas in the usual intertemporal utility function n, the length of life as of the planning date, is fixed, here it is an endogenous variable. In particular, death takes place when $H_i = H_{\min}$. Therefore, length of life depends on the quantities of H_i that maximize utility subject to certain production and resource constraints that are now outlined.

By definition, net investment in the stock of health equals gross investment minus depreciation:

$$H_{i+1} - H_i = I_i - \delta_i H_i, \tag{2}$$

where I_i is gross investment and δ_i is the rate of depreciation during the ith period. The rates of depreciation are assumed to be exogenous, but they may vary with the age of the individual.[2] Consumers produce gross investments in health and the other commodities in the utility function according to a set of household production functions:

$$\begin{aligned} I_i &= I_i(M_i, TH_i; E_i), \\ Z_i &= Z_i(X_i, T_i; E_i). \end{aligned} \tag{3}$$

In these equations, M_i is medical care, X_i is the goods input in the production of the commodity Z_i, TH_i and T_i are time inputs, and E_i is the stock of human capital.[3] It is assumed that a shift in human capital changes the efficiency of the production process in the nonmarket sector of the economy, just as a shift in technology changes the efficiency of the production process in the market sector. The implications of this treatment of human capital are explored in Section IV.

It is also assumed that all production functions are homogeneous of degree 1 in the goods and time inputs. Therefore, the gross investment production function can be written as

$$I_i = M_i g(t_i; E_i), \tag{4}$$

where $t_i = TH_i/M_i$. It follows that the marginal products of time and medical care in the production of gross investment in health are

[1] The commodity Z_i may be viewed as an aggregate of all commodities besides health that enter the utility function in period i. For the convenience of the reader, a glossary of symbols may be found in Appendix B.

[2] In a more complicated version of the model, the rate of depreciation might be a negative function of the stock of health. The analysis is considerably simplified by treating this rate as exogenous, and the conclusions reached would tend to hold even if it were endogenous.

[3] In general, medical care is not the only market good in the gross investment function, for inputs such as housing, diet, recreation, cigarette smoking, and alcohol consumption influence one's level of health. Since these inputs also produce other commodities in the utility function, joint production occurs in the household. For an analysis of this phenomenon, see Grossman (1970, chap. 6). To emphasize the key aspects of my health model, I treat medical care as the most important market good in the gross investment function in the present paper.

$$\frac{\partial I_i}{\partial TH_i} = \frac{\partial g}{\partial t_i} = g',$$
$$\frac{\partial I_i}{\partial M_i} = g - t_i g'. \tag{5}$$

From the point of view of the individual, both market goods and own time are scarce resources. The goods budget constraint equates the present value of outlays on goods to the present value of earnings income over the life cycle plus initial assets (discounted property income):[4]

$$\Sigma \frac{P_i M_i + V_i X_i}{(1+r)^i} = \Sigma \frac{W_i TW_i}{(1+r)^i} + A_0. \tag{6}$$

Here P_i and V_i are the prices of M_i and X_i, W_i is the wage rate, TW_i is hours of work, A_0 is discounted property income, and r is the interest rate. The time constraint requires that Ω, the total amount of time available in any period, must be exhausted by all possible uses:

$$TW_i + TL_i + TH_i + T_i = \Omega, \tag{7}$$

where TL_i is time lost from market and nonmarket activities due to illness or injury.

Equation (7) modifies the time budget constraint in Becker's time model (Becker 1965). If sick time were not added to market and nonmarket time, total time would *not* be exhausted by all possible uses. My model assumes that TL_i is inversely related to the stock of health; that is, $\partial TL_i/\partial H_i < 0$. If Ω were measured in days ($\Omega = 365$ days if the year is the relevant period) and if ϕ_i were defined as the flow of healthy days per unit of H_i, h_i would equal the total number of healthy days in a given year.[5] Then one could write

$$TL_i = \Omega - h_i. \tag{8}$$

It is important to draw a sharp distinction between sick time and the time input in the gross investment function. As an illustration of this difference, the time a consumer allocates to visiting his doctor for periodic checkups is obviously not sick time. More formally, if the rate of depreciation were held constant, an increase in TH_i would increase I_i and H_{i+1} and would reduce TL_{i+1}. Thus, TH_i and TL_{i+1} would be negatively correlated.[6]

[4] The sums throughout this study are taken from $i = 0$ to n.

[5] If the stock of health yielded other services besides healthy days, ϕ_i would be a vector of service flows. This study emphasizes the service flow of healthy days because this flow can be measured empirically.

[6] For a discussion of conditions that would produce a positive correlation between TH_i and TL_{i+1}, see Section III.

By substituting for TW_i from equation (7) into equation (6), one obtains the single "full wealth" constraint:

$$\Sigma \frac{P_i M_i + V_i X_i + W_i(TL_i + TH_i + T_i)}{(1+r)^i} = \Sigma \frac{W_i \Omega}{(1+r)^i} + A_0 = R. \tag{9}$$

According to equation (9), full wealth equals initial assets plus the present value of the earnings an individual would obtain if he spent all of his time at work. Part of this wealth is spent on market goods, part of it is spent on nonmarket production time, and part of it is lost due to illness. The equilibrium quantities of H_i and Z_i can now be found by maximizing the utility function given by equation (1) subject to the constraints given by equations (2), (3), and (9).[7] Since the inherited stock of health and the rates of depreciation are given, the optimal quantities of gross investment determine the optimal quantities of health capital.

B. *Equilibrium Conditions*

First-order optimality conditions for gross investment in period $i-1$ are:[8]

$$\frac{\pi_{i-1}}{(1+r)^{i-1}} = \frac{W_i G_i}{(1+r)^i} + \frac{(1-\delta_i) W_{i+1} G_{i+1}}{(1+r)^{i+1}} + \dots + \frac{(1-\delta_i) \dots (1-\delta_{n-1}) W_n G_n}{(1+r)^n} + \frac{Uh_i}{\lambda} G_i + \dots + (1-\delta_i) \dots (1-\delta_{n-1}) \frac{Uh_n}{\lambda} G_n; \tag{10}$$

$$\pi_{i-1} = \frac{P_{i-1}}{g - t_{i-1} g'} = \frac{W_{i-1}}{g'} \tag{11}$$

The new symbols in these equations are: $Uh_i = \partial U / \partial h_i =$ the marginal utility of healthy days; $\lambda =$ the marginal utility of wealth; $G_i = \partial h_i / \partial H_i = -(\partial TL_i / \partial H_i) =$ the marginal product of the stock of health in the production of healthy days; and $\pi_{i-1} =$ the marginal cost of gross investment in health in period $i-1$.

[7] In addition, the constraint is imposed that $H_n \leq H_{\min}$.

[8] Note that an increase in gross investment in period $i-1$ increases the stock of health in all future periods. These increases are equal to

$$\frac{\partial H_i}{\partial I_{i-1}} = 1, \frac{\partial H_{i+1}}{\partial I_{i-1}} = (1-\delta_i), \dots, \frac{\partial H_n}{\partial I_{i-1}} = (1-\delta_i)(1-\delta_{i+1}) \dots (1-\delta_{n-1}).$$

For a derivation of equation (10), see Part A of the Mathematical Appendix.

Equation (10) simply states that the present value of the marginal cost of gross investment in period $i-1$ must equal the present value of marginal benefits. Discounted marginal benefits at age i equal

$$G_i\left[\frac{W_i}{(1+r)^i}+\frac{Uh_i}{\lambda}\right],$$

where G_i is the marginal product of health capital—the increase in the number of healthy days caused by a one-unit increase in the stock of health. Two monetary magnitudes are necessary to convert this marginal product into value terms, because consumers desire health for two reasons. The discounted wage rate measures the monetary value of a one-unit increase in the total amount of time available for market and nonmarket activities, and the term Uh_i/λ measures the discounted monetary equivalent of the increase in utility due to a one-unit increase in healthy time. Thus, the sum of these two terms measures the discounted marginal value to consumers of the output produced by health capital.

While equation (10) determines the optimal amount of gross investment in period $i-1$, equation (11) shows the condition for minimizing the cost of producing a given quantity of gross investment. Total cost is minimized when the increase in gross investment from spending an additional dollar on medical care equals the increase in gross investment from spending an additional dollar on time. Since the gross investment production function is homogeneous of degree 1 and since the prices of medical care and time are independent of the level of these inputs, the average cost of gross investment is constant and equal to the marginal cost.

To examine the forces that affect the demand for health and gross investment, it is useful to convert equation (10) into a slightly different form. If gross investment in period i is positive, then

$$\frac{\pi_i}{(1+r)^i}\dot{i}=\frac{W_{i+1}G_{i+1}}{(1+r)^{i+1}}+\frac{(1-\delta_{i+1})W_{i+2}G_{i+2}}{(1+r)^{i+2}}+\dots$$
$$\frac{(1-\delta_{i+1})\dots(1-\delta_{n-1})W_nG_n}{(1+r)^n}+\frac{Uh_{i+1}G_{i+1}}{\lambda}+\dots$$
$$+(1-\delta_{i+1})\dots(1-\delta_{n-1})\frac{Uh_nG_n}{\lambda}. \quad (12)$$

From (10) and (12),

$$\frac{\pi_{i-1}}{(1+r)^{i-1}}=\frac{W_iG_i}{(1+r)^i}+\frac{Uh_iG_i}{\lambda}+\frac{(1-\delta_i)\dot{\pi}_i}{(1+r)^i}.$$

Therefore,

$$G_i\left[W_i+\left(\frac{Uh_i}{\lambda}\right)(1+r)^i\right]=\pi_{i-1}(r-\tilde{\pi}_{i-1}+\delta_i), \quad (13)$$

where $\tilde{\pi}_{i-1}$ is the percentage rate of change in marginal cost between period $i - 1$ and period i.[9] Equation (13) implies that the undiscounted value of the marginal product of the optimal stock of health capital at any moment in time must equal the supply price of capital, $\pi_{i-1}(r - \tilde{\pi}_{i-1} + \delta_i)$. The latter contains interest, depreciation, and capital gains components and may be interpreted as the rental price or user cost of health capital.

Condition (13) fully determines the demand for capital goods that can be bought and sold in a perfect market. In such a market, if firms or households acquire one unit of stock in period $i - 1$ at price π_{i-1}, they can sell $(1 - \delta_i)$ units at price π_i at the end of period i. Consequently, $\pi_{i-1}(r - \tilde{\pi}_{i-1} + \delta_i)$ measures the cost of holding one unit of capital for one period. The transaction just described allows individuals to raise their capital in period i *alone* by one unit and is clearly feasible for stocks like automobiles, houses, refrigerators, and producer durables. It suggests that one can define a set of single-period flow equilibria for stocks that last for many periods.

In my model, the stock of health capital cannot be sold in the capital market, just as the stock of knowledge cannot be sold. This means that gross investment must be nonnegative. Although sales of health capital are ruled out, provided gross investment is positive, there exists a used cost of capital that in equilibrium must equal the value of the marginal product of the stock.[10] An intuitive interpretation of this result is that exchanges over time in the stock of health by an individual substitute for exchanges in the capital market. Suppose a consumer desires to increase his stock of health by one unit in period i. Then he must increase gross investment in period $i - 1$ by one unit. If he simultaneously reduces gross investment in period i by $(1 - \delta_i)$ units, then he has engaged in a transaction that raises H_i and H_i *alone* by one unit. Put differently, he has essentially rented one unit of capital from himself for one period. The magnitude of the reduction in I_i is smaller the greater the rate of depreciation, and its dollar value is larger the greater the rate of increase in marginal cost over time. Thus, the depreciation and capital gains components are as relevant to the user cost of health as they are to the user cost of any other durable. Of course, the interest component of user cost is easy to interpret, for if one desires to increase his stock of health rather than his stock of some other asset by one unit in a given period, $r\pi_{i-1}$ measures the interest payment he forgoes.[11]

[9] Equation (13) assumes $\delta_i \tilde{\pi}_{i-1} \simeq 0$.

[10] For similar conclusions with regard to nonsalable physical capital and with regard to a nonsalable stock of "goodwill" produced by advertising, see Arrow (1968) and Nerlove and Arrow (1962).

[11] In a continuous time model, the user cost of health capital can be derived in one step. If continuous time is employed, the term $\delta_i \tilde{\pi}_{i-1}$ does not appear in the user cost formula. The right-hand side of (13) becomes $\pi_i(r - \tilde{\pi}_i + \delta_i)$, where $\tilde{\pi}_i$ is the in-

A slightly different form of equation (13) emerges if both sides are divided by the marginal cost of gross investment:

$$\gamma_i + a_i = r - \tilde{\pi}_{i-1} + \delta_i. \tag{13'}$$

Here $\gamma_i = (W_i G_i)/\pi_{i-1}$ is the marginal monetary rate of return on an investment in health and

$$a_i = \left[\frac{\left(\frac{Uh_i}{\lambda}\right)(1+r)^i G_i}{\pi_{i-1}} \right]$$

is the psychic rate of return. In equilibrium, the total rate of return on an investment in health must equal the user cost of health capital in terms of the price of gross investment. The latter variable is defined as the sum of the real-own rate of interest and the rate of depreciation.

C. *The Pure Investment Model*

It is clear that the number of sick days and the number of healthy days are complements; their sum equals the constant length of the period. From equation (8), the marginal utility of sick time is $-Uh_i$. Thus, by putting healthy days in the utility function, one implicitly assumes that sick days yield *disutility*. If healthy days did not enter the utility function directly, the marginal monetary rate of return on an investment in health would equal the cost of health capital, and health would be solely an investment commodity.[12] In formalizing the model, I have been reluctant to treat health as pure investment because many observers believe the demand for it has both investment and consumption aspects (see, for example, Mushkin 1962, p. 131; Fuchs 1966, p. 86). But to simplify the remainder of the theoretical analysis and to contrast health capital with other forms of human capital, the consumption aspects of demand are ignored from now on.[13]

If the marginal utility of healthy days or the marginal disutility of sick days were equal to zero, condition (13′) for the optimal amount of health capital in period i would reduce to

$$\frac{W_i G_i}{\pi_{i-1}} = \gamma_i = r - \tilde{\pi}_{i-1} + \delta_i. \tag{14}$$

stantaneous percentage rate of change of marginal cost at age i. For a proof, see Part B of the Mathematical Appendix.

12 To avoid confusion, a note on terminology is in order. If health were entirely an *investment commodity*, it would yield monetary, but not utility, returns. Regardless of whether health is investment, consumption, or a mixture of the two, one can speak of a *gross investment function* since the commodity in question is a durable.

13 Elsewhere, I have used a pure consumption model to interpret the set of phenomena that are analyzed in Sections III and IV. In the pure consumption model, the marginal monetary rate of return on an investment in health is set equal to zero (see Grossman 1970, chap. 3).

Equation (14) can be derived explicitly by excluding health from the utility function and by redefining the full wealth constraint as[14]

$$R' = A_0 + \Sigma \frac{W_i h_i - \pi_i I_i}{(1+r)^i}. \tag{15}$$

Maximization of R' with respect to gross investment in periods $i - 1$ and i yields

$$\frac{\pi_{i-1}}{(1+r)^{i-1}} = \frac{W_i G_i}{(1+r)^i} + \frac{(1-\delta_i) W_{i+1} G_{i+1}}{(1+r)^{i+1}} + \ldots + \frac{(1-\delta_i) \ldots (1-\delta_{n-1}) W_n G_n}{(1+r)^n}, \tag{16}$$

$$\frac{\pi_i}{(1+r)^i} = \frac{W_{i+1} G_{i+1}}{(1+r)^{i+1}} + \frac{(1-\delta_{i+1}) W_{i+2} G_{i+2}}{(1+r)^{i+2}} + \ldots + \frac{(1-\delta_{i+1}) \ldots (1-\delta_{n-1}) W_n G_n}{(1+r)^n}. \tag{17}$$

These two equations imply that (14) must hold.

Figure 1 illustrates the determinations of the optimal stock of health capital at any age i. The demand curve *MEC* shows the relationship between the stock of health and the rate of return on an investment or the marginal efficiency of health capital, γ_1. The supply curve S shows the relationship between the stock of health and the cost of capital, $r - \tilde{\pi}_{i-1} + \delta_i$. Since the cost of capital is independent of the stock, the supply curve is infinitely elastic. Provided the *MEC* schedule slopes downward,

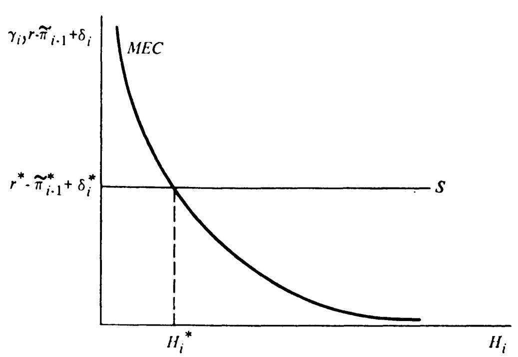

FIG. 1

14 Since the gross investment production function is homogeneous of the first degree, $P_i M_i + W_i TH_i = \pi_i I_i$.

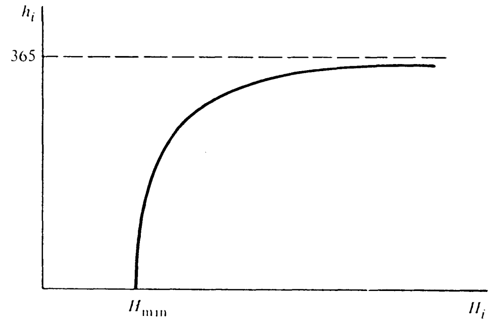

FIG. 2

the equilibrium stock is given by $H_i{}^*$, where the supply and demand curves intersect.

In the model, the wage rate and the marginal cost of gross investment do not depend on the stock of health. Therefore, the *MEC* schedule would be negatively inclined if and only if G_i, the marginal product of health capital, were diminishing. Since the output produced by health capital has a finite upper limit of 365 healthy days, it seems reasonable to assume diminishing marginal productivity. Figure 2 shows a plausible relationship between the stock of health and the number of healthy days. This relationship may be called the "production function of healthy days." The slope of the curve in the figure at any point gives the marginal product of health capital. The number of healthy days equals zero at the death stock $H_{\min}$, so that $\Omega = TL_i = 365$ is an alternative definition of death. Beyond $H_{\min}$, healthy time increases at a decreasing rate and eventually approaches its upper asymptote of 365 days as the stock becomes large.

In Sections III and IV, equation (14) and figure 1 are used to trace out the lifetime path of health capital and gross investment, to explore the effects of variations in depreciation rates, and to examine the impact of changes in the marginal cost of gross investment. Before I turn to these matters, some comments on the general properties of the model are in order. It should be realized that equation (14) breaks down whenever desired gross investment equals zero. In this situation, the present value of the marginal cost of gross investment would exceed the present value of marginal benefits for all positive quantities of gross investment, and equations (16) and (17) would be replaced by inequalities.[15] The remainder of the discussion rules out zero gross investment by assumption, but the conclusions reached would have to be modified if this were not the case. One justification for this assumption is that it is observed empirically that most individuals make positive outlays on medical care throughout their life cycles.

[15] Formally, $\gamma_i \leqq r - \tilde{\pi}_{i-1} + \delta_i$, if $I_{i-1} = I_i = 0$.

Some persons have argued that, since gross investment in health cannot be nonnegative, equilibrium condition (14) should be derived by using the optimal control techniques developed by Pontryagin and others. Arrow (1968) employs these techniques to analyze a firm's demand for nonsalable physical capital. Since, however, gross investment in health is rarely equal to zero in the real world, the methods I use—discrete time maximization in the text and the calculus of variations in the Mathematical Appendix—are quite adequate. Some advantages of my methods are that they are simple, easy to interpret, and familiar to most economists. In addition, they generate essentially the same equilibrium condition as the Pontryagin method. Both Arrow and I conclude that, if desired gross investment were positive, then the marginal efficiency of nonsalable capital would equal the cost of capital. On the other hand, given zero gross investment, the cost of capital would exceed its marginal efficiency.

The monetary returns to an investment in health differ from the returns to investments in education, on-the-job training, and other forms of human capital, since the latter investments raise wage rates.[16] Of course, the amount of health capital might influence the wage rate, but it necessarily influences the time lost from all activities due to illness or injury. To emphasize the novelty of my approach, I assume that health is not a determinant of the wage rate. Put differently, a person's stock of knowledge affects his market and nonmarket productivity, while his stock of health determines the total amount of time he can spend producing money earnings and commodities. Since both market time and nonmarket time are relevant, even individuals who are not in the labor force have an incentive to invest in their health. For such individuals, the marginal product of health capital would be converted into a dollar equivalent by multiplying by the monetary value of the marginal utility of time.

Since there are constant returns to scale in the production of gross investment and since input prices are given, the marginal cost of gross investment and its percentage rate of change over the life cycle are exogenous variables. In other words, these two variables are independent of the rate of investment and the stock of health. This implies that consumers reach their desired stock of capital immediately. It also implies that the stock rather than gross investment is the basic decision variable in the model. By this I mean that consumers respond to changes in the cost of capital by altering the marginal product of health capital and not the marginal cost of gross investment. Therefore, even though equation (14) is not independent of equations (16) and (17), it can be used to determine the optimal path of health capital and, by implication, the optimal path of gross investment.[17]

[16] This difference is emphasized by Mushkin (1962, pp. 132–33).

[17] This statement is subject to the modification that the optimal path of capital must always imply nonnegative gross investment.

Indeed, the major differences between my health model and the human capital models of Becker (1967) and Ben-Porath (1967) are the assumptions made about the behavior of the marginal product of capital and the marginal cost of gross investment. Both Becker and Ben-Porath assume that any one person owns only a small amount of the total stock of human capital in the economy. Therefore, the marginal product of his stock is constant. To rule out solutions in which the desired stock of capital is either zero or infinite, they postulate that the marginal cost of producing gross additions to the stock is positively related to the rate of gross investment. Since marginal cost rises, the desired stock of human capital is not reached immediately. Moreover, since the marginal product of capital is constant, gross investment is the basic decision variable in these models.[18] In my model, on the other hand, the marginal product of health capital falls because the output produced by this capital has a finite upper limit. Consequently, it is not necessary to introduce the assumption of rising marginal cost in order to determine the optimal stock.

To illustrate how the implications of the health and human capital models differ, suppose the rate of depreciation on either the stock of health or human capital rises. This upsets the equality between the cost of capital and its marginal efficiency. To restore this equality in the health model, the marginal product of health capital must rise, which would occur only if the stock of capital declines. To restore this equality in the human capital model, marginal cost must fall, which is possible only if gross investment declines.[19]

III. Life Cycle Variations in Depreciation Rates

Equation (14) enables one to study the behavior of the demand for health and gross investment over the life cycle. To simplify the analysis, it is assumed that the wage rate, the stock of knowledge, the marginal cost of gross investment, and the marginal productivity of health capital are independent of age. These assumptions are not as restrictive as they may seem. To be sure, wage rates and human capital are undoubtedly correlated with age, but the effects of shifts in these variables are treated in Section IV. Therefore, the results obtained in this section may be viewed as partial effects. That is, they show the impact of a pure increase in age on the demand for health, with all other variables held constant.

[18] For a complete discussion of these points, see Becker (1967, pp. 5–12) and Ben-Porath (1967, pp. 353–61). For models of the demand for physical capital by firms in which the marginal cost of investment and the amount of investment are positively correlated, see, for example, Eisner and Strotz (1963) and Gould (1968).

[19] Section III demonstrates that an increase in the rate of depreciation on health capital might cause gross investment to increase.

Since marginal cost does not depend on age, $\pi_{i-1} - 0$ and equation (14) reduces to

$$\gamma_i = r + \delta_i. \tag{18}$$

It is apparent from equation (18) that, if the rate of depreciation were independent of age, a single quantity of H would satisfy the equality between the marginal rate of return and the cost of health capital. Consequently, there would be no net investment or disinvestment after the initial period. One could not, in general, compare H_0 and H_1 because accumulation in the initial period would depend on the discrepency between the inherited stock and the stock desired in period 1. This discrepency in turn would be related to variations in H_0 and other variables across individuals. But, given zero costs of adjusting to the desired level immediately, H would be constant after period 1. Under the stated condition of a constant depreciation rate, individuals would choose an infinite life if they choose to live beyond period 1. In other words, if $H_1 > H_{\min}$, then H_i would always exceed the death stock.[20]

To permit the demand for health to vary with age, suppose the rate of depreciation depends on age. In general, any time path of δ_i is possible. For example, the rate of depreciation might be negatively correlated with age during the early stages of the life cycle. Again, the time path might be nonmonotonic, so that δ_i rises during some periods and falls during others. Despite the existence of a wide variety of possible time paths, it is extremely plausible to assume that δ_i is positively correlated with age after some point in the life cycle. This correlation can be inferred because, as an individual ages, his physical strength and memory capacity deteriorate. Surely, a rise in the rate of depreciation on his stock of health is merely one manifestation of the biological process of aging. Therefore, the analysis focuses on the effects of an increase in the rate of depreciation with age.

Since a rise in δ_i causes the supply curve of health capital to shift upward, it would reduce the quantity of health capital demanded over the life cycle. Graphically, an increase in the cost of capital from $r + \delta_i$ to $r + \delta_{i+1}$ in figure 3 reduces the optimal stock from H_i to H_{i+1}. The greater the elasticity of the *MEC* schedule, the greater the decrease in the optimal stock with age. Put differently, the slower the increase in the marginal product of health capital as H falls, the greater the decrease in the optimal stock.

Differentiation of equation (18) with respect to age quantifies the percentage rate of decrease in the stock of health over the life cycle:

$$\tilde{H}_i = -s_i \, \varepsilon_i \, \tilde{\delta}_i. \tag{19}$$

[20] The possibility that death can occur in period 1 is ruled out from now on.

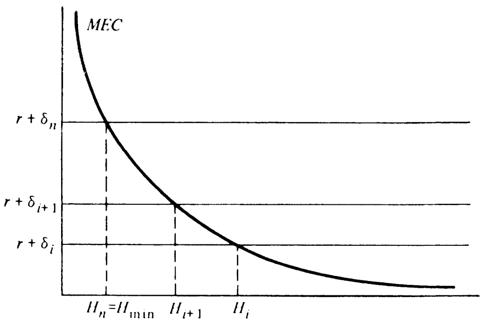

FIG. 3

In this equation, the tilde notation denotes a percentage time derivative ($\tilde{H}_i = (dH_i/di)\ (1/H_i)$, etc.), and the new symbols are: $s_i = \delta_i/r + \delta_i$ = the share of depreciation in the cost of health capital and

$$\varepsilon_i = -\frac{\partial \ln H_i}{\partial \ln(r+\delta_i)} = \frac{-\partial \ln H_i}{\partial \ln \gamma_i} = -\frac{\partial \ln H_i}{\partial \ln G_i} =$$

the elasticity of the MEC schedule (ln stands for natural logarithm).[21] Equation (19) indicates that the absolute value of the percentage decrease in H is positively related to the elasticity of the MEC schedule, the share of depreciation in the cost of health capital, and the percentage rate of increase in the rate of depreciation. If ε_i and $\tilde{\delta}_i$ were constant, the curve relating $\ln H_i$ to age would be concave unless $r = 0$, since[22]

$$\frac{d\tilde{H}_i}{di} = \tilde{H}_{ii} = -\, s_i(1 - s_i)\, \varepsilon\delta^2 < 0. \tag{20}$$

The absolute value of $\tilde{H}_i$ increases over the life cycle because depreciation's share in the cost of capital rises with age.

[21] From equation (18), $\ln(r + \delta_i) = \ln W + \ln G_i - \ln \pi$. Therefore,

$$\frac{\delta_i\, \tilde{\delta}_i}{r + \delta_i} = \frac{\partial \ln G_i}{\partial \ln H_i}\, \tilde{H}_i,$$

or

$$s_i\, \tilde{\delta}_i = -\frac{\tilde{H}_i}{\varepsilon_i}.$$

[22] Differentiation of (19) with respect to age yields

$$H_{ii} = \frac{\varepsilon\tilde{\delta}[(r + \delta_i)\delta_i\tilde{\delta} - \delta_i(\delta_i\tilde{\delta})]}{(r + \delta_i)^2},$$

or

$$\tilde{H}_{ii} = -\frac{\delta_i r \varepsilon \tilde{\delta}^2}{(r + \delta_i)^2} = -\, s_i(1 - s_i)\varepsilon\tilde{\delta}^2.$$

If δ_i grows continuously with age after some point in the life cycle, persons would choose to live a finite life. Since H declines over the life cycle, it would eventually fall to $H_{\min}$, the death stock. When the cost of health capital is $r + \delta_n$ in figure 3, $H_n = H_{min}$, and death occurs. At death, no time is available for market and nonmarket activities, since healthy time equals zero. Therefore, the monetary equivalent of sick time in period n would completely exhaust potential full earnings, $W_n\Omega$. Moreover, consumption of the commodity Z_n would equal zero, since no time would be available for its production if total time equals sick time.[23] Because individuals could not produce commodities, total utility would be driven to zero at death.[24]

Having characterized the optimal path of H_i, one can proceed to examine the behavior of gross investment. Gross investment's life cycle profile would not, in general, simply mirror that of health capital. In other words, even though health capital falls over the life cycle, gross investment might increase, remain constant, or decrease. This follows because a rise in the rate of depreciation not only reduces the amount of health capital *demanded* by consumers but also reduces the amount of capital *supplied* to them by a given amount of gross investment. If the change in supply exceeded the change in demand, individuals would have an incentive to close this gap by increasing gross investment. On the other hand, if the change in supply were less than the change in demand, gross investment would tend to fall over the life cycle.

To predict the effect of an increase in δ_i with age on gross investment, note that the net investment can be approximated by $H_i\tilde{H}_i$.[25] Since gross investment equals net investment plus depreciation,

$$\ln I_i = \ln H_i + \ln (\tilde{H}_i + \delta_i). \tag{21}$$

Differentiation of equation (21) with respect to age yields

$$\tilde{I}_i = \frac{\tilde{H}_i^2 + \delta_i\tilde{H}_i + \tilde{H}_{ii} + \delta_i\tilde{\delta}_i}{\tilde{H}_i + \delta_i}.$$

Suppose $\tilde{\delta}_i$ and ε_i were constant. Then from (19) and (20), the expression for $\tilde{I}_i$ would simplify to

[23] The above statement assumes that Z_i cannot be produced with X_i alone. This would be true if, say, the production function were Cobb-Douglas.

[24] Utility equal zero when $H = H_{\min}$ provided the death time utility function is such that $U(0) = 0$.

[25] That is,

$$H_{i+1} - H_i = H_i \frac{dH_i}{di}\frac{1}{H_i} = H_i\tilde{H}_i.$$

The use of this approximation essentially allows one to ignore the one-period lag between a change in gross investment and a change in the stock of health.

$$\tilde{I}_i = \frac{\tilde{\delta}(1 - s_i\varepsilon)(\delta_i - s_i\varepsilon\tilde{\delta}) + s_i^2\varepsilon\tilde{\delta}^2}{\delta_i - s_i\varepsilon\tilde{\delta}}. \tag{22}$$

Since health capital cannot be sold, gross investment cannot be negative. Therefore, $\delta_i \geqslant -\tilde{H}_i$.[26] That is, if the stock of health falls over the life cycle, the absolute value of the percentage rate of net disinvestment cannot exceed the rate of depreciation. Provided gross investment does not equal zero, the term $\delta_i - s_i\varepsilon\tilde{\delta}$ in equation (22) must exceed zero. It follows that a sufficient condition for gross investment to be positively correlated with the depreciation rate is $\varepsilon < 1/s_i$. Thus, $\tilde{I}_i$ would definitely be positive at every point if $\varepsilon < 1$.

The important conclusion is reached that, if the elasticity of the *MEC* schedule were less than 1, gross investment and the depreciation rate would be positively correlated over the life cycle, while gross investment and the stock of health would be negatively correlated. Phrased differently, given a relatively inelastic demand curve for health, individuals would desire to offset *part* of the reduction in health capital caused by an increase in the rate of depreciation by increasing their gross investments. In fact, the relationship between the stock of health and the number of healthy days suggests that ε is smaller than 1. A general equation for the healthy-days production function illustrated by figure 2 is

$$h_i = 365 - BH_i^{-C}, \tag{23}$$

where B and C are positive constants. The corresponding *MEC* schedule is[27]

$$\ln \gamma_i = \ln BC - (C + 1) \ln H_i + \ln W - \ln \pi. \tag{24}$$

The elasticity of this schedule is given by

$$\varepsilon = -\frac{\partial \ln H_i}{\partial \ln \gamma_i} = \frac{1}{(1 + C)} < 1,$$

since $C > 0$.

Observe that with the depreciation rate held constant, increases in gross investment would increase the stock of health and the number of healthy days. But the preceding discussion indicates that, because the

[26] Gross investment is nonnegative as long as $I_i = H_i(\tilde{H}_i + \delta_i) \geqslant 0$, or $\delta_i \geqslant -\tilde{H}_i$.

[27] If (23) were the production function, the marginal product of health capital would be

$$G_i = BCH_i^{-C-1},$$

or

$$\ln G^i = \ln BC - (C + 1)\ln H_i.$$

Since $\ln \gamma_i = \ln G_i + \ln W - \ln \pi$, one uses the equation for $\ln G_i$ to obtain (24).

depreciation rate rises with age, it is not unlikely that unhealthy (old) people will make larger gross investments than healthy (young) people. This means that sick time, TL_i, will be positively correlated with M_i and TH_i, the medical care and own time inputs in the gross investment function, over the life cycle.[28] In this sense, at least part of TL_i or TH_i may be termed "recuperation time."

Unlike other models of the demand for medical care, my model does not *assert* that "need" or illness, measured by the level of the rate of depreciation, will definitely be positively correlated with utilization of medical services. Instead, it derives this correlation from the magnitude of the elasticity of the *MEC* schedule and indicates that the relationship between the stock of health and the number of healthy days will tend to create a positive correlation. If ε is less than 1, medical care and "need" will definitely be positively correlated. Moreover, the smaller the value of ε, the greater the explanatory power of "need" relative to that of the other variables in the demand curve for medical care.

It should be realized that the power of this model of life cycle behavior is that it can treat the biological process of aging in terms of conventional economic analysis. Biological factors associated with aging raise the price of health capital and cause individuals to substitute away from future health until death is "chosen." It can be concluded that here, as elsewhere in economics, people reject a prospect—the prospect of longer life in this case—because it is too costly to achieve. In particular, only if the elasticity of the *MEC* schedule were zero would individuals fully compensate for the increase in δ_i and, therefore, maintain a constant stock of health.

IV. Market and Nonmarket Efficiency

Persons who face the same cost of health capital would demand the same amount of health only if the determinants of the rate of return on an investment were held constant. Changes in the value of the marginal product of health capital and the marginal cost of gross investment shift the *MEC* schedule and, therefore, alter the quantity of health demanded even if the supply curve of capital does not change. I now identify the variables that determine the level of the *MEC* schedule and examine the effects of shifts in these variables on the demand for health and medical care. In particular, I consider the effects of variations in market efficiency, measured by the wage rate, and nonmarket efficiency, measured by human capital, on the *MEC* schedule.

[28] Note that the time path of H_i or h_i would be nonmonotonic if the time path of δ_i were characterized by the occurrence of peaks and troughs. In particular, h_i would be relatively low and TH_i and M_i would be relatively high (if $\varepsilon < 1$) when δ_i was relatively high; these periods would be associated with relatively severe illness.

Before beginning the analysis, two preliminary comments are in order. First, the discussion pertains to uniform shifts in variables that influence the rate of return across persons of the same age. That is, if the variable X_i is one determinant, then

$$\frac{d \ln X_i}{d \ln X_{i-1}} = 1, \quad \text{all} \quad i.$$

Second, the discussion proceeds under the assumption that the real rate of interest, the rate of depreciation, and the elasticity of the *MEC* schedule are constant. These two comments imply that an increase in X_i will alter the amount of capital demanded but will not alter its rate of change over the life cycle.[29] Note from equation (21):

$$\frac{d \ln I}{dX} = \frac{d \ln H}{dX} \tag{25}$$

since the rate of depreciation and the percentage rate of net investment do not depend on X.[30] Equation (25) indicates that percentage changes in health and gross investment for a one-unit change in X are identical. Consequently, the effect of an increase in X on either of these two variables can be treated interchangeably.

A. Wage Effects

Since the value of the marginal product of health capital equals WG, an increase in the wage rate, W, raises the monetary equivalent of the marginal product of a given stock. Put differently, the higher a person's wage rate, the greater the value to him of an increase in healthy time. A consumer's wage rate measures his market efficiency or the rate at which he can convert hours of work into money earnings. Hence, it is obviously positively correlated with the benefits of a reduction in the time he loses from the production of money earnings due to illness. Moreover, a high wage rate induces an individual to substitute market goods for his own time in the production of commodities. This substitution continues until in equilibrium the monetary value of the marginal product of consumption time equals the wage rate. So the benefits from a reduction in time lost from nonmarket production are also positively correlated with the wage.

29 Strictly speaking, shifts in X_i would definitely have no effects on $\tilde{H}_i$ if and only if $\tilde{X}_i = 0$. Even though a uniform shift in X_i implies that there is no correlation between its level and rate of change, $\tilde{H}_i$ *might* be altered if $\tilde{X}_i \neq 0$. For a complete discussion of this point, see Grossmán (1970, p. 49).

30 Since the analysis in this section deals with variations in X among individuals of the same age, time subscripts are omitted from now on. Note also that (25), like the expression for I_i, ignores the one-period lag between an increase in gross investment and an increase in the stock of health.

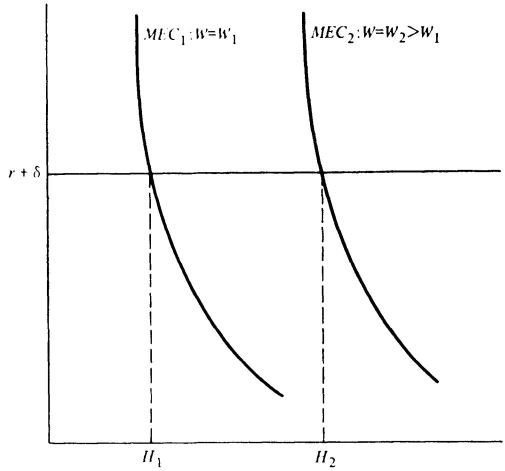

FIG. 4

If an upward shift in the wage rate had no effect on the marginal cost of gross investment, a 1 percent increase in it would increase the rate of return, γ, associated with a fixed stock of capital by 1 percent. In fact, this is not the case because own time is an input in the gross investment function. If K is the fraction of the total cost of gross investment accounted for by time, then a 1 percent rise in W would increase marginal cost, π, by K percent. After one nets out the correlation between W and π, the percentage growth in γ would equal $1 - K$, which exceeds zero as long as gross investment is not produced entirely by time.

Since the wage rate and the level of the MEC schedule are positively correlated, the demand for health would be positively related to W. Graphically, an upward shift in W from W_1 to W_2 in figure 4 shifts the MEC schedule from MEC_1 to MEC_2 and, with no change in the cost of health capital, increases the optimal stock from H_1 to H_2. A formula for the wage elasticity of health capital is[31]

$$e_{H,W} = (1 - K)\varepsilon. \tag{26}$$

This elasticity is larger the larger the elasticity of the MEC schedule and the larger the share of medical care in total gross investment cost.

Although the wage rate and the demand for health or gross invest-

[31] Differentiation of the natural logarithm of (18) with respect to $\ln W$ yields

$$\frac{d \ln(r+\delta)}{d \ln W} = 0 = 1 + \frac{\partial \ln G}{\partial \ln H}\frac{d \ln H}{d \ln W} - \frac{d \ln \pi}{d \ln W}$$

$$0 = 1 - K - \frac{e_{H,W}}{\varepsilon}.$$

ment are positively related, W has no effect on the amount of gross investment supplied by a given input of medical care. Therefore, the demand for medical care would rise with the wage. If medical care and own time were employed in fixed proportions in the gross investment production function, the wage elasticity of M would equal the wage elasticity of H. On the other hand, given a positive elasticity of substitution, M would increase more rapidly than H. This follows because consumers would have an incentive to substitute medical care for their relatively more expensive own time. A formula for the wage elasticity of medical care is

$$e_{M,W} = K\sigma_p + (1 - K)\varepsilon, \tag{27}$$

where σ_p is the elasticity of substitution between M and Th in the production of gross investment.[32] The greater the value of σ_p, the greater the difference between the wage elasticities of M and H.

Note that an increase in the price of either medical care or own time raises the marginal or average cost of gross investment. But the effects of changes in these two input prices are not symmetrical. In particular, an upward shift in the price of medical care lowers the *MEC* schedule and causes the demand for health to decline. This difference arises because the price of time influences the value of the marginal product of health capital while the price of medical care does not.

B. The Role of Human Capital

Up to now, no systematic allowance has been made for variations in the efficiency of nonmarket production. Yet it is known that firms in the market sector of an economy obtain varying amounts of output from the same vector of direct inputs. These differences have been traced to forces like technology and entrepreneurial capacity, forces that shift production functions or that alter the environment in which firms operate. Reasoning by analogy, one can say that certain environmental variables influence productivity in the nonmarket sector by altering the marginal products of the direct inputs in household production functions. This study is particularly concerned with environmental variables that can be associated with a particular person—his or her race, sex, stock of human capital, etc. While the analysis that follows could pertain to any environmental variable, it is well documented that the more educated are more efficient producers of money earnings. Consequently, it is assumed that shifts in human capital, measured by education, change productivity in

[32] For a proof, see Part C of the Mathematical Appendix. The corresponding equation for the wage elasticity of the own time input is

$$e_{TH,W} = (1 - K)(\varepsilon - \sigma_p).$$

This elasticity is positive only if $\varepsilon > \sigma_p$.

the household as well as in the market, and the analysis focuses on this environmental variable.

The specific proposition to be examined is that education improves nonmarket productivity. If this were true, then one would have a convenient way to analyze and quantify what have been termed the nonmonetary benefits to an investment in education. The model can, however, treat adverse as well as beneficial effects and suggests empirical tests to discriminate between the two.[33]

To determine the effects of education on production, marginal cost, and the demand for health and medical care, recall that the gross investment production function is homogeneous of degree 1 in its two direct inputs—medical care and own time. It follows that the marginal product of E, the index of human capital, would be

$$\frac{\partial I}{\partial E} = M \frac{\partial (g - tg')}{\partial E} + TH \frac{\partial g'}{\partial E},$$

where $g - tg'$ is the marginal product of medical care and g' is the marginal product of time.[34] If a circumflex over a variable denotes a percentage change per unit change in E, the last equation can be rewritten as

$$r_H = \frac{\partial I}{\partial E}\frac{1}{I} = \left[\frac{M(g - tg')}{I}\right]\left(\frac{g\hat{g} - tg'\hat{g}'}{g - tg'}\right) + \left(\frac{THg'}{I}\right)(\hat{g}'). \tag{28}$$

Equation (28) indicates that the percentage change in gross investment supplied to a consumer by a one-unit change in E is a weighted average of the percentage changes in the marginal products of M and TH.[35]

If E increases productivity, then $r_H > 0$. Provided E raises both marginal products by the same percentage, equation (28) would simplify to

$$r_H = \hat{g} = \hat{g}'. \tag{29}$$

[33] The model developed here is somewhat similar to the one used by Michael (1969).

[34] If I is homogeneous of degree 1 in M and TH, then from Euler's theorem

$$I = M(g - tg') + THg'.$$

Differentiation of this equation with respect to E, holding M and TH constant, yields the marginal product of human capital.

[35] Instead of putting education in the gross investment production function, one could let it affect the rate of depreciation or the marginal productivity of health capital. This approach has not been taken because a general treatment of environmental variables like education must permit these variables to influence all household commodities. Since depreciation rates and stock-flow relationships are relevant only if a particular commodity is durable, a symmetrical development of the role of environmental variables requires that they affect household production functions and not depreciation rates or stock-flow relationships. In a more complicated version of the model, the gross investment function, the rate of depreciation, and the marginal productivity of health capital might all depend on education. But the basic implications of the model would not change.

In this case, education would have a "neutral" impact on the marginal products of all factors. The rest of the discussion assumes "factor neutrality."

Because education raises the marginal product of the direct inputs, it reduces the quantity of these inputs required to produce a given amount of gross investment. Hence, with no change in input prices, an increase in E lowers average or marginal cost. In fact, one easily shows that

$$\hat{\pi} = -r_H = -\hat{g} = -\hat{g}', \tag{30}$$

where $\hat{\pi}$ is the percentage change in average or marginal cost.[36] So, if education increases the marginal products of medical care and own time by 3 percent, it would reduce the price of gross investment by 3 percent.

Suppose education does in fact raise productivity so that π and E are negatively correlated. Then, with the wage rate and the marginal product of a given stock of health held constant, an increase in education would raise the marginal efficiency of health capital and shift the *MEC* schedule to the right.[37] In figure 5, an increase in E from E_1 to E_2 shifts the *MEC* curve from MEC_1 to MEC_2. If the cost of capital were independent of E, there would be no change in the supply curve, and the more educated would demand a larger optimal stock (compare H_1 and H_2 in fig. 5).

The percentage increase in the amount of health demanded for a one-unit increase in E is given by[38]

$$\hat{H} = r_H \varepsilon. \tag{31}$$

Since r_H indicates the percentage increase in gross investment supplied by a one-unit increase in E, shifts in this variable would not alter the demand for medical care or own time if r_H equaled $\hat{H}$. For example, a person with ten years of formal schooling might demand 3 percent more health than a person with nine years. If the medical care and own time inputs were held constant, the former individual's one extra year of

[36] For a proof, see Part D of the Mathematical Appendix, where the human capital formulas are developed in more detail.

[37] It should be stressed that the model of nonmarket productivity variations presented here examines the *partial* effect of an increase in education with the wage rate held constant. Although these two variables are surely positively correlated, this correlation does not appear to be large enough to prevent one from isolating pure changes in nonmarket productivity at the empirical level. For some evidence on this point, see Grossman (1970, chap. 5) and Michael (1969, chaps. 4 and 5).

[38] If W and $r + \delta$ are fixed and if G depends only on H, then

$$\frac{d \ln(r+\delta)}{dE} = 0 = \frac{\partial \ln G}{\partial \ln H} \frac{d \ln H}{dE} - \frac{d \ln \pi}{dE},$$

or

$$0 = -\frac{\hat{H}}{\varepsilon} + r_H.$$

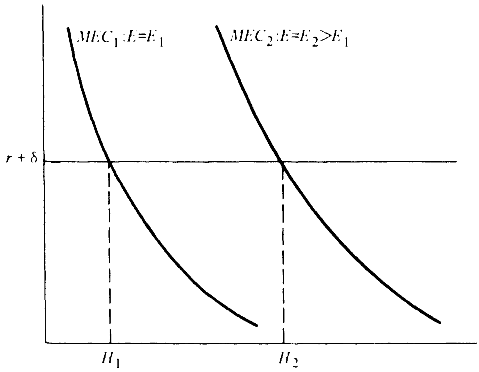

FIG. 5

schooling might supply him with 3 percent more health. Given this condition, both persons would demand the same amounts of M and TH. As this example illustrates, any effect of a change in E on the demand for medical care or time reflects a positive or negative difference between $\hat{H}$ and r_H:[39]

$$\hat{M} = T\hat{H} = r_H(\varepsilon - 1). \tag{32}$$

Equation (32) suggests that, if the elasticity of the MEC schedule were less than unity, the more educated would demand more health but less medical care. Put differently, they would have an incentive to offset *part* of the increase in health caused by an increase in education by reducing their purchases of medical services. Note that if r_H were negative and ε were less than 1, H would be negative and M would be positive. Since education improves market productivity, I have examined the implications of the hypothesis that r_H is positive. But the model is applicable whether r_H is positive or negative and gives empirical predictions in either case.

V. Summary and Conclusions

The main purpose of this paper has been to construct a model of the demand for the commodity "good health." The central proposition of the model is that health can be viewed as a durable capital stock that produces an output of healthy time. A person determines his optimal stock of health capital at any age by equating the marginal efficiency of this capital to its user cost in terms of the price of gross investment. Graphically, each person has a negatively inclined demand curve for health

[39] The terms M and TH are equal because, by the definition of factor neutrality, E has no effect on the ratio of the marginal product of M to the marginal product of TH.

capital, which relates the marginal efficiency of capital to the stock, and an infinitely elastic supply curve. The equilibrium stock is determined by the intersection of these two functions. The demand curve slopes downward due to diminishing marginal productivity of health capital.

Although in recent years there have been a number of extremely interesting explorations of the forces associated with health differentials (Adelman 1963; Fuchs 1965; Larmore 1967; Newhouse 1968; Auster, Leveson, and Sarachek 1969), these studies have not developed behavioral models that can predict the effects that are in fact observed. Consequently, the framework I have developed is important because of its ability to bridge the existing gap between theory and empiricism in the analysis of health differentials. My model explains variations in both health and medical care among persons in terms of variations in supply and demand curves for health capital. This paper has traced upward shifts in the supply curve to increases in the rate of depreciation on the stock of health with age, and it has traced upward shifts in the demand curve to increases in the wage rate and education.

One prediction of the model is that if the rate of depreciation increases with age, at least after some point in the life cycle, then the quantity of health capital demanded would decline over the life cycle. At the same time, provided the elasticity of the marginal efficiency of capital schedule were less than unity, expenditures on medical care would rise with age. A second prediction is that a consumer's demand for health and medical care should be positively correlated with his wage rate. A third prediction is that if education increases the efficiency with which gross investments in health are produced, then the more educated would demand a larger optimal stock of health. On the other hand, given a relatively inelastic demand curve, the correlation between medical outlays and education would be negative. It should be noted that one of the advantages of the model is that it enables one to study the effects of demographic variables like age and education without assuming that these variables are positively or negatively correlated with consumers' "tastes" for health. Instead, these variables enter the analysis through their impact on either the cost of health capital or its marginal efficiency, and one can make strong predictions concerning their effects on health levels or medical care.

It must be admitted that this paper has made a number of simplifying assumptions, all of which should be relaxed in future work. A more general model would treat the depreciation rate as an endogenous variable and would not rule out periods in which the optimal amount of gross investment is zero. Most important of all, it would modify the assumption that consumers fully anticipate intertemporal variations in depreciation rates and, therefore, know their age of death with certainty. Since in the real world length of life is surely not known with perfect foresight, it might be postulated that a given consumer faces a probability distribution of depreciation rates in each period. This uncertainty would give persons an incentive to protect themselves against the "losses" associated with higher than average depreciation rates by purchasing various types of insurance and perhaps by holding an "excess" stock of health.[40] But whatever modifications are made, it would be a mistake to neglect the essential features of the model I have presented in this paper. Any model must recognize that health is a durable capital stock, that health capital differs in important respects from other forms of human capital, and that the demand for medical care must be derived from the more fundamental demand for "good health."

Appendix A

Mathematical Appendix

A. Utility Maximization—Discrete Time

To maximize utility subject to the full wealth and production function constraints, form the Lagrangian expression

$$L = U(\phi_0 H_0, \ldots, \phi_n H_n, Z_0, \ldots, Z_n) + \lambda \left(R - \sum_i \frac{C_i + C_{1i} + W_i TL_i}{(1+r)^i} \right), \qquad \text{(A1)}$$

where $C_i = P_i M_i + W_i TH_i$ and $C_{1i} = V_i X_i + W_i T_i$. Differentiating L with respect to gross investment in period $i-1$ and setting the partial derivative equal to zero, one obtains

$$Uh_i \frac{\partial h_i}{\partial H_i} \frac{\partial H_i}{\partial I_{i-1}} + Uh_{i+1} \frac{\partial h_{i+1}}{\partial H_{i+1}} \frac{\partial H_{i+1}}{\partial I_{i-1}} + \ldots + Uh_n \frac{\partial h_n}{\partial H_n} \frac{\partial H_n}{\partial I_{i-1}} = \lambda \Bigg[\frac{(dC_{i-1}/dI_{i-1})}{(1+r)^{i-1}} + \frac{W_i(\partial TL_i/\partial H_i)/(\partial H_i/\partial I_{i-1})}{(1+r)^i} + \frac{W_{i+1}(\partial TL_{i+1}/\partial H_{i+1})/(\partial H_{i+1}/\partial I_{i-1})}{(1+r)^{i+1}} + \ldots + \frac{W_n(\partial TL_n/\partial H_n)/(\partial H_n/\partial I_{i-1})}{(1+r)^n} \Bigg]. \qquad \text{(A2)}$$

But

$$\frac{\partial h_i}{\partial H_i} = G_i, \frac{\partial H_i}{\partial I_{i-1}} = 1, \frac{\partial H_{i+1}}{\partial I_{i-1}} = (1 - \delta_i), \frac{\partial H_n}{\partial I_{i-1}}$$

[40] For an attempt to introduce uncertainty into a model that views health as a durable capital stock, see Phelps (in preparation).

$$= (1 - \delta_i) \ldots (1 - \delta_{n-1}), \frac{dC_{i-1}}{dI_{i-1}} = \pi_{i-1}, \quad \text{and} \quad \frac{\partial TL_i}{\partial H_i} = -G_i.$$

Therefore,

$$\frac{\pi_{i-1}}{(1+r)^{i-1}} = \frac{W_i G_i}{(1+r)^i} + \frac{(1-\delta_i) W_{i+1} G_{i+1}}{(1+r)^{i+1}} + \ldots + \frac{(1-\delta_i) \ldots (1-\delta_{n-1}) W_n G_n}{(1+r)^n} + \frac{Uh_i}{\lambda} G_i + (1-\delta_i) \frac{Uh_{i+1}}{\lambda} G_{i+1} + \ldots + (1-\delta_i) \ldots (1-\delta_{n-1}) \frac{Uh_n}{\lambda} G_n \tag{A3}$$

B. *Utility Maximization—Continuous Time*

Let the utility function be

$$U = \int m_i f(\phi_i H_i, Z_i) di, \tag{A4}$$

where m_i is the weight attached to utility in period i. Equation (A4) defines an additive utility function, but any monotonic transformation of this function could be employed.[41] Let all household production functions be homogeneous of degree 1. Then $C_i = \pi_i I_i$, $C_{1i} = q_i Z_i$,[42] and full wealth can be written as

$$R = \int e^{-ri} (\pi_i I_i + q_i Z_i + W_i TL_i) di. \tag{A5}$$

By definition,

$$I_i = \dot{H}_i + \delta_i H_i, \tag{A6}$$

where H_i is the instantaneous rate of change of capital stock. Substitution of (A6) into (A5) yields

$$R = \int e^{-ri} (\pi_i \delta_i H_i + \pi_i \dot{H}_i + q_i Z_i + W_i TL_i) di. \tag{A7}$$

To maximize the utility function, form the Lagrangian

$$L - \lambda R = \int [m_i f(\phi_i H_i, Z_i) - \lambda e^{-ri} (\pi_i \delta_i H_i + \pi_i \dot{H}_i + q_i Z_i + W_i TL_i)] di, \tag{A8}$$

or

$$L - \lambda R = \int Q(H_i, \dot{H}_i, Z_i, i) di, \tag{A9}$$

where

$$Q = m_i f(\phi_i H_i, Z_i) - \lambda e^{-ri} (\pi_i \delta_i H_i + \pi_i \dot{H}_i + q_i Z_i + W_i T_i L). \tag{A10}$$

Euler's equation for the optimal path of H_i is

[41] Strotz (1955–56) has shown, however, that certain restrictions must be placed on the m_i. In particular, the initial consumption plan will be fulfilled if and only if $m_i = (m_0)^i$.

[42] The variable q_i equals the marginal cost of Z_i.

$$\frac{\partial Q}{\partial H_i} = \frac{d}{di}\frac{\partial Q}{\partial \dot{H}_i}. \tag{A11}$$

In the present context,

$$\begin{aligned}
\frac{\partial Q}{\partial \mathrm{H_i}} &= Uh_iG_i - \lambda e^{-ri}\pi_i\delta_i + \lambda e^{-ri}W_iG_i, \\
\frac{\partial Q}{\partial \mathrm{H_i}} &= -\lambda e^{-ri}\pi_i, \\
\frac{d}{di}\frac{\partial Q}{\partial H_i} &= -\lambda e^{-ri}\dot{\pi}_i + \lambda e^{-ri} r\pi_i.
\end{aligned} \tag{A12}$$

Consequently,

$$G_i\left[W_i + \left(\frac{Uh_i}{\lambda}\right)e^{ri}\right] = \pi_i(r - \tilde{\pi}_i + \delta_i) \tag{A13}$$

which is the continuous time analogue of equation (13).

C. *Wage Effects*

To obtain the wage elasticities of medical care and the time spent producing health, three equations must be partially differentiated with respect to the wage. These equations are the gross investment production function and the two first-order conditions for cost minimization:

$$\begin{aligned}
I(M, TH; E) &= Mg(t; E) = (\tilde{H} + \delta)H, \\
W &= \pi g', \\
P &= \pi(g - tg').
\end{aligned}$$

Since I is linear homogenous in M and TH,

$$\begin{aligned}
\frac{\partial(g - tg')}{\partial M} &= -\frac{t\partial(g - tg')}{\partial TH}, \\
\frac{\partial g'}{\partial TH} &= -\frac{1}{t}\frac{\partial(g - tg')}{\partial TH}, \\
\sigma_p &= \frac{(g - tg')g'}{I\{[\partial(g - tg')]/\partial TH\}}.
\end{aligned}$$

Therefore, the following relationships hold:

$$\begin{aligned}
\frac{\partial(g - tg')}{\partial M} &= -\frac{t(g - tg')g'}{I\sigma_p}, \\
\frac{\partial g'}{\partial TH} &= -\frac{1}{t}\frac{(g - tg')g'}{I\sigma_p}, \\
\frac{\partial(g - tg')}{\partial TH} &= \frac{(g - tg')g'}{Iq_p}.
\end{aligned} \tag{A14}$$

Carrying out the differentiation, one gets

$$g'\frac{dTH}{dW}+(g-tg')\frac{dM}{dW}=-\frac{H(\widetilde{H}+\delta)\varepsilon}{\pi}\left(\frac{d\pi}{dW}-\frac{\pi}{W}\right),$$

$$1=g'\frac{d\pi}{dW}+\pi\left(\frac{\partial g'}{\partial TH}\frac{dTH}{dW}+\frac{\partial g'}{\partial M}\frac{dM}{dW}\right),$$

$$0=(g-tg')\frac{d\pi}{dW}+\pi\left[\frac{\partial(g-tg')}{\partial TH}\frac{dTH}{dW}+\frac{\partial(g-tg')}{\partial M}\frac{dM}{dW}\right]$$

Using the cost-minimization conditions and (A14) and rearranging terms, one has

$$\begin{aligned}
&I\varepsilon\frac{d\pi}{dW}+W\frac{dTH}{dW}+P\frac{dM}{dW}=\frac{I\varepsilon\pi}{W},\\
&I\sigma_p\frac{d\pi}{dW}-\frac{1}{t}P\frac{dTH}{dW}+P\frac{dM}{dW}=I\frac{\pi}{W}\sigma_p,\\
&I\sigma_p\frac{d\pi}{dW}+W\frac{dTH}{dW}-tW\frac{dM}{dW}=0.
\end{aligned}\tag{A15}$$

Since (A15) is a system of three equations in three unknowns—dTH/dW, dM/dW, and $d\pi/dW$—Cramer's rule can be applied to solve for, say, dM/dW:

$$\frac{dM}{dW}=\frac{\begin{vmatrix} I\varepsilon+W & +\dfrac{I\,\varepsilon\pi}{W}\\ I\sigma_p-\left(\dfrac{1}{t}\right)P & +\mathrm{I}\left(\dfrac{\pi}{W}\right)\sigma_p\\ I\sigma_p+W & -0\end{vmatrix}}{\begin{vmatrix} I\varepsilon+W & +P\\ I\sigma_p-\left(\dfrac{1}{t}\right)P & +P\\ I\sigma_p+W & -tW\end{vmatrix}}.$$

The determinant in the denominator reduces to $(I\sigma_p\pi^2I^2)/THM$. The determinant in the numerator is

$$\frac{I\sigma_p}{THM}\left(I\pi\sigma_pTHM+I\pi\varepsilon\frac{P}{W}M^2\right).$$

Therefore,

$$\frac{dM}{dW}=\frac{THM}{I\pi}\left(\sigma_p+\frac{\varepsilon PM}{WTH}\right).$$

In elasticity notation, this becomes

$$e_{M,W} = (1 - K)\varepsilon + K\sigma_p. \tag{A16}$$

Along similar lines, one easily shows that

$$e_{TH,W} = (1 - K)(\varepsilon - \sigma_p). \tag{A17}$$

D. *The Role of Human Capital*

To convert the change in productivity due to a shift in human capital into a change in average or marginal cost, let the percentage changes in the marginal products of medical care and own time for a one-unit change in human capital be given by

$$\frac{\partial(g - tg')}{\partial E}\frac{1}{g - tg'} = \frac{g\hat{g} - tg'\hat{g}'}{g - tg'},$$

$$\frac{\partial g'}{\partial E}\frac{1}{g'} = \hat{g}'.$$

If a shift in human capital were "factor neutral," the percentage changes in these two marginal products would be equal:

$$\hat{g} = \frac{g\hat{g} - tg'\hat{g}'}{g - tg'},$$

or

$$\hat{g}' = \hat{g} = r_H. \tag{A18}$$

The average cost of gross investment in health is defined as

$$\bar{\pi} = (PM + WTH)I^{-1} = (P + Wt)g^{-1}.$$

Given factor neutrality,

$$\frac{d\bar{\pi}}{dE}\frac{1}{\pi} = -g = -r_H. \tag{A19}$$

This coincides with the percentage change in marginal cost, since

$$\pi = P(g - tg')^{-1},$$

and

$$\frac{d\pi}{dE}\frac{1}{\pi} = -\left(\frac{g\hat{g} - tg'\hat{g}'}{g - tg'}\right) = -\hat{g}' = -\hat{g} = -r_H. \tag{A20}$$

Part B of Section IV outlines a derivation of the human capital parameter in the demand curve for medical care but does not give a rigorous proof. Taking the *total* derivative of E in the gross investment function, one computes this parameter thus:

$$\frac{dI}{dE}\frac{1}{I} = M\frac{(g - tg')}{I}M + \frac{THg'}{I}TH + r_H.$$

Since $\hat{M} = \widehat{TH}$ and $\hat{H} = \hat{I}$, the last equation can be rewritten as

$$\hat{H} = \hat{M} + r_H.$$

Solving for $\hat{M}$ and noting that $\hat{H} = r_H \varepsilon$, one gets

$$\hat{M} = r_H(\varepsilon - 1). \tag{A21}$$

Appendix B

Glossary of Mathematical Terms

n	Total length of life
i	Age
H_0	Inherited stock of health
H_i	Stock of health in period i
$H_{\min}$	Death stock
ϕ_i	Service flow per unit stock or number of healthy days per unit stock
h_i	Total number of healthy days in period i
Z_i	Consumption of an aggregate commodity in period i
I_i	Gross investment in health
δ_i	Rate of depreciation
M_i	Medical care
TH_i	Time input in gross investment function
X_i	Goods input in the production of Z_i
T_i	Time input in the production of Z_i
E_i	Stock of human capital
$g - t_i g'$	Marginal product of medical care in the gross investment production function
g'	Marginal product of time
P_i	Price of medical care
V_i	Price of X_i
W_i	Wage rate
A_0	Initial assets
r	Rate of interest
TW_i	Hours of work
TL_i	Sick time
Ω	Constant length of the period
R	Full wealth
G_i	Marginal product of health capital
Uh_i	Marginal utility of healthy days
λ	Marginal utility of wealth
π_i	Marginal cost of gross investment in health
$\tilde{\pi}_i$	Percentage rate of change of marginal cost
γ_i	Monetary rate of return on an investment in health or marginal efficiency of health capital
a_i	Psychic rate of return on an investment in health
$\sim$	A tilde over a variable denotes a percentage time derivative
s_i	Share of depreciation in the cost of health capital
ε	Elasticity of the MEC schedule
K	Fraction of the total cost of gross investment accounted for by time
σ_p	Elasticity of substitution betwen medical care and own time in the production of gross investment

$e_{H,W}$ .... Elasticity of H with regard to W
$e_{M,W}$ .... Elasticity of M with regard to W
$\hat{\vee}$ .... A circumflex over a variable denotes a percentage change per unit change in E
r_H .... Percentage change in gross investment for a one unit change in E
C_i .... Total cost of gross investment in health in period i
C_{1i} .... Total cost of Z_i
m_i .... Weight attached to total utility in period i
q_i .... Marginal cost of Z_i

References

Adelman, Irma. "An Econometric Analysis of Population Growth." *A.E.R.* 53 (June 1963): 314–49.

Arrow, Kenneth J. "Optimal Capital Policy with Irreversible Investment." In *Value, Capital and Growth: Papers in Honour of Sir John Hicks*, edited by J. N. Wolfe. Edinburgh: Edinburgh Univ. Press, 1968.

Auster, Richard D.; Leveson, Irving; and Sarachek, Deborah. "The Production of Health: An Exploratory Study." *J. Human Resources* 4 (Fall 1969): 411–36.

Becker, Gary S. *Human Capital.* New York: Columbia Univ. Press (for Nat. Bur. Econ. Res.), 1964.

———. "A Theory of the Allocation of Time." *Econ. J.* 75 (September 1965): 493–517.

———. *Human Capital and the Personal Distribution of Income: An Analytical Approach.* W. S. Woytinsky Lecture no. 1. Ann Arbor: Univ. Michigan, 1967.

Becker, Gary S., and Michael, Robert T. "On the Theory of Consumer Demand." Unpublished paper, 1970.

Ben-Porath, Yoram. "The Production of Human Capital and the Life Cycle of Earnings." *J.P.E.* 75 (August 1967): 353–67.

Eisner, Robert, and Strotz, Robert H. "Determinants of Business Investment." In *Impacts of Monetary Policy.* Englewood Cliffs, N.J.: Prentice-Hall (for Commission on Money and Credit), 1963.

Fuchs, Victor R. "Some Economic Aspects of Mortality in the United States." Mimeographed. New York: Nat. Bur. Econ. Res., 1965.

———. "The Contribution of Health Services to the American Economy." *Milbank Memorial Fund Q.* 44 (October 1966): 65–102.

Ghez, Gilbert R. "A Theory of Life Cycle Consumption." Ph.D. dissertation, Columbia Univ., 1970.

Gould, John P. "Adjustment Costs in the Theory of Investment in the Firm." *Rev. Econ. Studies* 25 (January 1968): 47–55.

Grossman, Michael. "The Demand for Health: A Theoretical and Empirical Investigation." Ph.D. dissertation, Columbia Univ., 1970.

Lancaster, Kelvin J. "A New Approach to Consumer Theory." *J.P.E.* 74 (April 1966): 132–57.

Larmore, Mary Lou. "An Inquiry into an Econometric Production Function for Health in the United States." Ph.D. dissertation, Northwestern Univ., 1967.

Michael, Robert T. "The Effect of Education on Efficiency in Consumption." Ph.D. dissertation, Columbia Univ., 1969; to be published by Nat. Bur. Econ. Res.

Mushkin, Selma J. "Health as an Investment." *J.P.E.* 70, no. 2, suppl. (October 1962): 129–57.
Muth, Richard. "Household Production and Consumer Demand Functions." *Econometrica* 34 (July 1966): 699–708.
Nerlove, Marc, and Arrow, Kenneth J. "Optimal Advertising Policy under Dynamic Conditions." *Economica* 29 (May 1962): 129–42.
Newhouse, Joseph P. "Toward a Rational Allocation of Resources in Medical Care." Ph.D. dissertation, Harvard Univ., 1968.
Phelps, Charles E. "The Demand for Health Insurance: Theory and Empirical Results." Ph.D. dissertation, Univ. Chicago (in preparation).
Strotz, Robert H. "Myopia and Inconsistency in Dynamic Utility Maximization." *Rev. Econ. Studies* 23 (1955–56): 165–180.