Volume Title: The Demand for Health: A Theoretical and Empirical Investigation

Volume Author/Editor: Michael Grossman

Volume Publisher: NBER

Volume ISBN: 0-87014-248-8

Volume URL: http://www.nber.org/books/gros72-1

Publication Date: 1972

Chapter Title: The Shadow Price of Health

Chapter Author: Michael Grossman

Chapter URL: http://www.nber.org/chapters/c3486

Chapter pages in book: (p. 11 - 30)

II

THE SHADOW PRICE OF HEALTH

In the previous chapter, I showed how a consumer selects the optimal quantity of health in any period of his life. In this chapter, I explore the effects of changes in the supply and demand prices of health in the context of the pure investment model. In Section 1, I comment on the demand curve for health capital in the investment model. In Section 2, I relate variations in depreciation rates with age to life cycle patterns of health and gross investment. I also examine the impact of changes in depreciation rates among individuals of the same age and briefly incorporate uncertainty into the model. In the third section, I consider the effects of shifts in market efficiency, measured by the wage rate, and nonmarket efficiency, measured by human capital, on supply and demand prices.

1. THE INVESTMENT DEMAND CURVE

If the marginal utility of healthy days or the marginal disutility of sick days were equal to zero, health would be solely an investment commodity. The optimal amount of H_i could then be found by equating the marginal monetary rate of return on an investment in health to the cost of health capital:

$$W_i G_i/\pi_{i-1} = \gamma_i = r - \tilde{\pi}_{i-1} + \delta_i. \tag{2-1}$$

Setting $Uh_i = 0$ in equation (1-13′), one derives equation (2-1). It can also be derived by excluding health from the utility function and redefining the full wealth constraint as

$$R' = A_0 + \sum \frac{W_i h_i - \pi_i I_i}{(1 + r)^i} \tag{2-2}$$

Maximization of R' with respect to gross investment in periods $i - 1$ and i yields condition (2-1).[1]

[1] For a proof, see Appendix B, Section 1. The continuous time version of (2-1) is

$$W_i G_i/\pi_i = \gamma_i = r - \tilde{\pi}_i + \delta_i,$$

where $\tilde{\pi}_i$ is the instantaneous percentage rate of change in marginal cost. This equation, too, is derived in Appendix B, Section 1.

Figure 1 illustrates the determination 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. The supply curve *S* shows the relationship between the stock of health and the cost of capital. Since the real-own rate of interest and the rate of depreciation are independent of the stock, the supply curve is infinitely elastic. Provided the MEC schedule slopes downward, the equilibrium stock is given by H_i^*, where the supply and demand curves intersect.

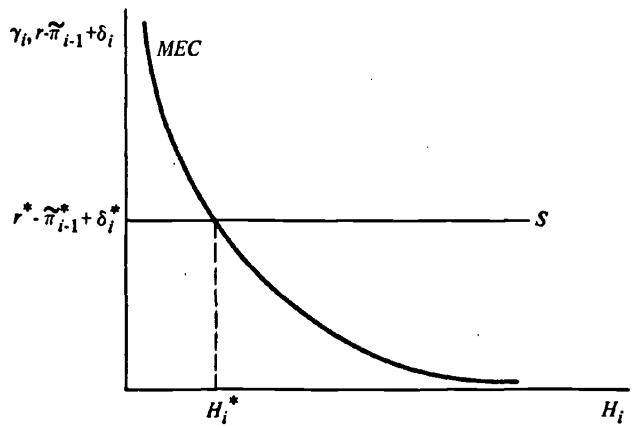

Figure 1

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 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

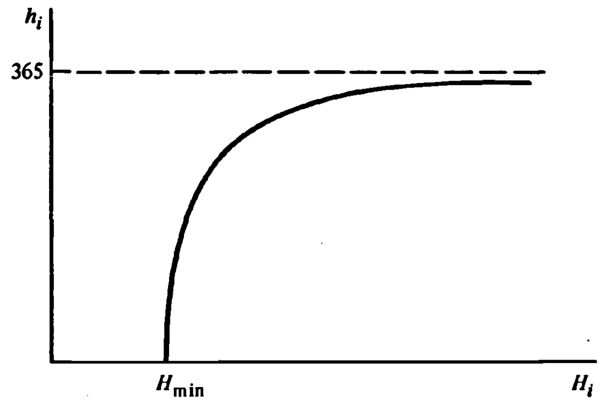

Figure 2

approaches its upper asymptote of 365 days as the stock becomes large.[2]

2. VARIATIONS IN DEPRECIATION RATES

Life Cycle Patterns

Equation (2-1) 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 3. 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.[3]

[2] Certain production functions might exhibit upper asymptotes but increasing or constant marginal productivity in some regions. In general, if discontinuities in the MEC schedule are ruled out, the sufficient condition for maximizing R' with respect to H_i requires diminishing marginal productivity in the vicinity of the equilibrium stock. For a complete discussion of this point, see Appendix B, Section 1.

[3] For an analysis of life cycle phenomena that allows wage rates and human capital to vary with age, see equation (2-21).

Since marginal cost does not depend on age, $\tilde{\pi}_{i-1} = 0$, and equation (2-1) reduces to

$$\gamma_i = r + \delta_i. \tag{2-3}$$

It is apparent from equation (2-3) 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 discrepancy between the inherited stock and the stock desired in period 1. This discrepancy 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 $H_1 > H_{\min}$, H_i would always exceed the death stock.[4]

To permit the demand for health to vary with age, assume that 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 early stages of the life cycle. Or 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 increases the cost of health capital, it would cause the demand for health to fall over the life cycle. Graphically, an upward shift 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 is the elasticity of the MEC schedule, the greater the decrease in the optimal stock with age. Put differently, the slower is the increase in the marginal product of health capital as H falls, the greater the decrease in the optimal stock.

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

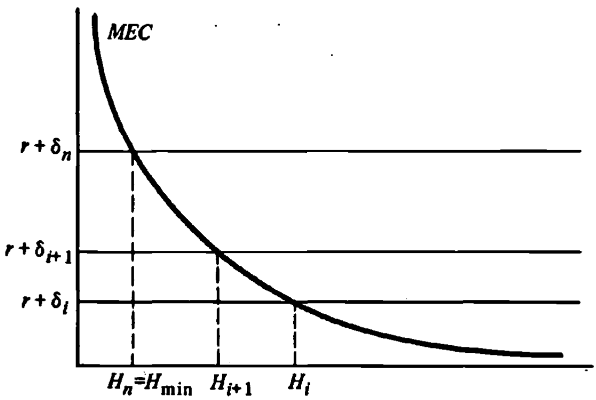

Figure 3

Differentiation of equation (2-3) with respect to age quantifies the percentage rate of decrease in the stock of health over the life cycle:

$$\tilde{H}_i = -s_i \varepsilon_i \tilde{\delta}_i. \tag{2-4}$$

A tilde denotes a percentage time derivative

$$\left(\tilde{H}_i = \frac{dH_i}{di}\frac{1}{H_i}, \text{etc.}\right),$$

where:

$$s_i = \frac{\delta_i}{r + \delta_i} = \text{the share of depreciation in the cost of health capital}$$

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

$= \text{elasticity of MEC schedule.}$[5]

Equation (2-4) 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

[5] From equation (2-3), $\ln (r + \delta_i) = \ln W + \ln G_i - \ln \pi$. Therefore,

$$\frac{\delta_i \tilde{\delta}_i}{r + \delta_i} = \frac{\partial \ln G_i}{\partial \ln H_i}\tilde{H}_i, \quad \text{or} \quad s_i \tilde{\delta}_i = -\frac{\tilde{H}_i}{\varepsilon_i}.$$

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[6]

$$\frac{\mathrm{d}\tilde{H}_i}{\mathrm{d}i} = \tilde{H}_{ii} = -s_i(1 - s_i)\varepsilon\tilde{\delta}^2 < 0. \tag{2-5}$$

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

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, $W\Omega$. 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.[7] Because individuals could not produce commodities, total utility would be driven to zero at death.[8]

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.

[6] Differentiation of (2-4) with respect to age yields

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

[7] This assumes that Z_i cannot be produced with X_i alone; which would be true if, say, the production function were Cobb-Douglas.

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

To predict the effect of an increase in δ_i with age on gross investment, note that net investment can be approximated by $H_i\tilde{H}_i$.[9] Since gross investment equals net investment plus depreciation,

$$\ln I_i = \ln H_i + \ln(\tilde{H}_i + \delta_i). \tag{2-6}$$

Differentiation of equation (2-6) 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 (2-4) and (2-5), the expression for $\tilde{I}_i$ would simplify to

$$\tilde{I}_i = \frac{\tilde{\delta}(1 - s_i\varepsilon_i)(\delta_i - s_i\varepsilon\tilde{\delta}) + s_i^2\varepsilon\tilde{\delta}^2}{\delta_i - s_i\varepsilon\tilde{\delta}}. \tag{2-7}$$

Since health capital cannot be sold, gross investment cannot be negative. Therefore, $\delta_i \geq -\tilde{H}_i$.[10] 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 (2-7) 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 one, 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 one. A general equation for the healthy days production function illustrated by Figure 2 is

$$h_i = 365 - BH_i^{-C}, \tag{2-8}$$

[9] That is, $H_{i+1} - H_i \simeq H_i(dH_i/di)(1/H_i) = H_i\tilde{H}_i$. 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.

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

where B and C are positive constants. The corresponding MEC schedule is[11]

$$\ln \gamma_i = \ln BC - (C + 1)\ln H_i + \ln W - \ln \pi. \tag{2-9}$$

The elasticity of this schedule is given by

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

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 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.[12] 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 one, 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 will reject a prospect—the prospect

[11] If (2-8) were the production function, the marginal product of health capital would be

$$G_i = BCH_i^{-C-1}, \qquad \text{or} \qquad \ln G_i = \ln BC - (C + 1)\ln H_i.$$

Since $\ln \gamma_i = \ln G_i + \ln W - \ln \pi$, one uses the equation for $\ln G_i$ to obtain (2-9).

[12] 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.

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.

Cross-Sectional Variations

The framework used to analyze life cycle variations in depreciation rates can easily be applied to examine the impact of variations in these rates among individuals of the same age. Assume, for example, a uniform percentage shift in δ_i across persons so that

$$\frac{\mathrm{d}\ln\delta_{i-1}}{\mathrm{d}\ln\delta_i} = 1, \text{ all } i.$$

It is clear that such a shift would have the same kinds of effects as an increase in δ_i with age. That is, persons of a given age who face relatively high depreciation rates would simultaneously reduce their demand for health but would increase their demand for gross investment if $\varepsilon < 1$. Differentiating equations (2-3) and (2-6) with respect to $\ln \delta_i$, one obtains[13]

$$\frac{\mathrm{d}\ln H_i}{\mathrm{d}\ln\delta_i} = -s_i\varepsilon \tag{2-10}$$

$$\frac{\mathrm{d}\ln I_i}{\mathrm{d}\ln\delta_i} = \frac{(1-s_i\varepsilon)(\delta_i - s_i\varepsilon\tilde{\delta}) + s_i^2\varepsilon\tilde{\delta}}{\delta_i - s_i\varepsilon\tilde{\delta}}. \tag{2-11}$$

According to (2-10) and (2-11), if ε were less than unity, H_i or h_i would be negatively correlated with TH_i and M_i (and TL_i would be positively correlated with these inputs) across individuals of the same age.

Uncertainty

The development of the model to this point has ruled out uncertainty. Consumers fully anticipate intertemporal and cross-sectional variations in depreciation rates and, therefore, know their age of death with certainty. In the real world, however, length of life is surely not known with perfect foresight. In order to explain variations in death time expectations, uncertainty must be introduced into the model. The easiest way to accomplish this is to postulate that a given consumer faces a probability

[13] Derivations of (2-10) and (2-11) are contained in Appendix B, Section 2.

distribution of depreciation rates in every period. For simplicity, let there be two depreciation rates, δ_{ia} and δ_{ib}, where $\delta_{ia} > \delta_{ib}$.[14] These depreciation rates correspond to two mutually exclusive outcomes, *a* and *b*. Since depreciation rates are not known with certainty, length of life can no longer be determined in a precise fashion. In particular, it would depend on the pattern of depreciation rates that actually occurs and would tend to be longer with patterns in which outcome *b* occurred more frequently than outcome *a*. But because depreciation rates rise with age, the stock of health would still tend to fall over the life cycle.

Besides creating dispersion in death time expectations, the existence of uncertainty has a number of additional implications. These follow from the state-preference approach to the problem of choice under uncertainty.[15] Since none of the major conclusions reached in this section and the next one tend to be altered, I will simply state here the main results of the analysis.[16] At any given age, health capital would tend to be lower and sick time and gross investment would tend to be higher in relatively undesirable "states of the world," i.e., outcomes with higher than average depreciation rates. The monetary value of the excess sick time and gross investment measures the "loss" associated with unfavorable states. Since this loss could be reduced by increasing the stock of health, consumers might have an incentive to hold excess stocks in relatively desirable states. In these states, the rate of return to an increase in H_i might be less than the cost of capital. Put differently, part of the demand for health capital would reflect a demand for self-insurance against losses in unfavorable states.

Consumers could also finance the monetary value of their losses by purchasing health insurance in the market. Conclusions reached by Ehrlich and Becker suggest that market health insurance and the stock of health should be substitutes; that is, an increase in market insurance

[14] In general, $\delta_{i+1a} > \delta_{ia}$ and $\delta_{i+1b} > \delta_{ib}$ since depreciation rates rise with age.

[15] This approach was developed by Jack Hirshleifer in "Investment Decisions under Uncertainty: Choice-Theoretic Approaches," *Quarterly Journal of Economics*, 79, No. 4 (November 1965). For an application to some general insurance problems, see Isaac Ehrlich and Gary S. Becker, "Market Insurance, Self-Insurance and Self-Protection," *Journal of Political Economy*, 80, No. 4 (July/August 1972).

[16] For derivations of these results, see Michael Grossman, "The Demand for Health: A Theoretical and Empirical Investigation," unpublished Ph.D. dissertation, Columbia University, 1970, Appendix B, pp. 131–135. Note that the assumption of perfect certainty is reintroduced at the end of this subsection. The empirical implementation of the model makes some attempt, however, to deal with uncertainty. See the discussion of this phenomenon in Chapter IV, Section 2, and Chapter V, Section 1.

would increase the optimal loss and reduce the stock of health.[17] Note finally that it is doubtful whether consumers can ever fully insure against all losses via the market. This follows because full insurance would not only finance the monetary value of excess gross investment and working time lost in states with high depreciation rates, but would also finance the monetary value of time lost from nonmarket activities. Hence, the demand for health capital may be substantial even when market insurance is available.

3. 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 cost 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.

Before beginning the analysis, two preliminary comments are in order. First, most of 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

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

Second, the discussion (through equation 2-20) 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.[18] Note that from equation (2-6)

$$\frac{\mathrm{d}\ln I}{\mathrm{d}X} = \frac{\mathrm{d}\ln H}{\mathrm{d}X}, \tag{2-12}$$

[17] See Ehrlich and Becker, "Market Insurance."

[18] 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 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, see footnote 30.

since the rate of depreciation and the percentage rate of net investment do not depend on X.[19] Equation (2-12) 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.

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.

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

[19] Since the main part of the analysis in this section deals with variations in X among individuals of the same age, time subscripts are omitted until equation (2-21). Note also that (2-12), like the expression for $\tilde{I}_t$, ignores the one period lag between an increase in gross investment and an increase in the stock of health.

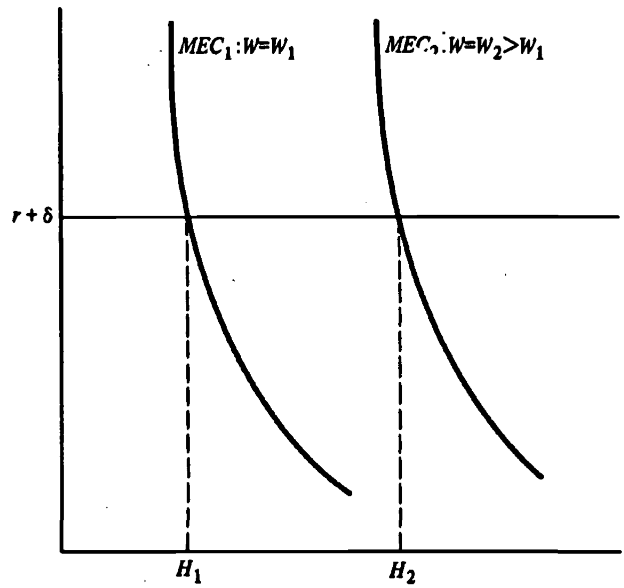

Figure 4

health capital, increases the optimal stock from H_1 to H_2. A formula for the wage elasticity of health capital is[20]

$$e_{H,W} = (1 - K)\varepsilon. \tag{2-13}$$

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 investment 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

[20] Differentiation of the natural logarithm of (2-3) with respect to ln W yields

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

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

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

$$e_{M,W} = K\sigma_p + (1 - K)\varepsilon, \tag{2-14}$$

where σ_p is the elasticity of substitution between M and TH in the production of gross investment.[21] The greater the value of σ_p the greater the difference between the wage elasticities of M and H.

Note that an increase in either the price of 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.

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 the 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.[22] While the analysis that follows could pertain to any

[21] For a proof, see Appendix B, Section 2. The corresponding equation for the wage elasticity of TH is

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

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

[22] Recall from Chapter I that at an operational level the stock of knowledge or human capital does not include health capital.

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

The specific hypothesis to be tested 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.[23]

To determine the effects of education on production, marginal cost, and the demand for health and medical care, assume the gross investment production function is homogeneous of degree one in its two direct inputs—medical care and own time. Then 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.[24] If a circumflex over a variable denotes a percentage change per unit change in E, the last equation can be rewritten as

$$r_H = \frac{\partial I}{\partial E}\frac{1}{I} = \left[\frac{M(g - tg')}{I}\right]\left[\frac{g\hat{g} - tg'\hat{g}'}{g - tg'}\right] + \left[\frac{THg'}{I}\right][\hat{g}']. \quad (2\text{-}15)$$

Equation (2-15) indicates that the percentage change in gross investment supplied to a consumer by a one unit change in E is a weighted average

[23] The model developed here is somewhat similar to the one used by Robert T. Michael in *The Effect of Education on Efficiency in Consumption*, New York, NBER, Occasional Paper 116, 1972. Michael's model examines the effects of education on the demand for consumption commodities and not investment commodities. His analysis, therefore, is more relevant to the consumption model of health presented in Chapter III.

[24] If I is homogeneous of degree one 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.

of the percentage changes in the marginal products of M and TH.[25] If E increases productivity, then $r_H > 0$. Provided E raises both marginal products by the same percentage, equation (2-15) would simplify to

$$r_H = \hat{g} = \hat{g}'. \tag{2-16}$$

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

$$\hat{\pi} = -r_H = -\hat{g} = -\hat{g}', \tag{2-17}$$

where $\hat{\pi}$ is the percentage change in average or marginal cost.[26] 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. 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 charge in the supply curve, and the more educated would demand a larger optimal stock (compare H_1 and H_2 in the figure). Note that E shifts the MEC schedule not because it is a determinant of consumers' "tastes" for health but because it is a determinant of nonmarket productivity.

[25] 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.

[26] For a proof, see Appendix B, Section 2, where the human capital formulas are developed in more detail.

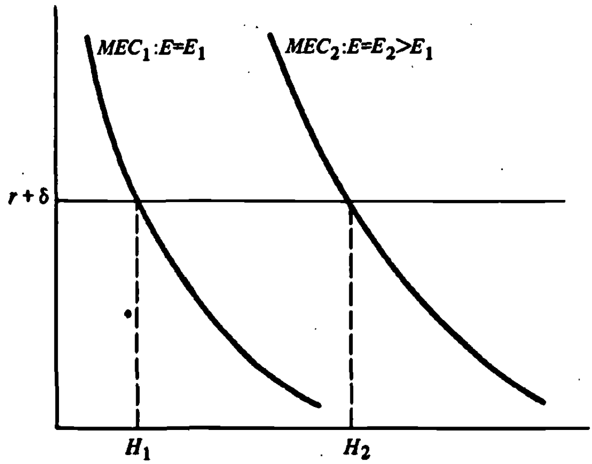

Figure 5

The percentage increase in the amount of health demanded for a one unit increase in E is given by[27]

$$\hat{H} = r_H \varepsilon. \tag{2-18}$$

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 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 differential between $\hat{H}$ and r_H.[28]

$$\hat{M} = \widehat{TH} = r_H(\varepsilon - 1). \tag{2-19}$$

[27] 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

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

[28] The terms $\hat{M}$ and $\widehat{TH}$ 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.

Equation (2-19) 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 one, $\hat{H}$ would be negative and $\hat{M}$ would be positive Since education improves market productivity, this study tests 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.

Joint Effects

This section has examined the *partial* effect of an increase in the wage rate or an increase in education on the demand for health. But, surely, these two variables are positively correlated, which raises two questions. First, what is the combined effect of an expansion in education, one that takes account of the impact of education on wage rates? Second, can variations in nonmarket efficiency be separated empirically from variations in market efficiency?

A formula for $\hat{H}$ that combines market and nonmarket efficiency changes is

$$\hat{H} = r_H\varepsilon + (1 - K)\varepsilon\hat{W}, \tag{2-20}$$

where $\hat{W}$ is the percentage change in the wage rate for a one unit change in E.[29] Equation (2-20) reveals the dual motive of the more educated for demanding more health capital. With nonmarket productivity constant, an increase in E causes the demand price of health capital to rise at a faster rate than the supply price, and with market productivity constant, E is negatively correlated with the supply price. Integration of (2-20) with respect to E yields

$$\ln H = r_H\varepsilon E + (1 - K)\varepsilon \ln W,$$

[29] If W is not held constant as E increases, then

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

or

$$0 = W(1-K) - \frac{\hat{H}}{\varepsilon} + r_H.$$

provided r_H, ε, and K are constant. Hence, by regressing $\ln H$ on E and $\ln W$, *partial* market and nonmarket efficiency parameters can be estimated. Of course, this procedure may break down if the correlation between E and $\ln W$ is extremely high. So it is the size of this correlation that ultimately determines whether pure changes in nonmarket productivity can be isolated at the empirical level.

Along similar lines, the analysis of life cycle variations in the demand for health can be modified to take account of the life cycle pattern of the stock of human capital. Letting $\dot{E}_i = \mathrm{d}E_i/\mathrm{d}i$, one would revise the formula for $\tilde{H}_i$ as follows:[30]

$$\tilde{H}_i = -s_i\varepsilon\tilde{\delta} + r_H\varepsilon\dot{E}_i + (1 - K)\varepsilon\hat{W}\dot{E}_i. \tag{2-21}$$

If equation (2-21) were applied to individuals who had completed their formal schooling, E would tend to increase at early ages due to on-the-job training. Eventually it would, however, decline as depreciation on the stock of human capital began to outweigh gross investment. Even if $\tilde{\delta}$ were always positive, the positive values of $\dot{E}_i$ at early stages of the life cycle might make $\tilde{H}_i$ positive during these stages. But ultimately the

[30] Replace equation (2-1) with its continuous time analogue:

$$\gamma_i = r - \tilde{\pi}_i + \delta_i.$$

Then if E_i is not fixed as i increases,

$$s_i\tilde{\delta} - \frac{\tilde{\pi}_{ii}\delta_i}{s_i} = \frac{\mathrm{d}\ln W_i}{\mathrm{d}E_i}\frac{\mathrm{d}E_i}{\mathrm{d}i} + \frac{\partial \ln G_i}{\partial \ln H_i}\frac{\mathrm{d}\ln H_i}{\mathrm{d}i} - \frac{\partial \ln \pi_i}{\partial E_i}\frac{\mathrm{d}E_i}{\mathrm{d}i} - \frac{\partial \ln \pi_i}{\partial \ln W_i}\frac{\mathrm{d}\ln W_i}{\mathrm{d}E_i}\frac{\mathrm{d}E_i}{\mathrm{d}i}$$

$$s_i\tilde{\delta} - \frac{\tilde{\pi}_{ii}\delta_i}{s_i} = \hat{W}_i\dot{E}_i - \frac{\tilde{H}}{\varepsilon} + r_H\dot{E}_i + (1 - K)\hat{W}_i\dot{E}_i.$$

Note that since the wage rate varies over the life cycle, $\tilde{\pi}_i$ would not, in general, equal zero. Instead, it would be given by

$$\tilde{\pi}_i = (K_i\hat{W}_i - r_H)\dot{E}_i.$$

The equation in the text assumes that $\tilde{\pi}_{ii} = \mathrm{d}\tilde{\pi}_i/\mathrm{d}i \simeq 0$. If r_H and $\hat{W}_i$ were independent of $\dot{E}_i$,

$$\tilde{\pi}_{ii} = (K_i\hat{W} - r_H)\frac{\mathrm{d}\dot{E}_i}{\mathrm{d}i} + (\dot{E}_i\hat{W})^2\frac{\mathrm{d}K_i}{\mathrm{d}\ln W_i}.$$

Provided the elasticity of substitution on production equaled unity, the second term on the right-hand side of the equation would equal zero. The first term could be ignored if $\mathrm{d}\dot{E}_i/\mathrm{d}i$ were relatively small or if the difference between $K_i\hat{W}$ and r_H were small.

Note also that even if $\hat{W}_i$ and $\dot{E}_i$ were independent of $\ln W_i$, a uniform percentage shift in wage rates across persons of the same age might alter $\tilde{H}_i$. This follows because it would change K_i and hence the real rate of interest unless the elasticity of substitution in production equaled one.

effect of net disinvestment in human capital would strengthen the effect of a rising depreciation rate.

4. GLOSSARY

A tilde over a variable denotes a percentage time derivative; a circumflex over a variable denotes a percentage change per unit change in E; and a dot over a variable denotes an absolute rate of change over the life cycle.

s_t	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
$e_{H,W}$	Elasticity of H with respect to W
$e_{M,W}$	Elasticity of M with respect to W
σ_p	Elasticity of substitution between medical care and own time in the production of gross investment
r_H	Percentage change in gross investment for a one unit change in E