Oxford University Press 2007 *Oxford Economic Papers* 59 (2007), 379–410 **379**
doi:10.1093/oep/gpm020

Health as human capital: synthesis and extensions[1]

By Gary S. Becker

Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637, USA, and Hoover Institution; e-mail: gbecker@uchicago.edu

1. Introduction

The first major collection of articles on human capital (Schultz, 1962) contained discussions of education, on the job training, migration, and health. While literally thousands of articles and books followed on human capital dimensions of education and training, there have been many fewer discussions of health as human capital. This is partly because the contribution on health in this volume is not particularly insightful, but it is also because the concept of health as human capital relies on somewhat different concepts than does education or training.

A major step forward occurred with Grossman's work (Grossman, 1972) that modeled optimal investment in increasing longevity. This article stimulated a large literature, but nevertheless, articles on health as human capital have been only a small fraction of those on education and training. In fact, most of the economics literature on health discusses ways to improve the delivery of health care services, such as HMO's or health savings accounts. Health care delivery is an important topic that interacts with considerations of health as human capital, but it is mainly a different topic.

The emerging field of health as human capital builds on three interrelated developments to create a dynamic and evolving field. These developments are (i) The analysis of optimal investments in health by individuals, drug companies, and to a lesser extent by governments that follows on Grossman's analysis, and also on the discussions in the insurance literature of self protection (see Ehrlich and Becker, 1972; and Ehrlich, 2000), and the literature on investments by pharmaceuticals. (ii) The value of life literature that analyzes how much people are willing to pay for improvements in their probabilities of surviving different ages (see, especially, Usher 1973; Rosen, 1988; and Murphy and Topel, 2006). (iii) The importance of complementarities in linking health to education and other types of human capital investments, and in linking investments in health to discount rates, to progress

[1]A revised version of the Hicks Lecture delivered at Oxford University on June 13, 2006.

in fighting different diseases (see the discussion of competing risks in Dow *et al.*, 1999; and Murphy and Topel, 2006) and to other sources of overall changes in survivorship rates.

This paper will present the resulting emerging theory of health as human capital. It builds on and integrates various contributions, but it adds to the literature by incorporating a few relevant ideas that have been ignored. The paper also refers to some of the evidence that demonstrates the empirical importance of health as human capital. This evidence covers general advances in health for the past several decades, and advances in treating a few specific diseases.

Table 1 presents a strong motivation for an interest in the economic value of improved life expectancy. It shows that life expectancy at birth hovered around 40 years in the 19th century, even in countries like the United Kingdom and the United States that were among the very richest. Life expectancy improved only by a few years during that century, but really took off during the 20th century. Life expectancy at birth hit the mid-sixties by 1950, and was close to 80 years by the beginning of the twenty first century. I believe the decline in mortality at all ages was among the most significant economic and social developments of the twentieth century.

2. Theory

2.1 The statistical value of life

The fundamental equation for determining the so-called 'statistical value of life' for any individual is the following:

$$V(W_o + \Delta W, S_o, r, A, \ldots) = V(W_o, S_o + \Delta S, r, A, \ldots) \tag{1}$$

Table 1 Average life expectancy in OECD countries

Average life expectancy	OECD Average
Early 19th century	38.8
Middle 19th century	41.0
Late 19th century	45.1
1900	48.5
1910	52.9
1920	53.7
1930	60.2
1940	60.8
1950	66.9
1960	70.8
1970	71.9
1980	74.0
1990	75.8
2000	78.2

Source: Human Mortality Database.

The right hand side gives the utility a person of age A who has a wealth of W_o is faced with an interest rate of r, and a vector of probabilities of surviving to different ages measures by an initial set of survivorship rates, S_o, plus an improvement of ΔS (see Becker *et al.*, 2005, that extends Usher, 1973). The left hand side gives the 'compensating variation' in wealth; by that is meant the increment in wealth from W_o that makes the person just as well off with S_o as he would be with the initial wealth and the improved survivorship.

Since V is rising in both W and S, $v = \Delta W / \Delta S$ is positive, and measures the 'statistical value of life' for this person who has wealth W_o, is faced with an improvement in survivorship of ΔS from a level of S_o, is age A, and faces interest rate r. Given the usual properties of utility functions, it follows that v is rising in initial wealth W_o, falls as r increases, and generally falls with age, although not necessarily monotonically. It is crucial to recognize that the statistical value of life considered by economists is not a constant even for the same person, but varies with his wealth, age, level of survivorship, magnitude of the changes in survivorship, and perhaps other variables.

For example, suppose $S_o = 0$, so that the initial survivorship condition is to die with certainty, while $\Delta S > 0$, so that the change makes it possible to live with perhaps a significant probability. The change may be due to a life-saving drug, surgery, etc. How large would ΔW be then? At one extreme, ΔW may equal W_o for any $\Delta S > 0$, so that a person would be willing to give all his wealth to avoid dying. Or ΔW may be less than W_o because utility is received from bequests. The large amounts that might be spent to delay death would help explain why a large fraction of health expenditures are made at older ages when people spend on surgeries or other medical interventions to avoid dying (see Becker *et al.*, 2006, in progress).

2.2 Expected utility

The common approach is take a more specialized formulation than eq. (1), and to specify the utility function to a discounted value of expected utilities at different ages, as in

$$U = \sum B^i S_i u_i(x_i, l_i), \tag{2}$$

where u_i is the utility at age i that depends on goods, x, and leisure, l, at that age, B is the discount rate, and S_i is the probability of surviving from the initial age to age i. The only uncertainty in this formulation is the uncertainty about length of life.

The unconditional probability of surviving to age i is the product of the conditional probabilities of surviving various ages:

$$S_i = s_0 s_{1\ldots} s_{i-1}, \tag{3}$$

where s_j is the probability of surviving age j, given that one survived to age $j-1$. If all the conditional probabilities were the same and equal to s, then $S_i = s^i$, and eq. (2) becomes

$$U = \sum B^i s^i u_i(x_i, l_i) = U = \sum (Bs)^i u_i(x_i, l_i) \tag{4}$$

Utility in each period is now multiplied by the product of the time discount rate and the conditional probability of surviving each age.

For expositional simplicity I mainly use a two period formulation, where $S_0 = 1$, and $S_1(h)$, where h is the number of units of health purchased to raise probability of surviving the second period.

$$U = u_0(x_0, l_0) + BS_1(h)u_1(x_1, l_1), \quad \text{with} \quad S_{1'} > 0, \ \text{and} \ S_{1''} \leqslant 0 \tag{5}$$

Of course, S_1 also depends on epidemics, random shocks to health, public health programs, and so forth.

The function u_1 gives the utility at age 1 if alive at that age. Hence this approach is essentially normalizing the utility from death to zero. This is fine if the utility of death is fixed and independent of events, age, probability of dying, and many other variables. However, there is an old literature on the 'fear of death', which implies that this fear might depend on age, the likelihood of dying, and other variables. If the fear of death changed with advancing age, with the probability of dying, or other variables, the normalization of the utility from death at 0 would not be tenable, and this fear could help explain the large expenditures on health at older ages (Becker *et al.*, 2006). I will neglect such considerations in this paper, although they are important in understanding how the value of life can vary in interesting ways with age (implications of the fear of terrorist attacks are considered by Becker and Rubinstein, 2006).

To simplify the budget constraint, I assume a full and fair annuity market that protects a person against the risk of running out of resources because he lives longer than expected, and against having unspent resources if he dies prematurely. If the capital market were also perfect, so that individuals can borrow and lend expected wealth at a fixed interest rate, the budget constraint with full insurance would be

$$x_0 + S_1 x_1/(1+r) + g(h) = w_0(1-l_0) + S_1 w_1(1-l_1)/(1+r) = W, \tag{6}$$

where w_i is the wage rate in period i, g gives expenditures on health, with the g function assumed to be convex: $g' > 0$, and $g' \geqslant 0$. I am assuming that all health spending occurs in the initial period, and that the price of x is 1.

If the utility function in eq. (5) is maximized with respect to the x's and l's subject to the constraint in eq. (6), one gets the usual first order conditions:

$$u_{0x} = B(1+r)u_{1x}, \ u_{0l}/u_{0x} = w_0, \quad \text{and} \ u_{1l}/u_{1x} = w_1 \tag{7}$$

The equilibrium marginal rate of substitution between goods today and goods in the future depends only on time preference and interest rates, and the equilibrium marginal rate of substitution between goods and time during any period depends only on the wage rate and utility function during that period.

In this case, full and fair insurance eliminates any effects of survival probabilities on savings, and the rates of substitution between present consumption and leisure time, and future consumption and leisure time. However, it is often failed to realize that this independence result with full insurance also requires separability over time in the utility function. With recursive utility due say to habits or addictions—that is, with consumption stocks—then even with fair market insurance, the probability of surviving in the future does affect the expected utility from consuming such goods in the present.

In particular, the greater the probability of surviving in the future, the smaller the incentive to consume harmfully addictive goods and the greater the incentive to consume beneficially addictive goods. So individuals with lower life expectancies should be more likely to be addicted to drugs, smoking, and other harmful goods even if they had full annuity insurance. I discuss such goods in Section 4.

The FOC for h is more novel:

$$\left(d\log S_1/dh\right)BS_1u_1 = u_{0x}\left\{g'(h) + (1/1+r)dS_1/dh(x_1 - w_1(1-l_1))\right\} \qquad (8)$$

The lhs of this equation gives the marginal benefit of increased spending on health. This benefit depends on the effect of health spending on survivorship, the discount rate (since expenditures on health occur today to influence survivorship in the future), the level of survivorship in the future, and the level of utility then. The dependence of marginal benefits on the level of utility rather than on marginal utility clearly implies that marginal benefits from health spending increase as wealth increases. I show that explicitly a little later on.

The marginal cost of health spending obviously depends on g' and u_{ox}, where the latter measures the opportunity cost of spending on health. The marginal cost of spending on health is also greater when future spending exceeds future income. The reason is that an increase in future survivorship would in expectation reduce the resources available for consumption in the present when future spending exceeds future income.

By using the FOC for x, eq. (8) can be written as

$$(1/1+r)\left(d\log S_1/dh\right)S_1u_1/u_{1x} = g'(h) + (1/1+r)dS_1/dh\,(x_1 - w_1(1-l_1)). \qquad (9)$$

There are two reasons why people spend resources to raise the probability of surviving through later periods. One is for the ‘self-protection’ gain with fair insurance (see Ehrlich and Becker, 1972). That is, up to a point, an increase in spending on health raises the expected value of full wealth, net of health expenditures g', since longer life adds an endowment of additional time. The other reason is due to the difference between average and marginal utility because u is concave. The intuition is that spending more in a given year only adds marginal utility, while additional years adds average utility. See Rosen, 1988, and Usher, 1973 for an emphasis on this difference, but with no discussion of the importance of the gain in time. Both reasons are considered by Murphy and Topel, 2006.

To show how these forces combine in eq. (9), assume that utility is homogeneous of degree γ in x and l, where $\gamma \leqslant 1$. (This functional form is used by Murphy and Topel, 2006). Then u_1 can be written as

$$u_1 = 1/\gamma(u_{1x}x_1 + u_{1l}l_1), \tag{10}$$

and

$$u_1/u_{1x} = 1/\gamma(x_1 + u_{1l}/u_{1x}l_1) = 1/\gamma(x_1 + w_1 l_1), \tag{11}$$

where the last term in this equation uses the FOC between x_1 and l_1. By substituting this into eq. (9) and combining terms one gets

$$(1/1 + r)dS_1/dh(1/\gamma - 1)(x_1 + w_1 l_1) = g'(h) - (1/1 + r)(dS_1/dh)w_1 \tag{12}$$

If net full income $w_0 + (1/1 + r)S_1 w_1 - g\ (h)$ is maximized alone by choosing h, the FOC is:

$$(1/1 + r)(dS_1/dh)w_1 - g'(h) = 0. \tag{13}$$

If $\gamma = 1$, then eq. (12) reduces to eq. (13). This is to be expected, for if $\gamma = 1$, so that u is CRS, there is no difference between marginal and average utility, and the only gain from spending on improving life expectancy is the gain from increasing the endowment of time.

On the other hand, if $\gamma < 1$, so that utility is a concave function of goods and leisure, there is an additional gain from improving life expectancy because then average utility exceeds marginal utility. In that case, the lhs of eq. (12) is strictly positive, so then the rhs of eq. (12) must also be strictly positive in equilibrium. Since the rhs of eq. (12) is 0 if net full wealth alone is maximized, and since g is a convex function of h, the equilibrium positive value for the rhs of eq. (12) implies that expenditure on health must exceed its wealth maximizing level when utility is concave, and average utility exceeds marginal utility. The difference between the optimal expenditure on health and its net full wealth maximizing level depends on the degree of convexity of the health expenditure function, the degree of concavity in the period utility function, interest rates, and other variables.

2.3 The statistical value of life

Many studies have estimated how much different people need to be paid to take on additional risks to their lives, such as risky construction jobs, enlistments in the military during wartime, driving faster (which raises the probability of a deadly accident), and other risks (see the survey by Viscusi and Aldy, 2003). If they require say \$500 (WTP) to accept an additional probability of dying equal to 1/10,000 ($dS = -1/10{,}000$), the statistical value of life $v = \mathrm{WTP}/dS = \$500/1/10{,}000 = \$5{,}000{,}000$.

Equation (9) basically contains dS and WTP, where the marginal WTP is given by the rhs of the equation, and dS is explicitly on the left hand side. Then $v = (1/1+r)u_1/u_{1x}$. If u is homogeneous of degree γ, then

$$v = (1/1+r)\ 1/\gamma(x_1 + l_1 w_1) = (1/1+r)\ 1/\gamma C_1, \tag{14}$$

where C_1 is full consumption in period 1.

Empirical estimates of the statistical value of life range from about \$2 million to \$9 million for a young person in the United States, with a central tendency of \$3–5 million. We can interpret eq. (14) as giving in a fuller life cycle framework an estimate of the statistical value of life that basically equals full wealth, adjusted upwards for the degree of concavity in the single period utility function ($1/\gamma$). Does that give an estimate for the typical young American in the \$3–5 million range?

Assume an average income of \$40,000 per year from 1900 annual working hours. For every hour worked, about 1.80 hours are spent in the household sector, where I ignore 68 hours per week for sleep and maintenance. Full income is then about \$110,000 per year if an hour of household time is valued at the hourly earnings implied by the data on annual earnings and hours. If γ is taken at about ½ (Becker *et al.*, 2005, use a smaller number), adjusted full income is then \$220,000 per year. When discounted at 5%, this gives a value of life of about \$4.4 million, which is smack in the middle of the empirical estimates.

A value of life for the average American of over \$4 million seems to many to be grossly too large, but that is because there still is a tendency to think in terms of earnings alone. The first book on the economic cost of early death, Dublin and Lotka's *The Money Value of Man* (1930), estimated this cost from the loss in future earnings due to early death. But the amount that a person is willing to pay to reduce the chances of dying considers not just lost earnings, but lost utility that also includes the value of leisure time, and the differences between average and marginal utilities. At \$40,000 earnings per year and a 5% annual discount rate, the present value of the lost earnings from early death is about \$1 million, less than a fourth of my back of the envelope estimate of \$4.4 million. Therefore, the vast majority of statistical value of life comes not from foregone earnings, but from the loss of leisure time, and differences between average and marginal utilities.

3. Complementarities

The health field is brimming with complementarities: by different diseases, across ages, between health, education and training, and even between the discount rate on future utilities and health. This section demonstrates a few of the more important complementarities, and the next one discusses their implications for economic inequality.

3.1 Between diseases

The probability of surviving to any age can be interpreted as the product of the probability of surviving different diseases, where these probabilities need not be independent. In the 2 period utility function we get

$$V = u_1 + ßS_1u_2 = u_1 + ßd_1(h_1)d_2(h_2)\ldots.\mathrm{d_n}(h_n)u_2, \tag{15}$$

The term d_i gives the probability of surviving disease I from the initial period through the second period, and $\mathrm{h_i}$ gives health inputs into reducing d_i, where $d'_i \geqslant 0$.

The full annuity budget constraint is then

$$x_1 + S_1x_2/(1+r) + g(h_1, h_2, \ldots.h_n) = w_1(1-l_1) + S_1w_2(1-l_1)/(1+r) \tag{16}$$

The FOC's for x_i and l_i are the same as in eq. (7), and I will not repeat them. Optimal spending on reducing the probability of each disease is given by the equations:

$$\begin{aligned}(1/1+r)\ \partial d_i/\partial h_i d_1 \ldots d_{i-1}d_{i+1}\ldots d_n u_1/u_{1x} = \partial g/\partial h_i \\ + (1/1+r)\ \partial d_i/\partial h_i d_1 \ldots d_{i-1}d_{i+1}\ldots d_n(x_1 - w_1(1-l_1)),\ \ i = 1, 2, \ldots \mathrm{n}\end{aligned} \tag{17}$$

The health spending function may have all kinds of substitutions and complementarities between spending on reducing the probabilities of different diseases. This equation clearly shows, however, the fundamental complementarities between improvements in fighting different diseases (see the discussion by Dow *et al.*, 1999; and Murphy and Topel, 2006). An increase in the probability of surviving disease j raises the marginal utility from spending more resources on reducing the probabilities of dying from other diseases. Unless the increase in the probability of surviving j sufficiently lowers the marginal product of spending on other diseases ($\partial g/\partial h_i$), increases in the probability of surviving some diseases would tend to raise the amount spent on fighting other diseases.

During the past 30 years there has been a major improvement in fighting cardiovascular diseases among older persons (see, e.g., Cutler and Kadiyala, 2003). This analysis suggests that more attention would then be paid to fighting other diseases of old age, such as cancer, diabetes, Alzheimer's, and so forth. In fact, the 'war on cancer' started after advances began in reducing deaths from heart attacks, and increasing attention began to be paid to these other diseases of old age. Similarly, as deaths from Aids in some African countries have risen dramatically in the past 20 years, many young Africans have become less concerned about preventing other diseases since the probability of dying from Aids is so high in these countries (see Oster, 2006).

A partly related determinant of the increased spending on diseases of old age is that more individuals have reached old age, in good part because deaths from various diseases at younger ages have gone down dramatically. The increased number of persons reaching older ages has raised the productivity of investing in learning more about how to reduce the death rate from diseases at old age; see Section 8.

3.2 Between ages

To show the complementarity between investments in health at different ages, assume that the probability of surviving the initial age is $s_0(h_0)$, and the conditional probability of surviving age 1 is $s_1(h_1,h_2)$, where h_0 is the real resources spent in the initial period on surviving that period, and h_1 and h_2 are the resources spent in periods 0 and 1 respectively to raise survivorship in period 1. Then the expected utility function becomes

$$V = s_0(h_0)u_0 + ßs_0(h_0)s_1(h_1, h_2)u_1, \tag{18}$$

With budget constraint

$$\begin{aligned} s_0x_0 + (1/1+r)s_0s_1x_1 + g(h_0, h_1) + s_0f(h_2)/1+r \\ = s_0w_0(1-l_1) + s_0s_1w_1(1-l_1)/1+r \end{aligned} \tag{19}$$

This constraint assumes that some activities on improving survivorship in period $1(h_1)$ occur in the initial period, and some survivorship-improving activities for that period (h_2) occur at the beginning of period 1.

The FOC's for the h's are

$$\begin{aligned} (\partial s_o/\partial h_0)u_0/u_{0x} + (1/1+r)s_1(h_1)u_1/u_{1x}\partial s_0/\partial h_0 \\ = \partial g/\partial h_0 + (1/1+r)\ \partial s_o/\partial h_0[x_0 + (1/1+r)s_1x_1 \\ -\{w_0(1-l_0) + s_1w_1(1-l_1)/1+r\}] \end{aligned} \tag{20}$$

$$\begin{aligned} (1/1+r)\ s_0(h_0)(u_1/u_{1x})\ \partial s_1(h_1, h_2)/h_1 \\ = \partial g/\partial h_1 + (1/1+r)s_0\ \partial s_1/\partial h_1[x_1 - w_1(1-l_1)] \end{aligned} \tag{21}$$

$$u_1/u_{1x}\ \partial s_1(h_1, h_2)/\partial h_2 = f'(h_2) + \partial s_1/\partial h_2[x_1 - w_1(1-l_1)] \tag{22}$$

Equation (20) shows the complementarity between surviving in the future and surviving in the present-sometimes called competing risks (see Dow *et al.*, 1999). For it states that an exogenous increase in s_1 induces greater spending on raising the survivorship during the initial period because it raises the expected utility of surviving the next period. This result is simply saying that if a person knows he is more likely to survive some future periods, he has more incentive to try to survive to these periods. Young adults in Africa have responded less to the risk of getting Aids than have adults in the United States in part because the risk of dying from other causes is so much higher in Africa (see Oster, 2006).

Equation (21) shows a similar complementarity between the probability of surviving the initial period and spending in the initial period on raising survivorship in the later period. For an increase in s_0 raises the marginal utility of spending to raise s_1 before the realization of s_0 is known. So if the probability of surviving childhood is low, it hardly pays to spend resources in childhood that would raise survivorship at older ages.

On the other hand, spending at later ages to raise survivorship at later ages does not directly depend on the probability of surviving to these ages. For survivorship is given when a person reaches the later age, and then his behavior would only depend on variables as he looks forward- this is clearly shown by the absence of s_0 in eq. (22). What is relevant then are variables in period 1-whatever happened in period 0 is a given. However, there could be an indirect effect if the productivity of h_2 depends on h_0 or h_1, or on variables other than h_0 that affect s_0 and also h_2's productivity.

4. Health and addictions

Full annuity insurance eliminates the risk of dying in the FOC's for goods and time in eq. (7) because of the assumption that the utility function is fully separable over time. When full separability is abandoned through assuming habits and addictions, or other links between present consumption and future utility, full market insurance does not eliminate the probability of surviving from the FOC's for addictive and habitual goods. For assume two goods, x and y, where x is addictive with consumption capital, or for any other reason x depends on the accumulation of consumption capital, as in

$$V = u_0\big(x_0, x_{-1}, y_0, l_0\big) + ßS_1 u_1\big(x_1, x_0, y_1, l_1\big), \tag{23}$$

The budget constraint would be

$$x_0 + S_1 x_1/(1+r) + y_0 + S_1 y_1/(1+r) + g(h) = w_0(1-l_0) + S_1 w_1(1-l_1). \tag{24}$$

The FOC's for the y's are the same as for the x's in eq. (7). The novel FOC for x_0 is

$$\begin{aligned} dV/dx_0 = \partial u_0/\partial x_0 + ßS_1 \partial u_1/\partial x_0 = \partial u_0/\partial y_0, \\ \text{and for } x_1 \text{ there is } ß\partial u_1/\partial x_1 = (1/1+r)\partial u_0/\partial y_0 \end{aligned} \tag{25}$$

The FOC for x_1 is basically the same as before since there is no third period to pick up any habit effect of consuming x in the second period. However, the condition for x_0 does depend on the probability of surviving through the second period, and the effect of x_0 on utility in the second period. The addiction is said to be harmful or beneficial as $\partial u_1/\partial x_0 <$ or > 0.

The greater the probability of surviving through the second period, the greater the weight placed on the habitual aspect of the consumption of x. This means that dV/dx_0, the total marginal utility from x_0, is greater for beneficial addictions when the probability of surviving the second period, S_1, is greater, while the total marginal utility of x_0 is smaller for harmful addictions when S_1 is greater.

For example, there is less harm from becoming addicted to activities that lower utility at older ages, such as heavy drinking, hard drugs, smoking, or fatty foods, if the probability of surviving to older ages is relatively low. Conversely, there is greater benefit from becoming addicted to activities, such as regular exercise, religion, or a spouse if the probability of surviving to older ages is high. For example, use of hard drugs by soldiers in Vietnam increased dramatically because probabilities of surviving to old age were significantly reduced by the war there. After they returned to a more normal life, addiction rates went down dramatically among the vast majority of soldiers who had been using these drugs (see, e.g., Robbins, 1993).

The foundation of this paper is that S_1 is not simply given but is affected by expenditures on health, h. An increase in h raises S_1, which in turn by the previous discussion would raise the demand for x_0 if x is beneficially addictive. It lowers the demand for x_0 if x is harmfully addictive. This means in essence that good health is complementary with good habits and addictions, and bad health is complementary with bad habits and addictions. Consumption capital in the form of good or bad habits and addictions, or in other forms, cannot be insured against by annuity markets.

It has long been observed that healthier people are less likely to smoke, to be overweight, to be on drugs, to neglect breakfast, and even to be more religious (see, e.g., Grossman and Kaestner, 1997). The explanation typically involves causation from these habits to better health, and such causation is probably of importance. However, the analysis in this section shows there is also causation from better health to better habits since the cost of bad habits is greater for persons in good health with better life expectancy.

5. Education and health

Complementarities are also important between expenditures on different forms of human capital. I will consider only the complementarities between health and schooling since these have received a lot of attention, although the same analysis applies to health and training, and health and migration. An increase in survivorship at later ages raises the returns from investments in education because educational costs come at earlier ages and returns at later ages. (see Becker, 1964, 1993: Meltzer, 1992; Ehrlich, 2000).

To show this with the two period example I have been using, suppose the cost of education, equal to E, occurs in the initial period, and that the investor is certain to survive this period. If the return next period is in the form of a higher wage rate, then $w_1(E)$ with $w_1' > 0$.

For several reasons, an increase in education also tends to increase survival rates. Abundant evidence indicates that more educated persons take better care of their health even with a given spending on medical care by visiting better doctors, taking their prescribed medicines more regularly, eating more nutritious diets, and in many other ways. This suggests that $S_1(h, E)$, with $\partial S_1/\partial E > 0$, so that an increase in schooling raises life expectancy.The utility function now becomes

$$V = u_0 + ßS_1(h, E)u_1, \tag{26}$$

and the budget constraint with full annuity insurance is

$$x_0 + S_1 x_1/1 + r + E + g(h) = w_0(1 - l_0) + S_1 w_1(E)(1 - l_1)/1 + r. \tag{27}$$

The FOC'S with respect to x_i, l_i, and h are the same as before. In addition, the optimal investment in education is giving by the equation

$$\begin{aligned}(1/1 + r)S_1 w_1'(E)(1 - l_1) + (1/1 + r)\, ß\partial S_1/\partial E u_1/u_{1x} \\ = 1 + (1/1 + r)\, \partial S_1/\partial E[x_1 - w_1(1 - l_1)]\end{aligned} \tag{28}$$

The lhs of eq. (28) gives the total benefit from increased expenditures on education. The first term is the usually discounted earnings due to a higher wage rate from greater education. The second term is the increased utility that comes from a higher survivor rate during the second period as a result of having greater education. The market effect is the higher earnings, and the psychic effect is the increased value place on a life with a higher probability of surviving in the future.

The rhs gives the cost of increasing education. The first term is the marginal dollar spent on education, and the second term is the increase in the expected resources that would be taken from the first period to finance greater spending during the second period relative to earnings in that period. Of course, spending in the second period could be lower than income in that period if ß were sufficiently low relative to r, if w_1 were high relative to w_0, etc.

Equation (28) implies that an increase in education raises life expectancy in two ways. Greater education raises expected wealth net of the spending on education. That produces a wealth effect that would increase spending on health, and thereby would raise survivorship in later years. Education also directly raises survivorship by making a person more productive at investing in health through better information about doctors, healthy lifestyles, and other information related to better health.

This analysis implies that greater education raises health, but not necessarily spending on health. The wealth effect from higher earnings does raise spending on health, but the effect on survivorship probabilities can imply better survivorship with lower equilibrium spending. It all depends on the elasticity of the derived demand for h, and how E affects the marginal product of h in raising S_1.

6. Health and the discount rate

There is debate in the health area about the interpretation of the empirical positive relation between health and education. Some argue the causation is from schooling to achieve better health, while others claim reverse causation from better health to more schooling. Both those forces appear in eq. (28). However, still another interpretation is that neither of these views is the most important, and the main causation is from persons with lower discount rates on future utility to both greater investment in schooling and greater investment in health by these people (see Fuchs, 1982).

This last interpretation takes the distribution of discount rates as given, and argues that persons with relatively low discount rates get selected into investing more in both education and health. Such selection happens, but it is a mistake to take discount rates as given. Discount rates reflect attitudes toward the future, and they are affected by spending of time, money, and energy on 'imagination capital' that helps to reduce how much future utilities are discounted in decision-making.

Becker and Mulligan (1997) show that the incentive to invest in raising discount rates is positively related to the magnitude of expected future utilities. Their analysis is applicable to the interaction with health and education that we are considering. Modify the utility function in eq. (23) to allow ß to depend on C, imagination capital, with ß' $\geqslant 0$, ß' < 0, and let $\Phi(C)$ be the cost of C, with $\Phi' > 0$, $\Phi'' > 0$. Then the budget constraint that includes spending on C would be:

$$x_0 + S_1 x_1/(1+r) + E + g(h) + \Phi(C) = w_0(1 - l_0) + S_1 w_1(E)(1 - l_1)/(1+r) \tag{29}$$

This budget constraint assumes that spending on imagination capital C is an investment, so that resources are spent earlier-perhaps in childhood- to influence discount rates at later ages.

Then the FOC for optimal investment in C is

$$d\log ß/dC\,(1/1+r)\,S_1 u_1/u_{1x} = \Phi'(C). \tag{30}$$

The marginal utility from investing in C depends positively on both the level of future utility, and the probability of surviving to the future. The former result implies that richer persons tend to discount the future less (as emphasized by

Becker and Mulligan, 1997). Since education raises wealth, given the FOC in eq. (28), this equation implies that more educated persons discount the future less. It is possible that E also directly enters ß, with $\partial ß/\partial E > 0$, which would be another reason why educated persons have lower discount rates.

The positive relation between the marginal utility of investing in imagination capital and the probability of surviving in the future implies that healthier persons will also have lower discount rates because they have greater incentive to invest in reducing the discount on future utilities. So healthier people would have lower discount rates not only because people with lower discount rates invest more in health, but also because healthier people invest more in lowering their discount rates.

This latter force has been entirely neglected in the health literature on the relation between health and the discount rate. Moreover, the analysis implies that healthier persons both invest more in education and in lowering their discount on future utilities. With lower discount rates, healthier people will invest more not only in human capital, but also in savings in the form of financial, housing, and other assets. Lower discount rates on future utilities reinforce the earlier conclusion that healthier persons are more likely than more sickly persons to develop beneficial habits and less likely to develop harmful habits.

In the expected utility formulation, since both ß and S_1 are discount rates on future utilities, it is no surprise that the marginal utility of spending to increase S_1 has the same form as the marginal utility to increase ß. However, while it is generally accepted that individuals can affect their probabilities of surviving different ages by how they spend their money, time, or effort, it is controversial whether individuals can affect their discount rates by how they spend their resources (see, e.g., the negative view on this by Elster, 1997).

That is just a legacy, I believe, of the tradition in economics of taking discount rates as given, and outside the control of individuals. It was once a tradition in economics too of taking death rates as given, and not affected by individual choices. That tradition has vanished as economists realized the many ways that individuals can affect their mortality rates. I expect the tradition of taking discount rates as 'exogenous' to also vanish as economists appreciate the many ways that parents affect their children's discount rates, and adults affect their own rates.

7. Inequality, health, and other human capital

One of the pioneering articles on human capital (Mincer, 1958) assumes that the present value of earnings, net of education costs, are the same at all levels of education when earnings are discounted by the market interest rate. In essence, this assumption eliminates all inequality in the present value of earnings by education level, net of education costs, including foregone earnings.

Further analysis and formulation of investments in human capital incorporated the older idea of non-competing groups into education and other human capital choices (see Becker, 1967, reprinted in Becker, 1993). Moreover, Becker and

Chiswick (1966) showed that under certain strict conditions, actual rates of return on investments in education, which greatly exceed market interest rates, could be estimated from data on differences in wages across education classes.

These formulations established the foundation for the conclusion that education is a significant source of inequality in earnings, even after netting out all education costs. Further research demonstrated that individuals differed significantly also in other determinants of wellbeing, such as health, savings, habits, and marital stability. If these characteristics were negatively correlated with each other and with education, then the overall degree of true inequality might be much less than that in earnings due to differences in education.

However, research has demonstrated that virtually all valued characteristics are positively rather than negatively correlated with education (see Elias, 2005). These include health and life expectancy, marital stability, achievements of children, law-abidingness, beneficial as opposed to harmful habits, savings rates, propensities to vote and other activities in elections, and still additional traits. Why this is so has not been answered in anything approaching proper generality.

I propose that the answer comes from the general complementarities among different classes of human capital. I have built the analysis in this paper around health as measured by the probabilities of surviving to later ages, and I expanded out from that to consider the relation between health and education, and other forms of human capital investment. The analysis shows that different forms of human capital investments are complements, not substitutes. These forms include not only education and health-which is well known- but also health and good habits, health and lower discount rates on future utilities, and improvements in life expectancy from different diseases and at different ages.

This means that people who have better life expectancies also have higher earnings and greater education, save a larger fraction of their permanent incomes, have ‘better’ habits, and also have greater conditional life expectancies, given that they reach any age. So characteristics like earnings, habits, discount rates, and savings do not offset inequality in life expectancy, but reinforce that inequality to contribute to a still widen inequality in overall welfare. Groups are ‘non-competing’ not in a single determinant of wellbeing but in a whole series of interrelated determinants.

8. Investment in R&D and population

8.1 A model of drug innovation

Much of health economics is concerned with health delivery systems. These include issues related to government financing of medical care, co-payment arrangements, health savings accounts, moral hazard and self selection in both private and public health coverage systems, incentives to economize on spending by hospitals and doctors, and many other issues. These are all important, and they deserve the attention they receive. Basically they all affect in different ways the costs of medical care to consumers, and their incentives to take various actions.

I will neglect them to concentrate on another issue: the role of population in affecting the incentives to spend on medical R&D. To generate this analysis, consider a pharmaceutical company that can spend R to produce a medical innovation also, I, in the form of pills that raise the probability of surviving particular ages A. Once innovated, the cost of producing each pill equals c. The demand for these pills is given by

$$D(I) = N(I, A)p^{-\varepsilon}, \tag{31}$$

where p is the market price of a pill, and ε is the elasticity of demand for these pills.

To determine the market price, I assume that the innovator is given a patent that provides effective monopoly power in the market for I for T years. After T years pass, generics enter and push the market price of each pill to c, its cost of production. So the innovator earns monopoly profits for T years and zero profits after that. Since profits on the pills related to I equal $\Pi = pD - cD$, given eq. (31), profits are maximized when price equals

$$p^* = (\varepsilon/\varepsilon - 1)c \tag{32}$$

Then the present value of discounted profits equal

$$W = \left[N(I, A)(\varepsilon/\varepsilon - 1)^{-\varepsilon}\left(c^{1-\varepsilon}/\varepsilon - 1\right)\right]\sum_{1}^{\mathrm{T}}(1/1 + r)^j. \tag{33}$$

Clearly, the wealth from selling the pills is greater the lower is c, the smaller is ε, the bigger is T, and the larger is N. For the innovation to be worth the investment, it is necessary that

$$W = \left[N(I, A)(\varepsilon/\varepsilon - 1)^{-\varepsilon}\left(c^{1-\varepsilon}/\varepsilon - 1\right)\right]\sum_{1}^{T}(1/1 + r)^j > R. \tag{34}$$

This is more likely to hold the longer is the effective patent life (T), the lower the elasticity of demand for the pill that is created, and the lower the cost of producing each pill.

This innovation is more likely to be profitable when N is larger, where N can be interpreted as the number of people demanding the pill when $p = 1$. This measure of demand depends on the number of people in the age groups (A) where the disease attacked by the pill is more prevalent, the fraction of this number that contract the disease, the income of the individuals getting the disease, public subsidies to buying the pill, and other variables. Obviously, incomes matter since as we have seen, willingness to pay to improve survivorship depends positively on incomes. Previous sections also show that the demand for this innovation would be greater, the greater the probabilities of surviving other diseases at these and later ages.

8.2 Population and the demand for innovations

I want to emphasize now the importance of sheer numbers, as measured by the number of persons in vulnerable ages groups for the particular diseases treated by innovations (for some of the theory on this applied to medical research, see Becker, 2000; Cerda, 2003; and Acemoglu and Linn, 2004). For the larger this population, the greater is the profitability of medical innovations, holding incomes, the probabilities of surviving other diseases, government subsidies, and other relevant variables constant. This is not surprising since innovations, and medical innovations in particular, have increasing returns to scale because the cost of producing innovations is independent of the scale of the demand for the innovated product or service.

The importance of the size of the market is recognized in The Orphan Drug Act of 1982. This Act stipulates that drugs which are innovated to treat diseases which have a market with fewer than 200,000 sufferers receive seven years of guaranteed market exclusivity. Studies have shown that this longer patent protection has stimulated additional expenditures on research on orphan drugs by pharmaceutical companies (see Grabowski, 2005). Still, the most profitable drugs are those that cater to large markets, such as drugs that treat high blood pressure, or erectile dysfunction, or high cholesterol.

Additional evidence on the size of the market comes from studies of the number of new chemical entities to treat diseases of different ages as a function of the relative number of persons at these ages. Figures 1 and 2 from Cerda (2003; also see Acemoglu and Linn, 2004) give the number of new chemical entities introduced in the United States to treat diseases contracted by persons at particular ages as a function of the number of persons at these ages. As the number of persons aged 45–64 increased, Fig. 1 shows that the number of drugs introduced to treat diseases that mainly affect people of these ages also increases. Figure 2 shows similar results for drugs introduced to treat diseases that mainly affect people aged 65 and older.

8.3 Population and the aggregate value of declines in mortality

Table 2 gives results from the well-known study by Murphy and Topel (2006) on the value of the gains to Americans from the decline in death rates between 1970 and 2000. They assume a statistical value of life to a young person of $5 million, and make various reductions to this number for older persons. They subtract all the growth in spending on medical care during this 30 year period. They then add together the value of the improvements in death rates at each age over the number of individuals at each age, with separate calculations for men and women. Their main results are in the fifth column of Table 2.

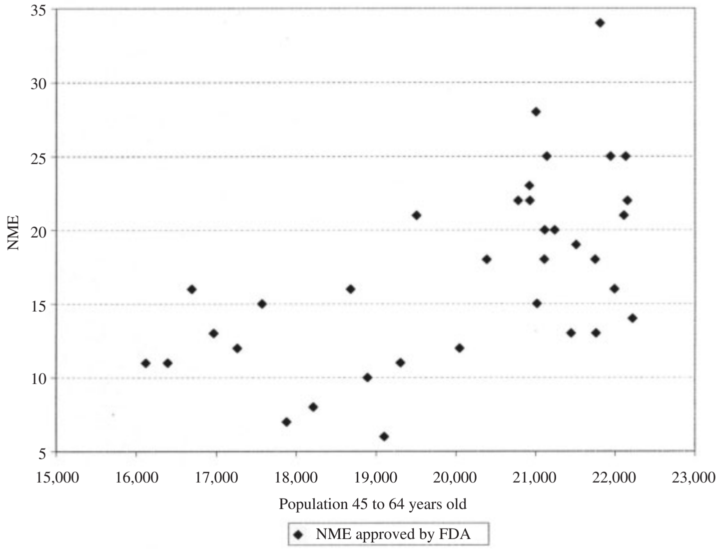

Fig. 1. New molecular entities approved by FDA and US Population of 45 to 64 years old, 1962–1997
Source: Cerda, 2003.

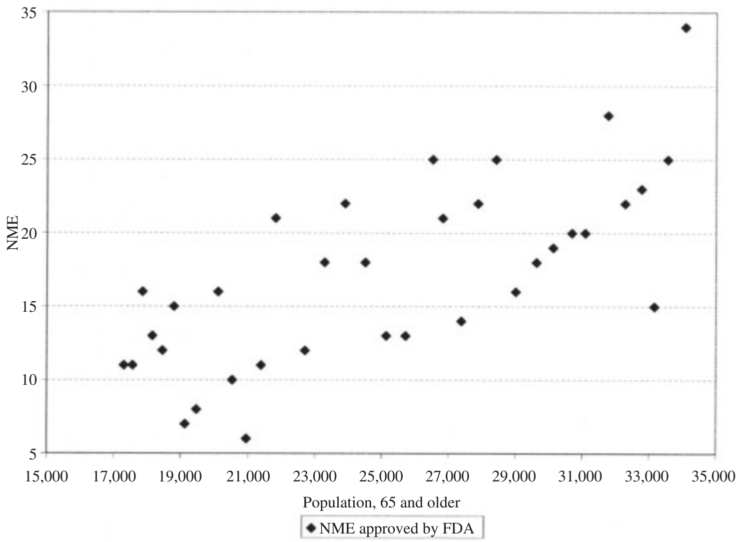

Fig. 2. New Molecular Entities approved by FDA and US Population of 65 and older, 1962–1997
Source: Cerda, 2003.

Table 2 Estimate gains net of the increase in health expenditure, 1970–2000 (in billions of dollars)

United States	1970–80	1980–90	1990–2000	1970–2000
Gross gains	$47,214	$24,538	$23,593	$96,345
Increase in expenditures	$8,206	$14,928	$11,591	$34,725
Gains net of expenditure growth	$39,008	$9,611	$12,001	$60,620
Expenditure increase as a % of gains	17.40%	60.80%	49.10%	36.40%
OECD	**1970–80**	**1980–90**	**1990–2000**	**1970–2000**
Gross gains	$126,276	$76,940	$70,264	$241,168
Increase in expenditures	$28,118	$15,889	$7,548	$51,555
Gains net of expenditure growth	$98,158	$61,051	$62,716	$189,613
Expenditure increase as a % of gains	22.27%	20.65%	10.74%	21.38%

Source: United States: Murphy and Topel, 2006; OECD: Gary S. Becker's Calculations.

These numbers are huge: combined for men and women they add up to about $60 trillion net of medical expenditures. Compare this to a GDP for the US of about $11 trillion, and one can appreciate how large is the aggregate value of the gain to the American people from the sizable improvements in their life expectancy since 1970. Their assumption of a $5 million statistical value of life for a young adult can be questioned, and perhaps a lower number is more appropriate (see Section 1.3). But their estimates of the gain from reduced death rates would be huge even if $3 million or $2 million rather than $5 million were used to measure the value of life.

For it is the scale of the US population of about 300 million persons that is the main driver of these aggregate gains, not the numbers used to measure the value of life to individuals of different ages. Even small gains to the average individual become large when multiplied by 300 million. This shows the power of increasing returns as population grows to magnify the gains from improvement in survivorship rates.

The full effects of improved mortality are even larger than this since the market for drugs is worldwide, and not restricted to the US. Suppose we consider just the OECD countries that excluding the US has a combined population of 867 million. To shortcut the calculation I assume a representative person in each country with a value of life that equals $5 million times the ratio of the *per capita* income in that country to the US *per capita* income. I value the change in life expectancy between 1970 and 2000 in each country by its calculated value of life, and multiply that result by the country's population. The results are summed over all OECD countries, including the US, where the increase in medical spending is subtracted to give the 'net' results in the bottom half of Table 2.

These values are about three times as big as the values for the US alone. Combined over men and women they total to about $190 trillion, an amazing sum. This means that enormous value is placed on the declines in mortality in

OECD nations during the past couple of decades. Moreover, this value is increasing in the size of their populations, given their *per capita* incomes.

From the view point of public and private spending on medical R&D, the size of the market in rich countries alone is enormous for innovations that can treat reasonably common diseases. A larger and more densely packed population also has some negative effects because that makes it easier to transmit infectious diseases. At the same time, however, a larger population also provides stronger incentives to find ways of treating infectious and all other diseases. So the negative attitudes toward larger world populations (even for a given *per capita* income) that is common among demographers, scientists, and many others is largely wrong with regard to the determinants of innovations in basic and applied medical research.

8.4 The cost of medical innovations and population

I have assumed so far that the cost of a medical innovation is fixed at R, and I have compared that cost to the revenue from innovations. This cost may vary positively (diminishing returns) or negatively (increasing returns) with the stock of past medical innovations, and negatively with advances in scientific knowledge, such as the mapping of the human genome. The cost of innovation may also vary with mortality rates since it may be more difficult to lower mortality rates further when they are already low.

Suppose that the cost of medical innovations increases as the stock of these innovations grows and mortality rates fall. Studies do in fact indicate that there has been a significant increase over time in the cost of producing a medical innovation that receives FDA approval (see DiMasi, *et al.* 2003). At the same time, however, these innovations result in a larger number of persons surviving to different ages. The increased numbers and also the lower mortality rates raise the demand for additional innovations. Which force dominates, cost or demand, then depends on how sensitive each is to changes in innovation stocks and mortality rates.

To discuss this in terms of our maximizing condition eq. (31), express this equation as

$$N(T)g - R(T) = W(R, T), \tag{35}$$

where g is the discounted profit per ‘pill’, W is the net wealth from an innovation with cost R at time T. If g remains fixed over time but N and R change for the reasons discussed in the two previous paragraphs, then we have

$$dW/dT \geqslant 0 \text{ as g } dN/dT - dR/dT \geqslant 0. \tag{36}$$

Perhaps very little pharmaceutical research was conducted up to the mid-point of the 20^{th} century because R was relatively high compared to g and population as measured by N in eq. 36. This made W negative for any significant investments in medical R&D. But advances in medical-biological knowledge were lowering R, and the growth of population and in *per capita* incomes that resulted from technological advances and declines in mortality were raising *N*. As a result, *W* changed from negative to positive for many projects, and the boom in pharmaceutical and bio-tech research began.

As I mentioned earlier, in recent years the cost of pharmaceutical R&D has been growing, but so too has demand because of the growth in incomes and population. Since expenditures on medical R&D have been growing, any growth in the cost of R&D for a representative innovation relative to discounted profits per pill (R/g in eq. (35)) may have been compensated by a growth in N, due in good measure to population growth, especially at older ages where medical interventions are more needed.

8.5 Explaining the declines over time in age-specific death rates

The analysis in previous sections can be applied to understanding the pattern of the decline over time in age-specific death rates. Figures 3 and 4 show the improvements in life expectancy at birth during the 20th century, and also improvements in life expectancy at ages 45 and 65. Life expectancy at birth incorporates life expectancies at all later ages. This life expectancy rose rather steadily, and remarkably, during this past century. Life expectancies at ages 45 and 65, however, changed little during the first half of that century, but both increased rapidly during the second half. This means that mortality declined mainly at younger ages during the first half of the century, and mainly at older ages during the second half.

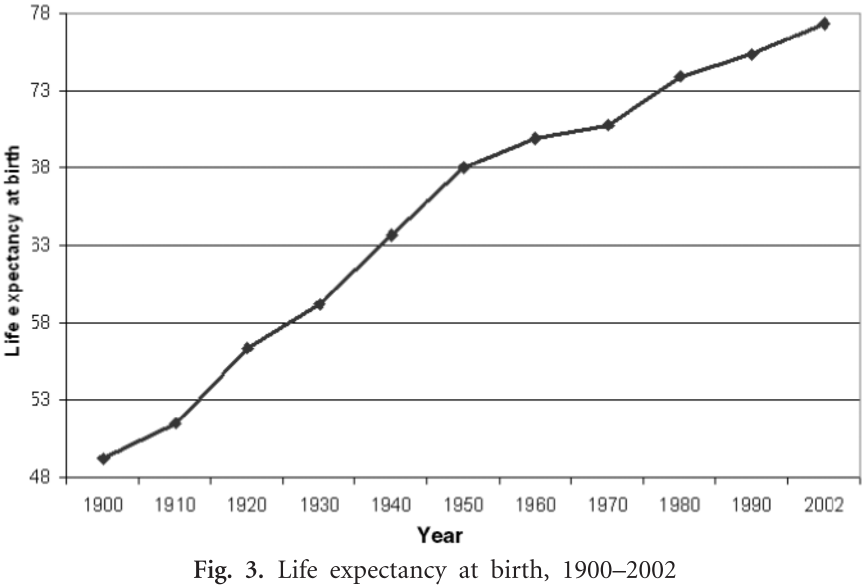

Fig. 3. Life expectancy at birth, 1900–2002
Source: National Vital Statistics Report, 2002

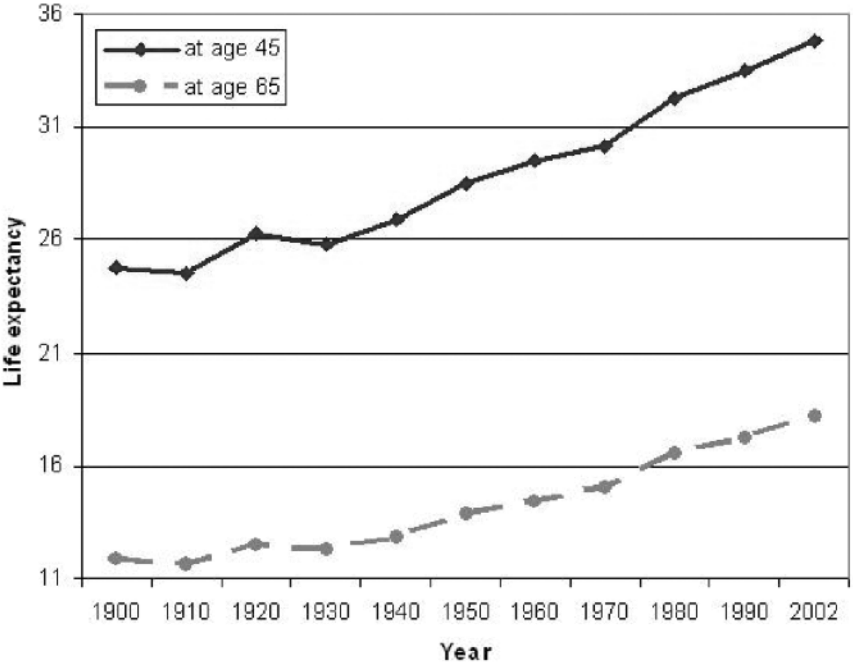

Fig. 4. Life expectancy at ages 45 and 65, 1900–2002
Source: National Vital Statistics Report, 2002

This pattern of changes in mortality appears to fit the pattern that would be predicted by the theory developed in this paper. At the beginning of the 20th century, there were many young people because of high birth rates and a young adult population. At the same time, infant and child death rates were quite high. For both reasons, the market for vaccinations and other innovations that would reduce deaths from childhood was large. Effort was concentrated on diseases of young age, such as diphtheria and scarlet fever, greater cleanliness in delivering children, and in the quality of children's diet. The return to investments in reducing mortality among young persons was very high at that time because there were lots of young people, and their survival of childhood would bring benefits at later ages.

By 1950 deaths in childhood were reduced so much that a small fraction of persons died before age fifty. As a result of this and the growth in the number of older persons, the emphasis in medical research shifted toward diseases that struck primarily at older ages, such as cardiovascular disease, strokes, and various forms of cancer. During the second half of the 20th century, substantial progress was made in fighting these and other diseases of older ages. For the first time, new drugs began to play a major role in fighting diseases, such as antibiotics, drugs to lower blood pressure and cholesterol, drugs to fight breast and prostate cancer, drugs against lymphoma, and so on for other diseases- for the reasons discussed in the previous section. The result was the substantial progress shown in Figure 4 in reducing death rates after age 45, and increasingly after age 65.

We have seen in Section 3.1 that a decline in death rates from one disease increases how much individuals are willing to spend to reduce the probability of dying from other diseases that strike at the same or later ages. This means that the development of a drug that reduces the probability of dying from one disease increases the demand for drugs to fight 'competing' diseases. The price to consumers of drugs that fight disease J that affects people at age A would enter into the demand function for disease I in eq. (31) with a negative exponent. The negative exponent indicates that diseases I and J are complements:

$$D_i(I) = N(I, A)p_j^{-e}p_i^{-\varepsilon} = N\big(I, A, p_j\big)p_i^{-\varepsilon}, \quad \text{with } \partial N/\partial p_j < 0 \tag{37}$$

Even if improvements in reducing the mortality from one disease did not affect an individual's demand for treatments that combat other diseases, there would be an increased aggregate demand for treatments for other diseases that are prevalent at the same or later ages as this disease. The reason is that more persons would survive these and later ages, because more survive disease I, and the larger number of survivors raise the demand for treatments of other disease at these ages (see Becker, 2000 and Cerda, 2003). That is, N in eqs (31) and (34) is affected by the number of survivors to age A, which increases when treatments improve for diseases like heart conditions and strokes that affect people at age A. So returns to population scale, in addition to complementarities between diseases at the individual level, produce complementarities in the aggregate demand functions to treat different diseases that strike at related ages.

One of the paradoxes of behavior in recent decades in the United States is that on the one hand, people are greatly concerned about their health, as shown by the increase in exercise clubs and the popularity of exercise videos, and from almost daily reports on drugs, diets, and other health factors in major newspapers and TV news programs. On the other hand, a considerable amount of behavior appears to contribute to worse health, such as the consumption of boutique ice creams with high fat content, the reduction of exercise by teenagers and their substitution toward sedentary activities like computer games and chat rooms, and of course the resulting significant increase in weight of adults, and even more so of teenagers.

This unusual combination of heightened concern about health, and behavior that appears to reduce health can be understood by recognizing both that people are forward looking, and that they expect new drugs to be developed in the future. Consider the sizable increase in average weight (see Philipson and Posner, 2003), including a significant growth in obesity, during the past 25 years. This has been a source of concern by health officials and others since being overweight, and particularly being obese, is currently associated with increased risk of diabetes, high blood pressure, heart problems, and other diseases. I say 'currently' because it is reasonable to expect that drugs to prevent many of the negative consequences of obesity will be developed in the future. After all, the past several decades have seen the development of much better drugs to control high blood pressure, drugs to reduce cholesterol levels, and drugs to make diabetes itself more manageable.

To show how the anticipation of medical progress can affect current behavior, consider the utility function

$$V = u_0(x_0, l_0) + ßS_1(h, I_1, x_0)u_1(x_1, l_1), \tag{38}$$

where I_1 are the drugs available in period 1. Drug innovation means that $I_1 > I_0$. Assume that x is harmful to health, so that $\partial S_1/\partial x_0 < 0$, and that drugs raise health, so that $\partial S_1/\partial I_1 > 0$. I make the important assumption that an increase in I_1 reduces the negative effect of x_0 on S_1; that is, $\partial^2 S_1/\partial x_0 \partial I_1 > 0$.

The FOC for x_0 is

$$\partial V/\partial x_0 = u_{0x} + ßu_1 \partial S_1/\partial x_0 \tag{39}$$

The negative effect of x_0 on S_1 discourages consumption of x in the initial period. However, an expected increase in the availability of substitute drugs in the future, ΔI_1, would reduce that negative effect compared to what it would be if I remained fixed at its initial value, I_0. Hence the increase in I would increase x_0. Interpret x, for example, as playing computer games or eating a lot that contribute to a growth in weight over time. This weight increase with the medical knowledge in the initial period might significantly raise the prospects of dying prematurely in the future, while the growth in knowledge could significantly lower that mortality risk.The harmful effect of x on future health would discourage eating a lot and playing computer games even though they give pleasure. However, an expectation that medical knowledge (I) would improve over time would encourage an increase in x_0 compared to its level with no medical improvement, as long as consumer behavior were sufficiently forward looking, and consumers have enough information about trends in the introduction of drugs and other medical advances.

Figure 5 shows the change in new molecular entities approved by the FDA from 1941 to 1996. This growth rate decreased during the 1960s and increased starting in the 1970s. If consumers were aware of the continuing growth in available drugs, then younger persons who were very concerned about their life expectancy might nevertheless increase activities that would appear to lower their future survival rates. The acceleration in the rate of introduction of new drugs should have caused an increased consumption of goods that harm future health. Perhaps that helps explain the acceleration in weight gain among teenagers that began about 1980.

The analysis of population scale and complementarities across ages is relevant to the old and difficult question of what are the limits to length of human life? Answers are usually given in strictly biological terms, such as potentials for genetic modification, or evolutionary forces that selected for individuals who used their 'energy' at reproductive ages and have little remaining at older ages. According to these evolutionary explanations, humans tended to deteriorate rapidly after they became of little use either directly or indirectly in reproduction. Biology is

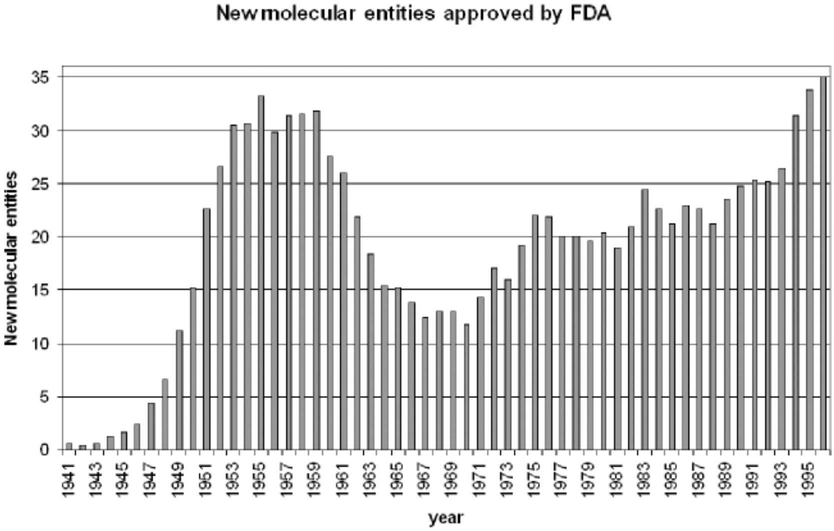

Fig. 5. New molecular entities approved by FDA, 1941–1996
Source: Cerda, 2003, smoothed

obviously of enormous importance in determining length of life, but incentives and economic considerations are also relevant, and they have been neglected.

I interpret the biological explanation as implying reasonably that the cost of finding effective treatments for diseases that strike adults rises with their age, and is much harder at old ages than at younger ages. However, if the number of people reaching older ages rises substantially because of progress in fighting diseases that strike at earlier ages, the aggregate willingness to pay for treatments at old age also would rise substantially. This much higher aggregate willingness to pay for treatments to reduce deaths at older ages would justify spending on research that tries to find ways to treat diseases of old age, research that was not worthwhile in the past when there were many fewer older persons.

On this interpretation, progress against diseases of very old age is slowed considerably by biological factors, but it is speeded up by the sharp growth in the number of individuals at ages where they would benefit substantially from such medical progress. The net effect on the extension of life expectancy at old ages depends on both biological and economic forces. The net effect has clearly been positive in recent decades since great progress is being made to extend life at older ages that would have been considered impossible by biologists and demographers not that long ago.

8.6 The potential cost of an avian flu pandemic

The flu pandemic of 1918–19 that is estimated to have killed 50 million persons worldwide (see Kolata, 1999) was the major epidemic by far of the 20^{th} century. The number killed is about 2.8% of world population at the time (see the second row of Table 3). There is present concern about a major avian flu pandemic during the

Table 3 Estimated value of loss from a potential avian flu epidemic

	US Estimate	World Estimate
Number of deaths from the flu 1918–19 divided by population at that time	2.80% × 1.8bn = 50m	2.80% × 1.8bn = 50m
Estimate of possible deaths from avian flu	300m × 2.8% = 8.4m	6 bn × 2.8% = 168m
Ratio of US GDP *per capita* over World GDP *per capita*		$8,800 / $40,100 = 0.22
Estimate of value of loss from a potential avian flu epidemic	$3m × 8.4 m = $25 tr	$3m × 168m × 0.22 = $110tr

Source: Gary S. Becker Calculations.

next few years if that virus becomes more easily transmitted from fowls to humans and between humans. I do not know the likelihood of such a major pandemic, but the Congressional Budget Office has taken the possibility seriously. It estimates that if an avian flu pandemic emerges with the same relative impact as the flu pandemic in 1918–19, US GDP would decline about 5% for one year (CBO, 2005).

The CBO does not, however, attempt to estimate the value of the loss in life from such an enormous pandemic. We do that now. Row 3 of Table 3 shows that 168 million people would die from an avian flu pandemic of comparable virulence to the one in 1918–19, for that number is 2.8% of the present world population of just over six billion persons. 2.8% of the US population would equal about 8 ½ million people. If following my earlier discussion, each life is valued conservatively at $3 million, row 5 of the table shows that this means an aggregate loss to the US of around $25 trillion. Since this figure dwarfs the 5% loss in GDP, which is about $550 billion, the number of people who would die is clearly the principle loss from a major pandemic.

To make a similar calculation for the world, assume that each country would lose the same 2.8% of it population, and the statistical value of life in a country equal $3 million times the ratio of the country's *per capita* income to the US' *per capita* income. Combining these calculations and summing over all countries gives the estimate in row 5 of Table 3 of a worldwide loss of about $110 trillion from an avian flu pandemic comparable in its destructive power to the flu pandemic during and after World War I.

While this is an enormous loss, how much precautionary actions are justified depends on the probability of having such a serious pandemic. If the probability is 1/1000 of having a pandemic during the next decade of the same order of severity as the flu pandemic of 1918–19, then the expected worldwide loss in the value of lives would be about $100 billion, which is big enough to justify hurry up efforts to develop vaccines and other protections. On the other hand, if the probability of such a pandemic is only 1/100,000, then the expected worldwide loss is only $1 billion, and crash vaccine and other programs do not seem urgent. I leave it to the epidemiologists to supply information that could lead to credible estimates of what reasonable probabilities are.

9. National income accounts and changes in world inequality

National income accounts, developed in the 1930s, were designed to measure transactions that go through the marketplace. It was recognized from the beginning that many important determinants of welfare do not go through the market place, such as time spent at home tending children, in leisure, and other activities. These accounts also do not incorporate the value of increased life expectancy since any extra hours available for non-market work are not counted, and many other benefits of living longer are also ignored.

Nevertheless, both levels and changes over time in measures like GDP *per capita* are used to make comparisons of levels and changes in utility or welfare. To get a better measure of welfare, one should try to measure the commodities produced through household production, commodities that depends on the goods purchased in the marketplace-what GDP tries to measure- the time directly supplied to households, and household technology that may differ across households and change over time.

Welfare during any year for any person would be approximated by full income deflated by an index of commodity prices that depend on the prices of goods, wage rates, and household technology:

$$R = \sum \pi_{\mathrm{i}} Z_1 = (wT + K)/\pi. \tag{40}$$

In this equation, w is the wage rate, K is non-earned income, Z_i is the ith home produced commodity that depends on goods purchased, household time, and productivity in producing the ith commodity, π_{i} is the average shadow price of Z_i which depends on the prices of goods that help produce Z_i and the wage rate, and π is a price index of the different π_{i}. Equation (40) is the household production underpinning of eq. (10) where utility is shown to depend on full consumption of household commodities. A quantity index of the Z_i would be a measure of full commodity consumption.

A measure such as in eq. (40) in principle would pick up changes in the cost of producing better mental and physical health through a reduction in the shadow price of better health quality. It would help pick up much of the gain from improvements in survivorship probabilities because when summed over all people it would measure the real full income of all survivors, not just their money incomes alone, including any effects on the relevant price deflator for full incomes. My empirical discussion to follow will not attempt to redo the household accounts to incorporate longer life expectancy (see Usher, 1973, for an interesting early discussion). I will simply report calculations about changes in international income inequality that are greatly affected by whether or not adjustments are made for improvements in life expectancy.

It is well known that the *per capita* incomes of nations that were relatively poor in 1960 did not grow much faster, if at all faster, during the subsequent 40 years than the *per capita* incomes of nations that were relatively rich then. The exact results depend on which nations are included, whether or not the country data are weighted by population -since the two largest nations, China and India, grew rapidly from 1985 onwards- and other adjustments. Row 1 in Table 4 gives the regression coefficient in a population weighted regression for about 100 countries of the growth in *per capita* income on the level of *per capita* income in 1960. These data show no regression to the mean in *per capita* income from 1960–1990, and some degree of regression to the mean from 1960–2000 because of the rapid growth in the '90s of China and India.

It is also known that mortality declined faster in poorer countries during the past 40 years than in richer countries (see, e,g, Preston, 1975; and Becker, *et al.*, 2005). The UN computes aggregate social indicators by weighting the percentage changes over time in both *per capita* income and life expectancy by 1/3, and also weights the percentage growth in years of schooling by 1/3 as well. These weights are completely arbitrary (see the criticism by Philipson and Soares, 2002), and involve much double counting since a considerable part of the benefit from higher education is captured by the improvements in income and health.

The results in the second row of Table 4 do not weight education changes separately, and combine increases in life expectancy and increases in *per capita* income by using estimates of willingness to pay for each country. Becker, *et al.* (2005) assume the statistical value of life for the United States is $3 million, and for each other country the value of life of the average person is assumed to be $3 million multiplied by the ratio of *per capita* income in that country to US *per capita* income. They add the value of the increase in life expectancy to the increase in *per capita* GDP to get a measure of the aggregate growth in *per capita* full income that is based on willingness to pay rather than the arbitrary weights used by the UN.

The results of their calculation are reproduced in the second row of Table 4. The degree of regression to the mean in full income is significant even between 1960–90, and from 1960–2000 it is twice the size of the estimated degree of regression to the mean in *per capita* incomes. These results show that when the national income accounts are adjusted for improvements in life expectancy, there has been substantial convergence in inequality between rich and poor nations during the past 40 years, even while there has been little convergence in *per capita* incomes.

Becker *et al.* (2005) also shows that the degree of convergence varies greatly across disease categories: it was large for infectious and respiratory diseases, and actually negative for cardiovascular diseases and cancers. The theory presented

Table 4 Evolution of cross-country inequality in full income, 1960–2000

Regression to the mean over 1960	1990	2000
Income *per capita*	−0.01	−0.13
Full income	−0.1	−0.26

Source: Becker *et al.*, 2005.

earlier can explain this divergence across diseases. Since poorer countries have relatively few persons who reach the older ages where heart disease and cancers are more predominant, poorer countries are less concerned about diseases of older ages than about diseases of childhood, and so put fewer resources into reducing mortality from diseases of older ages. In fact, much of the progress in poorer countries was made in reducing mortality during childhood and at younger adult ages-the Aids epidemic in certain African countries is an important exception. Let me add that the concentration of much of the gains at very young ages in poorer countries suggests that Becker, *et al.* (2005) estimates of the statistical value of lives for these countries is too high.

I believe that a significant part of the convergence in mortality rates occurred because of globalization. By that I mean that poorer countries 'imported' knowledge and technologies developed by the richer nations. Poorer nations gained knowledge about the importance of clean water, and cleanliness more generally, antibiotics for infectious and respiratory diseases, drugs to fight Aids developed in recent years, anti-malaria drugs, and many others. Not all poor countries benefited equally, and a good study is needed to explain why, for example, mortality declined rapidly in some Asian countries, while mortality declined much more slowly in much of sub-Sahara Africa and some other poor countries. But there can be little doubt that without the interaction with rich nations, the mortality experience of poorer nations would have been far worse during the past half century (see Papageorgiou *et al.*, 2004).

10. Conclusions

Developments during the past decade have created a vibrant field that looks at health as human capital, although several pioneering studies go back a few decades earlier. The point of departure of this field is that individuals heavily influence their mortality rates and the quality of their health, subject, of course, to their genetic make-ups, developments in the medical field, epidemics, luck, and many other considerations. This paper concentrates on mortality, although major advances have occurred in the analysis of the quality of health as well.

The major foundation of the analysis is optimal behavior by consumers where they maximize utility over time, subject to the resources they have, and to actions they can take to affect their survivor rates at different ages. Using the results of the optimization analysis, we derive optimal investments in lowering mortality. This enables one to calculate the willingness to pay for improvements in probabilities of surviving to different ages, called the statistical value of life. This value of life tends to decline with age and interest rates, rises with income, is higher when period utility functions are more concave, and depends on other variables too.

The analysis also demonstrates a series of major complementarities between improved survivorship probabilities and many other aspects of behavior. Higher survivorship at adult ages would induce greater investment in education because expected returns on education investment would be greater. Higher survivorship also induces greater investment in beneficial goods—goods that add to future utility—including beneficial habits and addictions, and discourage investments in harmful goods—goods that lower future utility—including harmful habits and addictions.

Higher survivorship leads too to greater investment in 'imagination capital' that lowers the discount on future utilities. The lower discount rates of persons with greater survivorship probabilities will lead them to save more, even with full and fair annuities, and is an additional reason why good habits and greater education are complementary with longer life expectancy.

Of considerable importance are the complementarities between the survivorship rates of different diseases and at different ages. An increase in the probability of surviving one disease raises the expected benefit from improving the probability of surviving other diseases. For example, improvements in preventing deaths from heart attacks encouraged much greater efforts to find ways to reduce deaths from strokes and various cancers. Similarly, if survival rates at older ages improve, this would raise efforts to raise survivorship at earlier ages.

These various complementarities imply that inequality in a society has several dimensions that are positively correlated with each other. In particular, persons with better survivorship probabilities at younger adult ages tend to get more education and hence have higher earnings, they would discount the future less and hence would save more, they would have more beneficial habits, and they would tend to have better survivorship rates at older ages too.

Population at different ages is an important determinant of the amount and type of investments in medical R&D by pharmaceutical companies and by government sponsored medical research. The role of population implies also that a pandemic caused say by the avian flu virus could cause the deaths of many more person (about 170 million) than the terrible flu pandemic of 1918–19 that is estimated to have killed about 50 million persons worldwide. Such a pandemic would cost worldwide about $100 trillion in terms of the statistical value of lives lost.

International comparisons of changes in world inequality over time concentrate on *per capita* GDP, although they sometimes also consider the number or fraction of persons with income below $1 or $2 dollars per day. In either case they are seriously incomplete by not considering also death rates and life expectancies. Changes in life expectancy across different countries should be added to the growth in *per capita* incomes by weighting improvements in life expectancies by the willingness to pay appropriate to a country's income level. World inequality measured by the growth in this 'full' income declined much more rapidly since 1960 than would be suggested by changes in *per capita* incomes alone.

Acknowledgements

I am indebted especially to Kevin Murphy for many discussions on human capital aspects of health. I received valuable comments at seminars at Oxford and Humboldt University. Kristopher Hult and Yong Wang provided excellent assistance.

References

Acemoglu, D. and Linn, J. (2004) Market size in innovation: theory and evidence from the pharmaceutical industry, *Quarterly Journal of Economics*, **119**, 1049–90.

Becker, G.S. (1964) *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, National Bureau of Economic Research, New York; 2nd edition, NBER, New York, 1975; 3rd edition, University of Chicago Press, Chicago, 1993.

Becker, G.S. (1967) Human capital and the personal distribution of income, University of Michigan Press, Ann Arbor. Reprinted in G.S. Becker (1993) *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 3rd ed. University of Chicago Press (for NBER), Chicago.

Becker, G.S. (2000) The new population economics I: effects of population on mortality, Summer, unpublished manuscript.

Becker, G.S. and Chiswick, B.R. (1996) Education and the distribution of earnings, *American Economic Review*, **56**, 358–69.

Becker, G.S. and Mulligan, C.B. (1997) The endogenous determination of time preference, *The Quarterly Journal of Economics*, **112**, 729–58.

Becker, G.S. and Rubinstein, Y. (2006) Fear and the response to terrorism, manuscript.

Becker, G.S., Murphy, K.M., and Philipson, T.J. (2006) Spending at old age, manuscript.

Becker, G.S., Philipson, T.J. and Soares, R.R. (2005) The quantity and quality of life and the evolution of world inequality, *American Economic Review*, **95**, 277–91.

Cerda, R. (2003) Drugs, population and market size, unpublished PhD dissertation.

Cutler, D.M. and Kadiyala, S. (2003) The return to biomedical research: treatment and behavioral effects, in K.M. Murphy and R.H. Topel (eds), *Measuring the Gains from Medical Research: An Economic Approach*, University of Chicago Press, Chicago.

DiMasi, J.A., Hansen. R.W., and Grabowski, H.G. (2003) The Price of innovation: new estimates of drug development costs, *Journal of Health Economics*, **22**, 151–85.

Dow, W.H., Philipson, T.J., and Sala-i-Martin, X. (1999) Longevity complementarities under competing risks, *The American Economic Review*, **89**, 1358–71.

Dublin, L.I. and Lotka, A.J. (1930) *The Money Value of a Man*, The Ronald Press Co., New York.

Ehrlich, I. and Becker, G.S. (1972) Market insurance, self-insurance, and self-protection, *The Journal of Political Economy*, **80**, 623–48.

Ehrlich, I. (2000) Uncertain lifetime, life protection, and the value of life saving, *Journal of Health Economics*, **19**, 341–67.

Elias, J. (2005) The effects of ability and family background on non-monetary returns to education, unpublished PhD dissertation.

Elster, J. (1997) Review of Becker (1996), *University of Chicago Law Review*, **64**, 749–64.

Fuchs, V.R. (1982) Time preference and health: an exploratory study, in V.R. Fuchs (ed.), *Economic Aspects of Health*, National Bureau of Economic Research, New York, 93–120.

Grabowski, H.G. (2005) Increasing R&D incentives for neglected diseases: lessons from the Orphan Drug Act, in K.E. Maskus and J.H. Reichman (eds), *International Public Goods, and Transfer of Technology Under a Globalized Intellectual Property Regime*, Cambridge University Press, Cambridge, 457–80.

Grossman, M. (1972) On the concept of health capital and the demand for health, *The Journal of Political Economy*, **80**, 223–55.

Grossman, M. and Kaestner, R. (1997) Effects of education on health, in J.R. Berhman and N. Stacey (eds) *The Social Benefits of Education*, University of Michigan Press, Ann Arbor, MI.

Kolata, G. (1999) *Flu: the Story of the great Influenza Pandemic of 1918 and the Search for the Virus that Caused it*, Farrar Straus and Giroux, New York.

Meltzer, D.O. (1992) Mortality decline, the demographic transition, and economic growth., dissertation, University of Chicago, December.

Mincer, J. (1958) Investment in human capital and personal income distribution, *The Journal of Political Economy*, **66**, 281–302.

Murphy, K.M. and Topel, R.H. (2003) The economic value of medical research, in K.M. Murphy and R.H. Topel (eds) *Measuring the Gains from Medical Research: An Economic Approach*, University of Chicago Press, Chicago.

Oster, E. (2006) HIV and sexual behavior change: why not Africa?, unpublished manuscript.

Papageorgiou, C., Savvides, A., and Zachariadis, M. (2004) International medical R&D spillovers, unpublished manuscript, October.

Philipson, T.J. and Posner, R.A. (2003) The long run growth of obesity as a function of technological change, *Perspectives in Biology and Medicine, Summer*, **46**, 87–108.

Philipson, T.J. and. Soares, R.R. (2002) World inequality and the rise in longevity, in B. Pleskovic and N. Stern (eds), *Annual World Bank Conference on Development Economics 2001/2002*, Oxford University Press, New York, 245–59.

Preston, S.H. (1975) The changing relation between mortality and level of economic development, *Population Studies*, **29**, 231–48.

Robbins, L.N. (1993) Vietnam veterans' rapid recovery from heroin addiction: a fluke or normal expectation, *Addiction*, **88**, 1041–54.

Rosen, S. (1988) The value of changes in life expectancy, *Journal of Risk and Uncertainty*, **1**, 285–304.

Schultz, T.W. (1962) *Investment in Human Beings*, University of Chicago Press, Chicago.

US Congress, Congressional Budget Office (CBO) (2005) *A Potential Influenza Pandemic: Possible Macroeconomic Effects and Policy Issues*, US Government Printing Office, Washington, DC.

Usher, D. (1973) An imputation of the measure of economic growth for changes in life expectancy, in M. Moss (ed.) *The Measurement of Economic and Social Performance*, Studies in Income and Wealth, Volume 38, Conference on Research in Income and Wealth, National Bureau of Economic Research and Columbia University Press, New York, 193–225.

Viscusi, W.K. and Aldy. J.E. (2003) The value of a statistical life: a critical review of market estimates throughout the world, *Journal of Risk and Uncertainty*, **27**, 5–76.