Chapter 3

<u>Demand for healthcare</u>

3.1 Introduction

As we have already seen in chapter 1, to most of us healthcare means visiting our family doctor and taking medications, going for medical diagnostic tests like blood-work or a magnetic resonance imaging (MRI) session, and having to check into the hospital for a minor procedure or a serious operation. These components of healthcare require the efforts of doctors, nurses, various technologists and all other inputs required to operate the family doctor's practice, the hospital and the diagnostic clinic. Briefly, healthcare includes services supplied by the medical profession or, in other words, the medical care. However, healthcare also includes health-enhancing activities, from exercise and vitamin intake to good sleep to eating well. These self-initiated activities may also need market services such as a gym and goods such as a good bed and healthy food. Thus, as healthcare requires the purchase of various goods and services, economic analysis classifies the purchasing need as demand for healthcare.

The demand for healthcare does not originate from primitive preferences but, rather, it is a demand derived from the more primitive demand for health. However, it also differs from most inputs in production where the output is also a flow. Individuals directly demand health[1], a **stock** variable or the level of one's health at a given moment in time, whereas the demand for healthcare is a **flow** or a certain amount of healthcare over a given time period. The healthcare demand is rather similar to a worker's demand for human capital where training, education and on-the-job learning are all flow inputs combined with one's time and effort to produce human capital. Similarly, human capital enhances the individual's earning potential by boosting one's wages or salaries whereas the health stock increases one's healthy time available for work and leisure. Taken in the long run context, sustained periods of health positively affect the individual's earnings both in terms of wages and his ability to work. Where the health stock and human capital differ is the direct demand for health stock. While human capital may not be a prerequisite for leisure activities, health stock necessarily is.

Since being healthy is a desired state by individuals under all circumstances, work or leisure, such a desire generates the first reason for the demand for health stock. The second reason is that, normally, individuals have to work for a living and work is better performed if the individual in question is healthy. Therefore, the first reason is health stock as consumption good whereas the second as an investment good[2].

Healthcare, as a produced good, exhibits the following properties. First, as discussed above, healthcare demand is a derived demand, i.e. health is demanded and healthcare is

---

[1] Or the flow of daily good health as is modeled in Grossman [1972, 2000].
[2] See Grossman [2000] for a technical analysis of these two distinct cases.

demanded because it produces health. Second, health is produced by using various inputs, one of them being healthcare in the larger sense and medical care in the narrow. Finally, the replenishment of the health stock introduces a dynamic relationship between the health stock and healthcare. The combination of healthcare and other health inputs produces the health investment. The individual's health profile over time can then be represented as a **stock adjustment** model where the stock of health varies for the better if the individual positively contributes to his health over a given period whereas reckless behaviour lowers the stock. This relationship between one's stock of health and the flow of health investment yields a simplified version of the health stock model of healthcare demand originally developed by Grossman [1972]. A graphic summary of the model is given in figure 3.1 below.

Individuals consume various goods and services by purchasing them and allocating their valuable time to consume them. Going to the movies as well as jogging involve substantial amounts of leisure time as well as purchased inputs like movie tickets, transport, and running shoes. The model lumps such goods and services into home goods represented by B and health investment goods by I. B is consumed by combining one's time T_B with the purchased inputs X and I by combining time T_I with inputs M. The consumption of goods exhibits properties of production functions in that the time and purchased inputs are combined to yield the consumption. The two production functions in the simple model are $B(X,T_B;E)$ and $I(M,T_I;E)$ where E denotes environmental variables, such as noise and pollution that would, respectively, spoil the production of home goods B and health investment I. At the centre of figure 3.1 lies the link between health investments and the state of one's health. The genetically programmed erosion of human health over time is represented by δH, i.e. humans lose a varying fraction of their health stock H over a given period of time. However, health investment I contributes to the stock. Therefore, the sum I - δH yields the rate of change of the health stock. The health stock is not only good in itself. The healthier the person, the more healthy time is available either for work or for leisurely activities B and I, the latter being the critical contributor to health stock. T_H can thus be split between work time T_W and the time allocated to the consumption of home goods and health investment, respectively T_B and T_I. We note that T_W generates the income used to purchase inputs X and M. Finally, since individuals value their consumption of goods and services B as well as their state of health H, they allocate their resources between B and I, towards B because it yields utility, towards I because it contributes to the health stock. Now that the Grossman model has been intuitively introduced, the purchased inputs M can be interpreted loosely as healthcare.
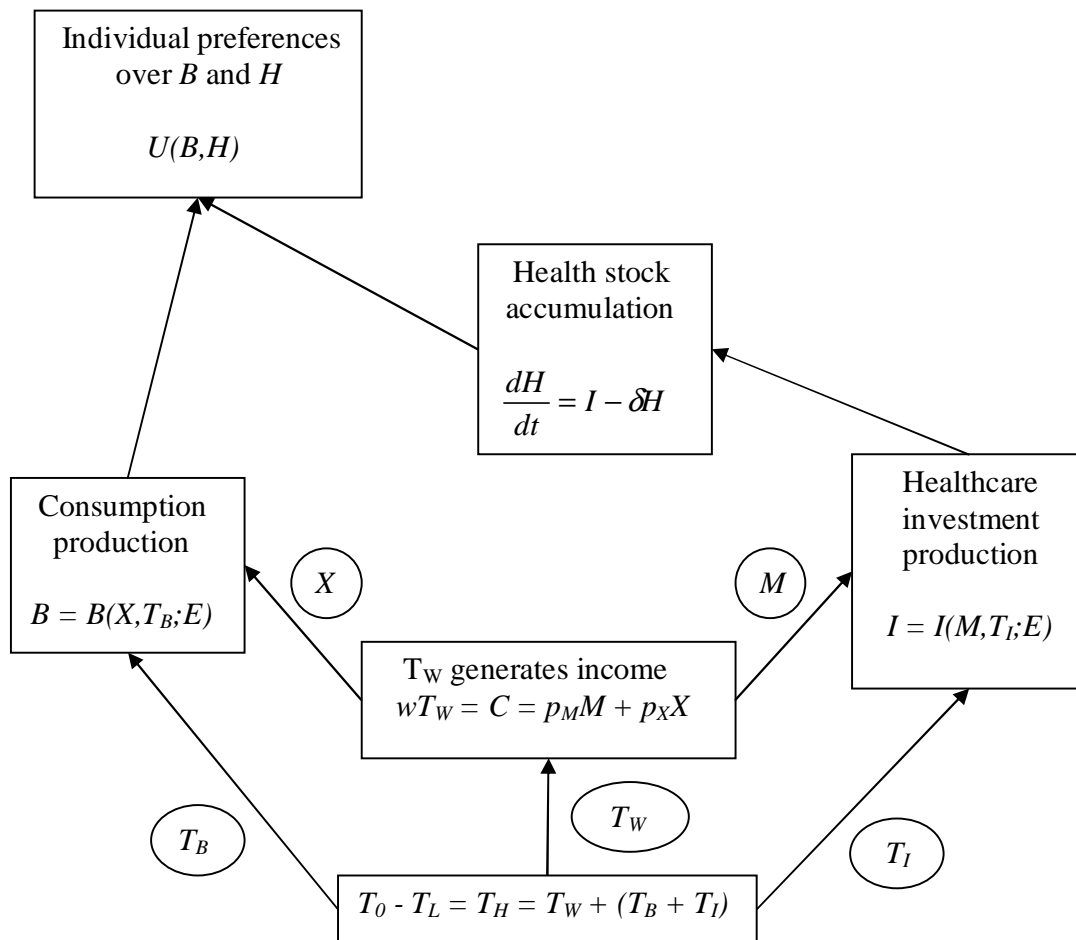
Figure 3.1 Health stock model of healthcare demand schematically summarized

However, an important distinction must be drawn between medical care and the more comprehensive concept of healthcare. Often used interchangeably, both are gross investments into one's health. Medical care is the collection of health-restoring, health-preserving and health-enhancing services provided by applied medicine. As such, medical care consists of the available medical technology, running typically from symptoms to diagnosis to treatments, but also including preventive technologies. Thus it can be preventive[3] or curative. Healthcare, however, beyond medical care involve layers of individual choices over work, consumption and leisure. For example, choices of workplace, vocation, work tempo, consumption of healthy food and the allocation of

---

[3] Contrary to popular belief prevention is not uniformly less costly and less invasive than treatment (Laupacis [1996], Marshall [1996]).

leisure time to health investment all fall into the healthcare category without being medical interventions. Moreover, these choices tend to be overwhelmingly preventive rather than the mostly curative modern medicine. Thus, whereas all choices enhancing one's health constitute healthcare, a subset of services mostly provided by medicine constitute the medical care.

The second section of this chapter will progressively develop the health stock model of demand for healthcare. As summarized in figure 3.1, the model internalizes the ability of individuals to choose their health profile as well as the inherent dynamics of one's health stock. The section will thus examine individuals' preferences for health as well as their allocations of time and money towards healthy activities and medical care in order to derive their demand for healthcare and medical care. The third section considers examples. The effects of non-monetary and monetary factors on the demand for healthcare will be examined. For instance, the response of healthcare demand to changing preferences and rising wages will be analyzed. The fourth section traces the effects of user fees, a demand management tool. The conclusions section reemphasizes the fundamentals covered in the chapter and provides links to the demand for healthcare insurance covered in chapter four. .

3.2 The health stock model of healthcare

Individual's preferences

When individuals enjoy various goods and services, this enjoyment is normally translated into a demand to purchase and consume. However, this enjoyment is stronger, the better the individual's health. Technically speaking, demands for goods and services are health-state dependent. Thus, there exists some complementarity between an individual's state of health and her consumption of goods and services through this state-dependency. This complementarity relationship does strongly suggest that health, in itself, is desirable and, hence, individuals would be prepared to allocate resources to enhance health.

Yet, there definitely exists some substitutability between health and consumption through tradeoffs between health-enhancing goods vs. the rest. For instance, over-exertion and stress in pursuit of higher income frequently appear at the expense of health or, simply, as lower levels of health investment. This substitutability may involve both dimensions of health, as purely a consumption good as well as investment into income-generating health capital. For expositional purposes, we will henceforth refer to health as a consumption good entering an individual's utility function alongside other goods and services without conditioning individuals' utility functions by health status.

Thus, an individual's utility function U(B,H) is defined over ordinary goods and services B (henceforth home-goods, consistent with Grossman [1972] terminology) and health H^4, with utility increasing in both B and H and the utility function yielding convex towards

---

[4] We'll simplify the Grossman [1972] notation that enters health stock services, rather than the health stock itself, into the utility function.

the origin indifference curves representing diminishing marginal rate of substitution between B and H. Referring back to figure 3.1, this tradeoff represents individuals' preferences both current and intertemporal. Currently, enjoying a certain level of income, individuals can choose to marginally sacrifice health investment goods (e.g. less sleep causing short-term drop in alertness) for an increase in home-goods. However, this is not the full opportunity cost of the increase in home-goods because current lower investment in health would induce a long-term drop in the health stock with the reduced healthy time consequence. The present discounted value of the reduced healthy time decrease in the future combined with the current loss of health constitute the opportunity cost of an increase in home-goods consumption. Since individuals thus decide over intertemporal allocations, the modeling must intrinsically be intertemporal.

The dynamics (or the time profile) of the health stock requires that the single-period individual utility function U(B,H) be modified so as to reflect individuals' intertemporal tradeoffs and their discounting of future utilities. Two other intertemporal channels, in the general optimization problem, beyond individuals' valuation of the future are the depreciation of the health stock and the possibility of countering such ultimately inevitable depreciation through health investments. Health as stock can be accumulated or rather decumulated over time and health as consumption good[5] can be consumed at different points in time; individuals characteristically take account of their future health for both these reasons. Health as consumption good then necessitates all future consumptions be taken into account and health as investment good generates healthy time required for work and hence income. The individual's lifetime utility can then be modeled as the present discounted value of future utilities[6] (as the continuous version of Ried [1998])

$$\int_0^T e^{-\theta t} U(B,H) dt \tag{3.1}$$

where t is the instantaneous time unit (or the moment in time), θ is the time discount rate, $e^{-\theta t}$ the discount factor (or, simply, the individual's subjective weight attached to every moment in the future) and T the individual's residual life expectancy. The utility function U(B,H) in equation 3.1 must be interpreted as the instantaneous utility of the individual and the whole expression then is the weighted sum of utilities over the residual lifetime. The weights decline over time, signifying that today's enjoyment is more valuable than in a future period. Hence, the individual's choice of current levels of consumption and health investment are, therefore, not independent of their future expected values. For instance, a lower health investment today may increase current consumption without lowering current health but its opportunity cost is a fall in the future stream of health stock in turn lowering not only the future consumption but also the individual's future earning capacity and hence his future consumption. These tradeoffs are moderated by the discount weights $e^{-\theta t}$ that assign higher utility weights to today's health and consumption. Thus, the maximand in equation 3.1 is fairly straightforward except for the time horizon T. There are two issues regarding T. The first is whether there is an optimal length of

---

[5] The health as consumption good in Grossman [1972] is modeled as flow of services from stock.
[6] Grossman [1972, 2000] uses a discrete time framework for the individual lifetime problem.

horizon. For example, euthanasia is an endogenous choice of end-of-life whereas a terminal illness is a relatively randomly-timed exogenous end to life. Most individuals choose their health investments considering a normal or average residual life expectancy. This optimality question links the time profile of the depreciation rate to the individual's willingness to invest in health so as to aim at a health stock perhaps well above the survival minimum beyond a certain advanced age. Alternatively, de-investment in the form of harmful addictions or negligence today boost the health depreciation rate (Becker & Murphy [1988]). The second issue related to residual life expectancy T involves a modeling technicality. Whereas the time profile of the depreciation rate is not deterministic in so far as we scientifically know, the question remains as to whether it should be analyzed as such as. A random evolution of the depreciation rate would add considerable modeling complexity yet, with advances in genetics, some of sources of randomness are becoming predictable. We will return to the discussion of the depreciation rate below in the section entitled Health stock profile and health investment.

Allocation of time and income

Every individual is endowed with the same amount of time T_0 regardless of the unit of time chosen for analysis. Though, for the sake of realism, T_0 must be long enough to allow days of morbidity as, typically, days can be characterized as healthy or unhealthy. A longer period chosen would then yield meaningful periods of illness versus wellbeing. The total time endowment will now be broken down into components as

$$T_0 = T_L + T_W + (T_B + T_I) \qquad (3.2)$$

where T_L is the ill days time, T_W the work time, T_B the time allocated to produce the home good B and the time T_I allocated to produce the health investment good I. The home good B can be produced by a combination of T_B and of market-purchased inputs X thus obeying the production function

$$B = B(X, T_B; E). \qquad (3.3)$$

Similarly, for health investment I, the production function is given as

$$I = I(M, T_I; E) \qquad (3.4)$$

where M is the set of market-purchased inputs, including medical care. In both equations 3.3 and 3.4, E represents the exogenous factors such as the individual's genetic inheritance and environmental context. As such, changes in any component of E will induce shifts in the production function. For example, for the same input levels, the quantity of health investment produced will be lower in an unhealthy environment, e.g. a polluted environment, than in a healthy one. Both production functions should normally exhibit decreasing returns to scale. Although easily understandable for home-goods, decreasing returns to scale in the production of health investment must be related to the human body's capacity to absorb one's health-enhancing activities. For instance, while

6

moderate to medium levels of exercise go a long way, medium to high levels of exercise do not enhance the health stock as much. The convex-towards-the-origin shape of the production possibility frontier (PPF) in figure 4(b) easily follows from the decreasing returns to scale property of the production functions. The shape of the PPF corresponds to the increasing opportunity cost for either the home-goods or the health-investment.

Individuals face time and income constraints. The time constraint was introduced above in equation 3.2 whereas the income constraint is given by

$$wT_W = p_M M + p_X X \tag{3.5}$$

where w, p_M and p_X are, respectively, the wage rate, the price of medical care (or health-enhancing goods at large) and the price of other goods and services. The constraint means that the individual cannot spend more than her earned income wT_W on purchased inputs M and X. The two constraints are not independent. First, time and income allocated to healthy activities may reduce time lost to illness and boost time available for work. Moreover, in the longer term, it may enhance productivity through improvements in health stock and thus increase wages. Finally, while it will be ignored here, the intertemporal planning of time and income allocations allows further interactions between the two constraints over time.

Individuals strive to allocate their income and time resources efficiently between the production of these health-enhancing and other goods as presented in figure 3.2 below. The rectangular box, representing the time and money resources given a distribution of one's time between work (hence money income) and leisure, is called the Edgeworth box. This box shows the feasible distribution of resources towards the production of B and I. Times allocated to production plus and time allocated to work for income add up to T_H, healthy time available. And the time allocated to work yields the total income available to purchase the two inputs M and X anywhere in the box. Moreover, to simplify the exposition, the input prices are normalized to unity (i.e. $p_M = p_X = 1$) by adjusting the units of M and X. This simplification allows us to identify every quantity M or X with the expenditure on that input. Furthermore, whereas the horizontal dimension of the Edgeworth box being the total leisure time $T_{LE} = T_B + T_I$, the vertical dimension is equal to the total expenditure on those inputs. Concisely, every point in the box is an allocation of time and money. Of course, one the ultimate objective being efficient production levels for B and I, inefficient input allocations must be ruled out.

One such inefficient allocation is the bundle (M^0, T_I^0) in figure 3.2 below. Given this bundle, the individual can produce (B^0, I^0) as given by the isoquants passing through the bundle. However, since marginal rates of substitution at the bundle are not equal, a reallocation of more time but less money to B production along the broken arrow induces higher levels of both B and I. In fact, the bundle (M^*, T_I^*) allows the production of $I^{LOW} > I^0$ whereas at the same bundle (X^*, T_B^*) allows $B^{HIGH} > B^0$. The initial bundle as well as the final bundle are feasible and use all the available inputs, except that the $\{(M^*, T_I^*), (X^*, T_B^*)\}$ bundle allows a higher production level of both goods. We note that, at this latter bundle, the isoquants B^{HIGH} and I^{LOW} are tangent, i.e. the marginal rates of

technical substitution are equal. In other words, any departure from this bundle will lead to a decrease in one output or the other along the diagonal curve (or possibly both if one moves off the curve) whereas from $\{(M^0,T_I^0),(X^0,T_B^0)\}$ to $\{(M^*,T_I^*),(X^*,T_B^*)\}$ both outputs increased. We conclude that $\{(M^*,T_I^*),(X^*,T_B^*)\}$ is an efficient bundle. In fact, all bundles along the diagonal curve are efficient. This diagonal curve in the Edgeworth box is then called the **contract curve** and it consists of those input bundles that are efficient.
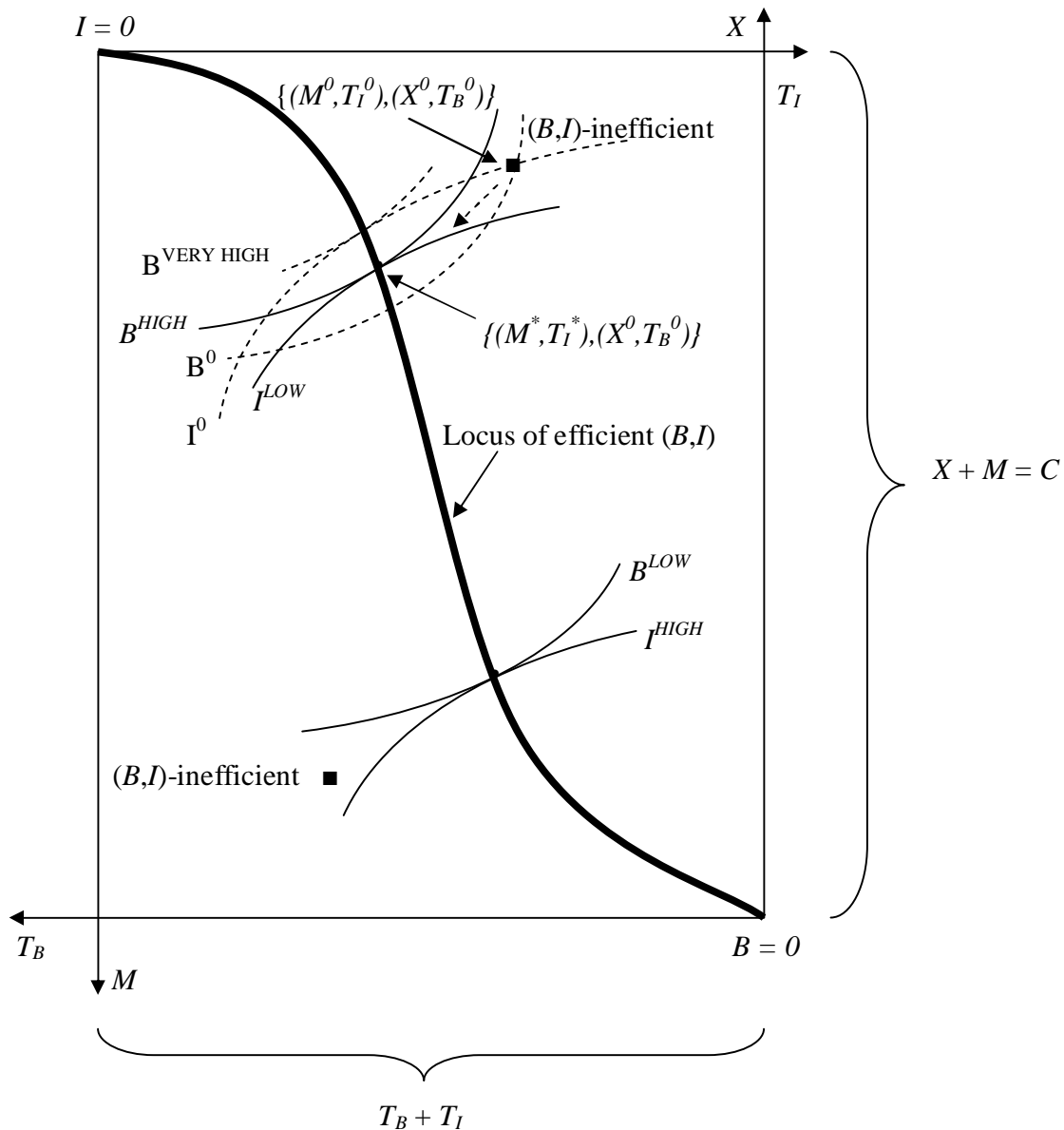

Figure 3.2   Efficient choices of B-I production

Intuitively, at any allocation off the contract curve, such as $\{(M^0, T_I^0), (X^0, T_B^0)\}$, a reallocation of time from the production of, say, B to I requires a smaller increase in X to preserve the B production than the release of M to preserve the I production. Thence, either B or I can be increased without the other being decreased or, as we saw above, both can be increased along the broken arrow. This move is a **Pareto improvement**, i.e. more of each output can be produced with the same total resources but by redistributing them. Thus, the individual would never choose a bundle off the contract curve. Note that movements along the contract curve aren't Pareto-improving because an increase in one comes at the expense of decreasing the other. As such, their relative evaluation depends on the individual's preferences over (B,I) bundles.

An example can illustrate the efficiency along the contract curve. Suppose I is an individual's fitness activity in a gym and B a home entertainment activity like hosting friends. Let the initial bundle off the contract curve be $\{(M^0, T_I^0), (X^0, T_B^0)\}$. Given (M^0, T_I^0) this individual reaches $I^0 = I(M^0, T_I^0)$, a sufficient level of fitness. We note that if she spends less time in the gym by reducing T_I while purchasing the services of a personal fitness advisor by increasing M, she can preserve the same fitness level if the substitution time for money towards B is realized along the isoquant I^0. The individual then reaches the bundle where the two isoquants $B^{VERY\ HIGH}$ and I^0 are tangent along the contract curve. The significant increase in B derives from input substitutions. Along I^0, little increase in money is required to compensate decreases in time because the individual is already spending too much time in the gym. Yet, the production of B steeply increases with time rather than money because, at (X^0, T_B^0), an extra unit of time is very valuable. Overall then, the individual realizes that she is spending too much time rather than guidance in the gym and rectifies the situation by releasing time towards the home-good that dearly requires more time. The final bundle yields as much fitness I^0 as before but much more in B at $B^{VERY\ HIGH}$, surely a Pareto improvement.

A remark concerning the size of the Edgeworth box is now in order. Since individuals are endowed with $T_H = (T_0 - T_L)$ of healthy time, they can allocate it between work time T_W and leisure time $T_{LE} = T_B + T_I$. This time allocation problem will be further studied in the following section as well as in Appendix A as a movement along the contract curve may have second-round effects. A change in preferences inducing, say, an increase in I, will induce not only a first-round change in the bundle (B,I) but, also, via the effect of I on the health stock H and subsequently on T_H. And, of course, if the health stock increases, so do the healthy time and the dimensions of the Edgeworth box.

Health stock profile and health investment

Humans, like all living beings, have a generic lifecycle. In the absence of major illnesses and medical interventions, we are born, get healthier and stronger to our mid-twenties and then decline towards the natural death through a net loss of various cells in our bodies. Where healthcare essentially intervenes is when illness sets in or when our health stock can be improved with the consequential retardation of cell death. In technical terms, human health stock typically accumulates at first and decumulates later, with the eventual demise of the body. This health stock profile is neither completely exogenous, nor

completely endogenous. The function of health investment as introduced above is to boost this profile and, assumed away in the model under consideration, is the effect of the individual's activities on the depreciation rate.[7] This latter effect can lower the rate or retard its inevitable increase.

In order to understand the basics of healthcare demand, we would be justified in significantly simplifying the framework used. Thus, the health stock profile can be represented by the following equation of motion of the health stock H over time

$$\frac{dH}{dt} = I - \delta H.$$ (3.5)

Equation 3.5 yields the time rate of change of health stock or the net investment as the difference between the gross investment in health stock, I, and the total stock depreciation δH where δ, a fraction, denotes the current rate of stock depreciation. The rate of depreciation of the health stock is not time independent, being typically low early in life and getting larger later. Three observations on δ are in order. First, biological reality makes it so that, towards the end of life, no matter how high investments may have been, the health stock will eventually fall below a threshold level signifying death. Thus, when the gross investment I required for preventing negative net investment becomes prohibitively costly or simply impossible, the individual starts approaching the end of life. Second, though the depreciation rate δ is fundamentally determined by genetic makeup, it will also be influenced by cumulative health investments and the environment in which the individual in question lives. Finally, δ may even be negative early in life as the person becomes stronger, acquires immunities and, significantly, learns to produce health investment more efficiently.

The solution to equation 3.5, as part of the individual's overall problem set up below in equation 3.7 and explained in Appendix B, yields the health stock profile. This trajectory starts at birth with an initial health stock for a particular individual, typically rises over time towards mid-twenties, more or less follows a plateau till late middle ages, then starts a steady decline and, finally, takes a tumble late in life. Of course, various relatively random events, from catastrophic illnesses and accidents, may cut one's life short. Throughout individual's life, the health stock is the major determinant of morbidity as represented by ill time T_L and, hence, time available for work T_W. A higher health stock translates into a longer healthy time T_H through a healthy time production function $T_H = T(H)$. Not only is there an upper limit to one's time available but also health stock improvements exhibit diminishing returns in that increases in the health stock reduce morbidity slower or, on other words, increase healthy time slower. This relationship is illustrated in figure 3.3 below.
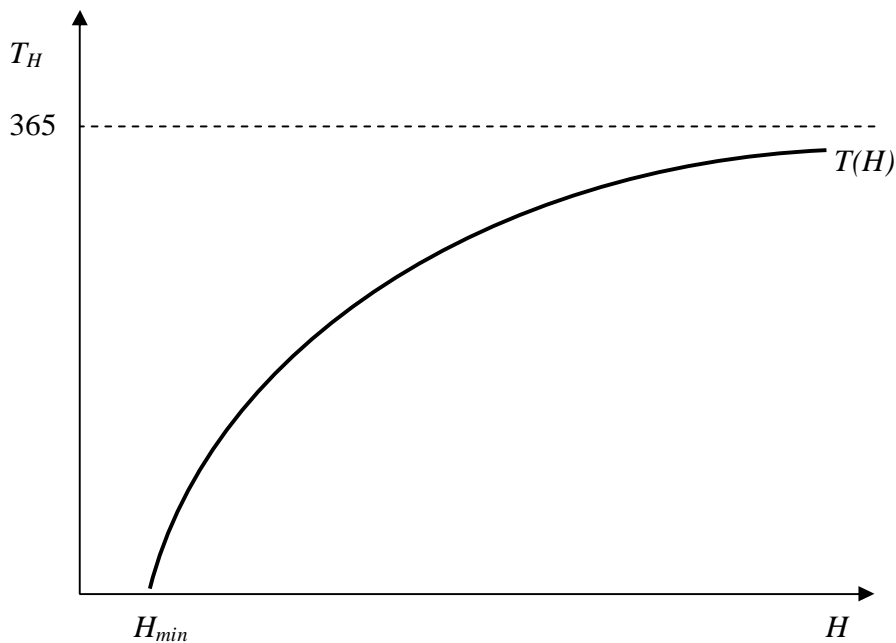
_____

[7] Chang [1996].

Figure 3.3 Healthy days production function

Figure 3.3, by linking the health stock and healthy time, does in fact link the income and time constraints because a higher health stock enables one to earn more and relax the income constraint. Moreover, T_{min} in the diagram is the end of life threshold although one's health could deteriorate as a result of catastrophic illness or accident and not just old age.

Now, in order to concentrate on the fundamentals of individuals' decisions concerning health, we will restrict our attention to snapshots of reality called steady states where a given health stock prevails. This simplifies the exposition and allows the use of simple diagrams. Every snapshot corresponds to a particular steady state where, given a steady rate of depreciation, health stock is momentarily constant, that is $dH/dt = I - \delta H = 0$ or, simply, $I = \delta H$. This is a reasonable approximation to reality in that, barring random occurrences of serious illness, most individuals' working lives are characterized by steady yet diverse levels of health. As health investment is kept equal to total health depreciation, the derivation of the demand for healthcare can be graphically derived. (Appendix B presents the simplified but still dynamic health stock model.)

Since health investment equals the health stock replacement δH, the health stock variable H can simply be replaced by I/δ in the individual maximand U(B,H). As a result, any exogenous changes, from prices to tastes, will simply induce changes in health investment I and H instantaneously to a new level instead of H adjusting over time. Since the intertemporality of the problem is thus suppressed, the individual's current preferences can then be represented by

11

$$V(B,I) = U(B, I/\delta).\tag{3.6}$$

The individual enjoys increasing utilities from higher levels of home-goods and health investment given that health investment boosts health stock, this latter being the ultimately desired consumption and investment good for the individual. The indifference curves generated by V(B,I) are illustrated below in figures 3.4(b) and 3.5(b).

This simplification is, just to reiterate, used for expositional purposes only. However, the constancy of the health stock in the absence of major illnesses is not far from reality. As exemplified by the difficulties of lowering one's weight or cholesterol levels over short periods of time, this constancy constitutes a fairly realistic approximation over the short term. Moreover, high levels of investment may prove infeasible simply due to the time constraint and the limitation induced by income-earning potential. On the other hand, de-investment in health is clearly feasible and self-destruction can bring the health stock down fairly easily not only in individual cases but also at community level as recently exemplified in 1990s' turmoil in Russia. Within a few years, Russian life expectancy fell by about six years due to deterioration in diet and obesity, to increased addiction to alcohol and harmful substances and stress as well as a lack of medical care (Brainerd [2005]). What happened in Russia can be easily put into context with the help of equation 3.5 above. Falling incomes did not allow Russians to invest in health as current consumption needs took precedence. Moreover, the health depreciation rate increased, even in such a short term, as a result of dreadful social and economic circumstances. Thus, both I falling and δ rising, the health stock took a tumble, lowering life expectancy drastically.

Before we look at the full problem of the individual it is worth noting that another simplification is the dropping of the possibility of intertemporal income transfers keeping with our earlier simplification of the dynamics of the problem. Individuals can and do, in general, save and borrow with an eye to the ultimate evening out of their marginal utilities over time (see Grossman [1972, 2000], Ried [1998]). Just like in simple consumption problems where the consumer reallocates his marginal dollar from a lower utility generating consumption to a higher one, health care would require reallocation of that dollar from a healthy state where the return is low to an unhealthy state where the extra dollar would generate a high utility return by enabling purchase of healthcare. Similarly, one might want to transfer the extra dollar from his youth when illness is unlikely to occur to old age when it is more likely. This intertemporal behaviour corresponds to a lifetime utility maximization over various goods and services. The simple and most pervasive examples being a house mortgage, retirement savings and health insurance, though this latter is not only against old age ailments but also for income-smoothing against lumpy and unexpected healthcare expenses for random and serious illnesses.

Given the steady-state assumption simplifying equation 3.5 and hence the individual's maximand as in equation 3.6, the individual's full but now static problem of utility maximization can be written as follows

$$\max_{\{X,M,T_B,T_I\}} V(B,I)$$

$$\text{s.t.} \quad B = B(X,T_B;E) \tag{3.7}$$
$$I = I(M,T_I;E)$$
$$T_0 = T_L + T_W + T_{LE}$$
$$wT_W = p_M M + p_X X$$
$$T_{LE} = T_B + T_I.$$

The first two constraints are the production functions for home-goods and health investment. The inputs for the production functions, i.e. time and purchased goods or money, have to satisfy the availability constraints, the first being the overall time constraint and the second the budget constraint. Finally, the last constraint limits the time available for the production of B and I to the leisure time, after subtracting the ill and work times from total time available to the individual. The variable E in the production functions acts as a shift variable, depending on given genetic and environmental backgrounds. The problem is illustrated in figure 3.4 below.

Two remarks will relate this simplified version of the model to the original. First, as health stock is kept constant, time ill T_L is also constant. As a result, healthy time T_H is given. When the individual allocates it between work and other activities, it suffices to choose $T_B + T_I$ in order to determine work time or vice versa.

Second, complementing the first remark, the opportunity cost of time spent in the production of B or I must be equal to the wage rate. In figure 3.4(a), the slopes of isoquants are equal at $\{(M^*,T_I^*),(X^*,T_B^*)\}$ therefore the last hour spent in either activity is worth an equal value. If, however, the opportunity cost of time spent in B (or, equally, in I) in terms of purchased goods exceeded the wage rate, the individual would work less and consume more. From figure 3.4(a), the marginal rates of substitution in production are equalized and the value of marginal time spent in leisure activities is equal to consumption goods sacrificed at the margin. This marginal willingness to pay is then equal to the wage rate, as seen in figure 3.4(c) where the indifference curves shown are induced preferences, i.e. they are derived from V(B,I). Since individuals choose efficient combinations (B,I), any bundle chosen necessitates a corresponding bundle of leisure time and income. A corresponding apportionment of the available healthy time would yield this latter bundle of leisure time and income.

Figure 3.4(a) below illustrates the efficient production choices made by the individual. Given the availability of time $T_{LE}^* = T_B^* + T_I^*$ (from panel (c)) the individual splits it between the productions of B and I (from panels (a) and (b)), combining the time inputs with the allocation of available income wT_W^* to purchase, respectively, the inputs X^* and M* into the two lines of production. Panel (b) exhibits the production possibilities boundary (PPF) derived from the Edgeworth box in panel (a). Given decreasing returns to scale technologies in the productions of B and I, the PPF is concave towards the origin. The individual then picks the utility-maximizing bundle (B^*,I^*) along the PPF. This choice simultaneously determines first the health stock H^* and, secondly, the allocation of
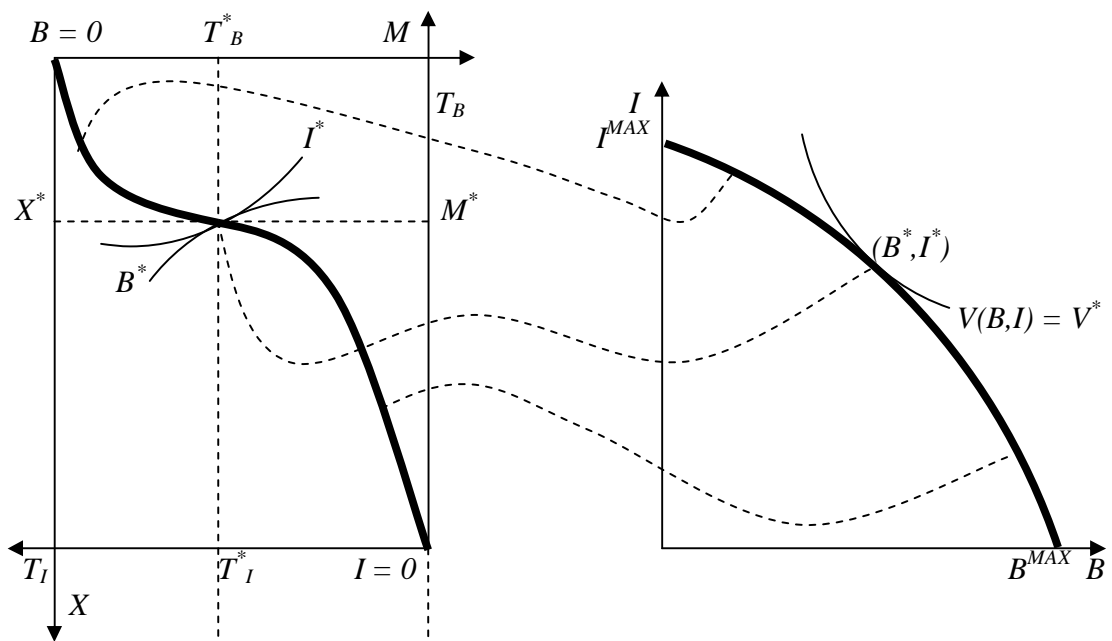
(a) Efficient allocation of time and inputs

(b) Utility maximized over individual PPF
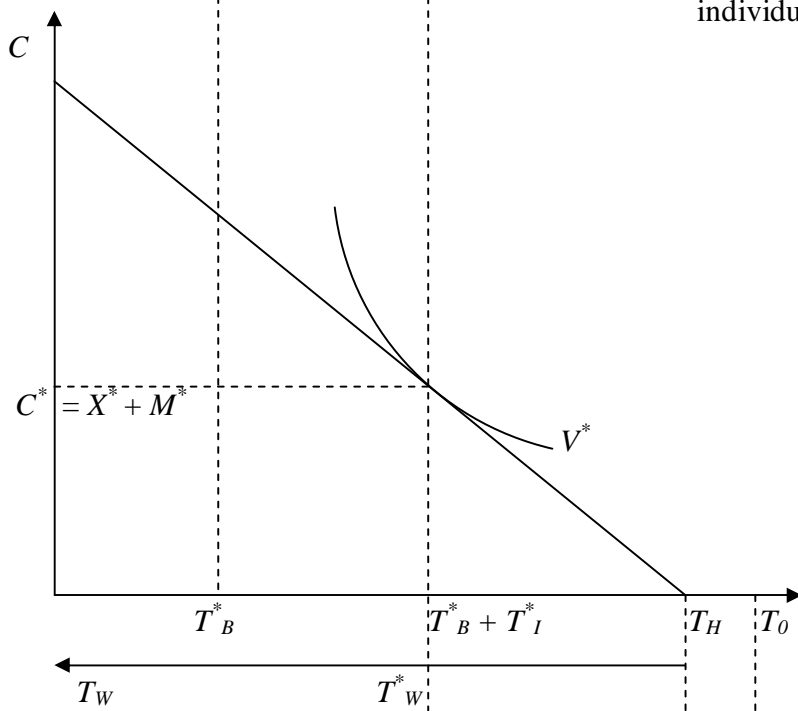
(c) Leisure-work allocation of time

**Figure 3.4   Individual demand M^* for healthcare**

14

one's time between work T_W^* and leisure T_{LE}^* as well as the purchased inputs M^* and X^*. We note that, given the parameters of the model, M^* is the quantity demanded of healthcare (medical care) if M is interpreted to be healthcare (medical care).

A simple comparative statics exercise will be used to illustrate (see figure 3.5 below) the functioning of choices in response to changes in the individual's environment. What happens when, say, preferences change? Since preferences are exogenously specified, this initial change will first trigger a complex process of reallocations. Time endowment is the individual's fundamental resource in that time is a primary input into the production of B and I as well as being the marketable labour resource that enables one to work and purchase X and M.

As shown in panel (b), the individual's marginal rate of substitution in consumption (i.e. her marginal willingness to pay for B in terms of I) has increased. This induces the lower level of health investment I^2 and hence the lower health stock H^2 (not shown on the diagram). However, from healthy days production function in figure 3.3, T_H falls from T_H^1 to T_H^2 thus inducing the smaller Edgeworth box as in panel (a). We note that this Edgeworth box is smaller both in terms of time and expenditures as the individual prefers to reduce leisure time considerably from T_{LE}^1 to T_{LE}^2 as well as work time slightly from $T_W^1 = T_H^1 - T_{LE}^1$ to $T_W^2 = T_H^2 - T_{LE}^2$ rather than, say, reducing work considerably and leisure slightly in panel (c). Returning to panel (a), we also note that the individual's final preferences reveal that she is indifferent between the original bundle (B^1, I^1) and the final bundle (B^2, I^2). However, if the PPF had not shrunk, the individual could have been better off at his new bundle *(B^2, I^2)* with his new indifference curves than with the old bundle *(B^1, I^1)*. This process of adjustment will later be used (in section 3.3 below) to analyze exogenous (e.g. pollution) and endogenous (e.g. addiction) harmful consumption activities.

Here is an example to gain further insight into this adjustment process. Considering our mundane attitudes to emerging health information, our recent experience with the Atkins diet may be easily understood to illustrate the Grossman model adjustments. The low-carbohydrate Atkins diet, upon proving as an effective short-term weight loss tool, became a fad in the early part of this decade but then fizzled away as serious health concerns with the diet ensued. Of course, when the diet became popular, consumers were not yet aware of the negative health consequences. The Atkins Company went bankrupt in 2005. At the diet's heyday, food manufacturers flooded the markets with low-carbohydrate food items upon consumers' increasing demand following their change of preferences relative to weight-loss products and services. If we maintain the current wisdom that the Atkins diet may, overall, be a consumption good B rather than a health investment good I, the consumers' initial shift to the Atkins diet may be seen as an increase in their marginal willingness to pay for the diet in terms of health investment goods forgone as in figure 3.5(b). However, the negative effects of the diet would have shrunk the PPF thus eventually yielding a level of utility no higher than the original.

This section introduced the concept of health care demanded as a demand for input, an input into the production of desirable health investment goods. In fact, individuals
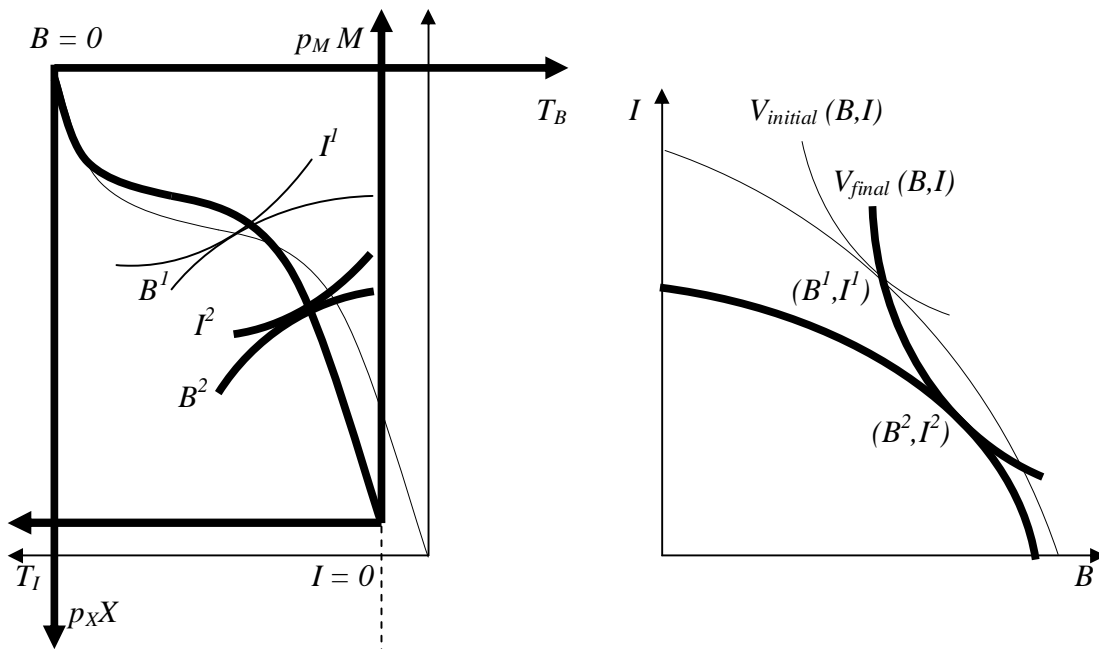
(a) Efficient reallocation of time and inputs

(b) Utility maximized over
smaller individual PPF
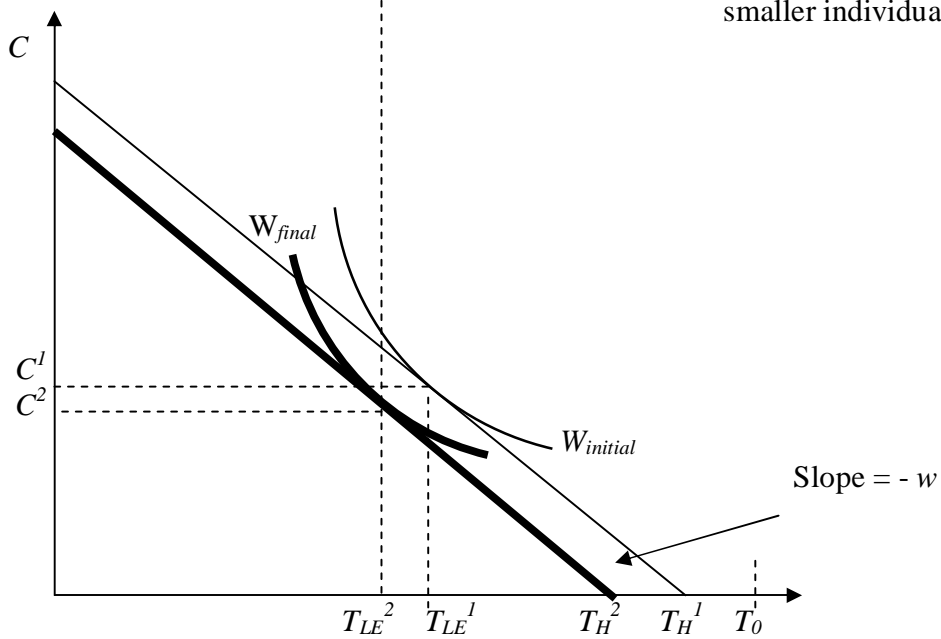
(c) Leisure-work reallocation of time

Figure 3.5 Changes in preferences

16

demand consumption goods and health, this latter also being demanded as it enables one to earn income. We can therefore derive the healthcare demand function from this same framework by varying the price of healthcare.

Demand for healthcare

Demand for healthcare in the health stock model is demand for purchased inputs into healthy activities, ranging from medical care to jogging.[8] We will now derive the demand curve for healthcare by simply defining M as healthcare[9]. A fall in the price p_M of healthcare, from p_M^0 to p_M^1, can be thought of triggering three rounds of induced changes. Firstly, holding all other variables constant, the fall in own price of M induces the increase in demand from M^0 to M^{int} in figure 3.6 below. In panel (a), neither the leisure time T_{LE}^0 nor the resources allocated to the production of B have changed thus (M^{int}, T_I^{int}) is not a feasible choice. Thus, secondly, the extra income and time to be expended on producing I have to come at least partially from reductions in X and T_B, i.e. at the expense of B. Thirdly, however, since the relative price of M with respect to the cost of leisure time is lower, there will be a substitution of T_I for M hence reducing the induced demand pressure on leisure time. This modified work leisure choice could thus go either way, i.e. more or less work, because although leisure time has just become relatively more expensive (compared to M), the increasing production of I demands more of time input T_I despite being substituted by M.

Thus, the B-I frontier in figure 3.5(b) shifts upward (except at B^{MAX}) for a higher level of I at the expense of less B all the while increasing the individual's utility. Consequently, induced by the lower relative price of M, the demand for M will increase, though by less than the initial M^{int} and settle at M^1.

Perhaps an anecdotic example would help. Suppose the initial impetus comes from cheaper access to fitness clubs. This certainly lowers the price of M and tends to boost the demand for M towards M^{int}. However, the generic individual needs more time to combine a higher usage of M with T_M. Since healthy time T_H is fixed in the short run, the extra time must cut into T_B or T_W. Either case, the demand for M is moderated by limited time resources or, in other words, by having to compete with other uses of healthy time available. The demand for healthcare thus increases from M^0 to M^1, a lower level than M^{int}.

We note that this demand curve is a short-run phenomenon. For instance, following the price drop, healthcare demand increases. Since healthcare is health-improving, the health stock and, consequently, healthy time will increase as in figure 3.7 below, partially

---

[8] Feldman & Dowd [1993], Rice [1993].
[9] We are still abstracting from health insurance. As Dusansky & Koc [2006] demonstrate, the presence of health insurance renders even the free-at-the-point-of-service medical care sensitive to price hence a negatively-sloped demand curve. Since insurance premia must reflect the cost of insurance coverage, individuals respond by adjusting their coverage by purchasing more or less depending on, respectively, whether premia fall or rise.
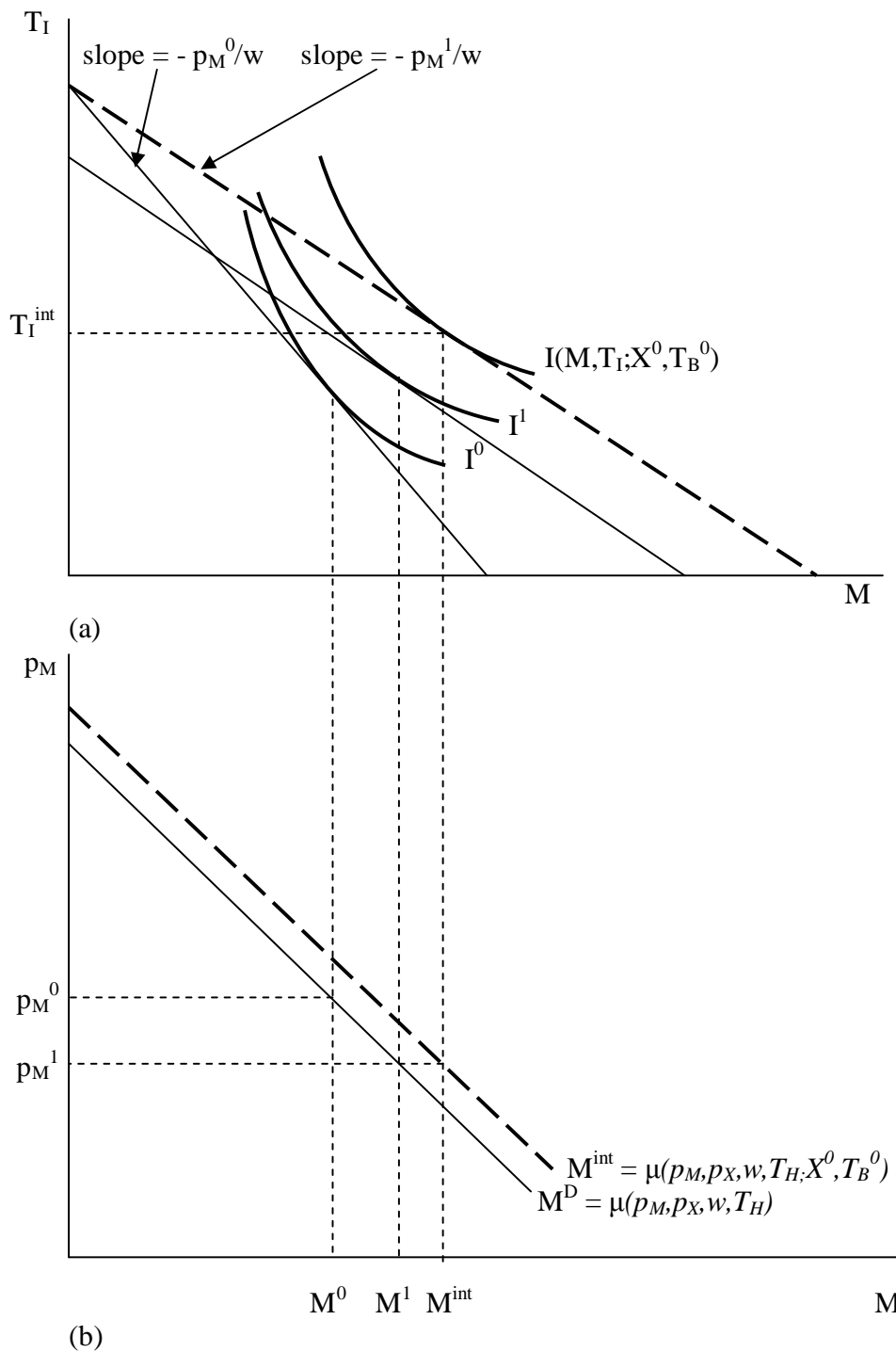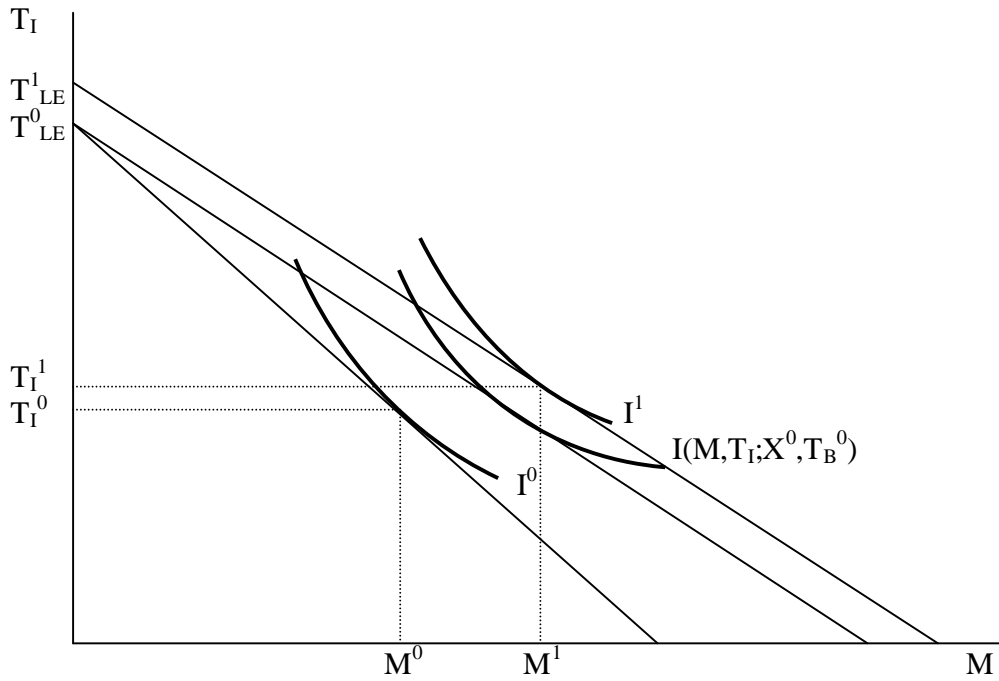
T_I

slope $= - p_M^0/w$ slope $= - p_M^1/w$

T_I^{int}

$I(M,T_I;X^0,T_B^0)$

I^1

I^0

M

(a)

p_M

p_M^0

p_M^1

$M^{int} = \mu(p_M,p_X,w,T_H;X^0,T_B^0)$
$M^D = \mu(p_M,p_X,w,T_H)$

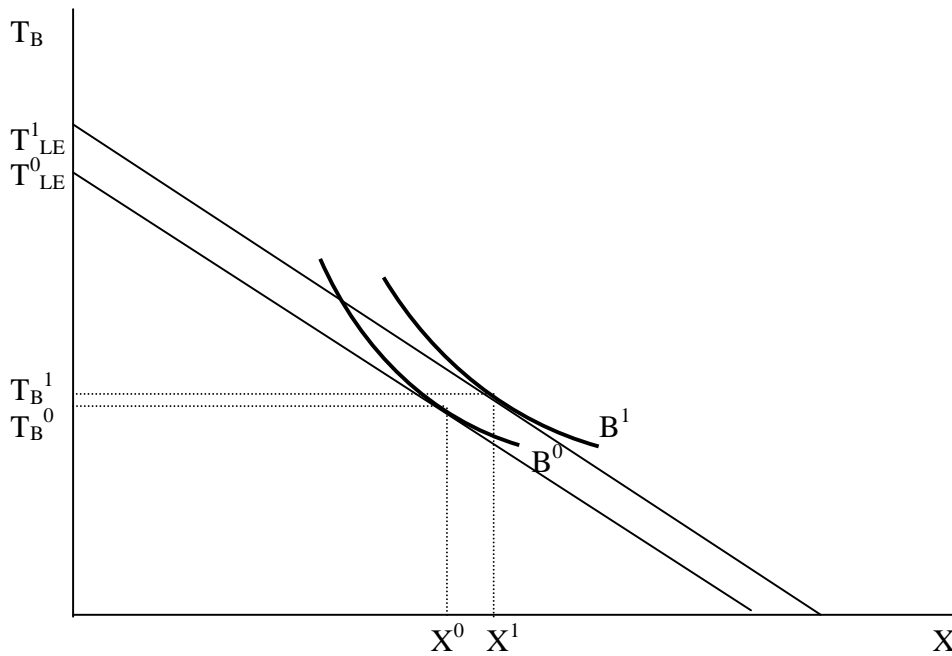M^0 M^1 M^{int} M

(b)

Figure 3.6   Derivation of the healthcare demand curve

18

relaxing the constraint that ruled out an increase in demand from M^0 to M^{int}. Consistent with the generic higher long-run price elasticity, the long-run increase in quantity

(a)

(b)

Figure 3.7   Towards work-leisure choice

demanded will exceed M^1. A difference not to be overlooked is the effect of increased healthy time on the consumption of other goods, including home-goods B. The increased healthy time does act like an income effect in that, not only it increases time available for all leisure goods but also increases income available for consumption overall through a potential income increase. Thus, in the long run, the B-I frontier would fully shift out, i.e. even beyond B^{MAX} in figure 3.5(b) above.

Another perspective on the consequences of a fall in the price of healthcare is that consumption, i.e. $C = p_X X + p_M M$, now being cheaper, part of the savings will be spent on leisure time $T_{LE} = T_B + T_I$ and, consequently, demand for healthcare is expected to go up if healthcare is not an inferior good. Empirical evidence clearly shows that healthcare is normal, at least a necessity at the individual level and possibly a luxury good at the community level[10].

The price sensitivity of the demand for healthcare ought to be negatively correlated with the seriousness of the decision. For instance, decisions concerning serious medical interventions are typically suggested by the physician-agent and followed by patient-principal approval. Even then, the price sensitivity is somewhat recovered due to the intermediation by health insurance[11] as, for serious interventions, insurance is characteristically needed to cover high costs.

3.3 Health stock model in action

The example introduced above, that of fitness clubs, essentially illustrated the full analysis of the case where a price change induced optimal adjustments in individual's choices both in the short as well as in the long run. This section will introduce two further examples, one on adjustments in response to a change in preferences and the other one in response to environmental factors.

Tobacco use has been in the forefront of policy discussions for decades and recent reductions in use in many countries (through regulation as well as education and information transmission) prove to be, at least partly, policy successes. Other harmful and addictive substances like narcotic drugs have also occupied the policy arena, not only for their negative health effects, but also in relation to individual safety and national security issues.[12] Environmental factors, on the other hand, exhibit the fundamental difference that individuals have less choice in consuming those typically external bads such as air pollution, noise, tainted water and other contaminants. Choice to avoid negative environmental factors is closely related to human capital as mobility is positively correlated with income.

The analysis of the effects of consuming healthy and unhealthy goods can be accomplished by reclassifying them as either input X in the home-goods production or as

---

[10] European Commission [2006], Feldman & Dowd [1993], Getzen [2000], Olsen [1993], Rice [1993].
[11] Dusansky & Koc [2006].
[12] Chaloupka & Wechsler [1997], Grossman & Chaloupka [1998], Pacula [1998].

positive or negative spillovers emanating from variable E, the environmental factor, in the health-investment production function. This is a useful distinction from the market goods, purchased and consumed by choice. On a related note, market goods like tobacco (broccoli) produce two effects: First, the direct utility generation and, second, the joint product of health deterioration (improvement). Changes to the parameters related to such products (e.g. their prices) induce a complex adjustment process in the individual's problem.

We consider an initial, exogenous change in the composition of home-goods towards unhealthy but still desirable goods such as over-consumption of alcohol or even food itself leading to obesity with negative health stock ocumes.[13] This change may come as a result of changes in preferences as in figure 3.5(b) above. The initial impact through negative spillover consists of a less efficient production of health-investment goods in that, with the same allocation of time T_I and expenditure $p_M M$, the production function $I(M,T_I;E)$ shifts down or, in other words, for each combination of the inputs less health-investment I is produced. In figure 3.5(c) above, healthy-time T_H thus falls from T_H^1 to T_H^2. This induces a smaller Edgeworth box as in figure 3.5(a). Individual's preferences determine whether one side or both will shrink because, T_{LE} and T_W being under downward pressure, with less healthy time T_H, the individual may continue to produce and consume the same or more of home-goods thereby letting the health-investment good to drop substantially. This change is illustrated in figure 3.5(b) by the shrinkage of PPF and the individual's reasonably strong preferences towards the home-good B whose consumption does actually increase.

The above exercise illustrates that an exogenous shift towards addictive and harmful substances results in a final decrease in health. Following a chain of changes similar to the derivation of demand for healthcare, the same result would obtain as a result of a fall in the price of a substance X used in the production and consumption of the harmful home-good B. The model thus predicts that, when cheaper, harmful but desirable substances would induce, unsurprisingly, decreases in health stock.

We now consider the impact of a particular exogenous environmental factor. As an example, the effect of an increase in air pollution (or external smoking[14]) will now be analyzed. As stated before, one way exogenous factors may impact decisions is through E in the production functions of home-goods and health investment. An increased level of, say, ambient pollution will not only affect outdoor sports and recreational activities negatively by reducing their health-improving effects, but may also prevent outdoor leisure activities from family outings to shopping trips. Assuming that health investment is more severely affected, figure 3.8(a) shows the shrinkage of the Edgeworth box, figure 3.8(b) emphasizes that while PPF shrinks, it shrinks more for I than B, and figure 3.8(c) shows that, as a result of decreased health investment, health stock and hence healthy time fall. Despite the overall shrinkage of the PPF in figure 3.8(b), the individual chooses the significantly lower level of health investment I^1 while she preserves her consumption
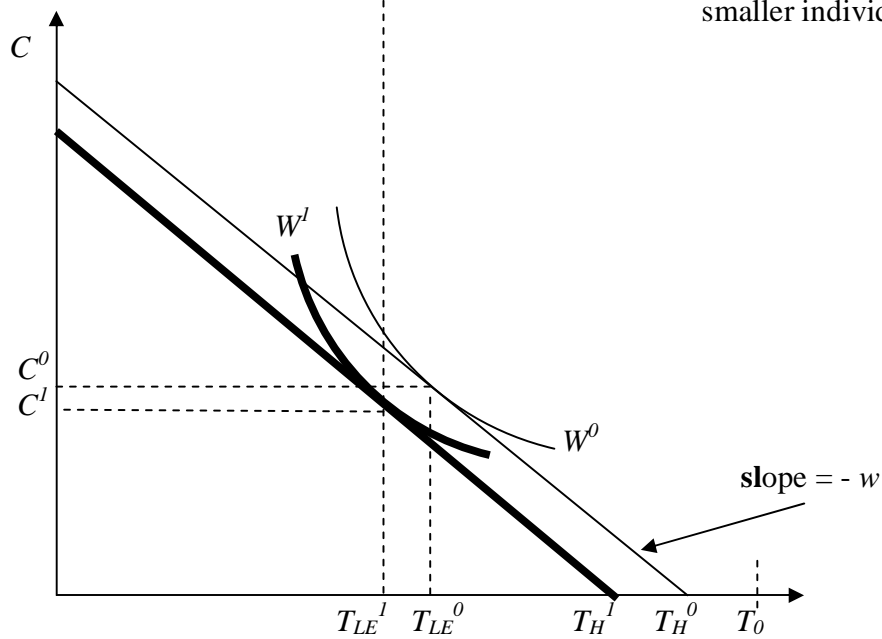
---

[13] Auld & Powell [2006], European Commission [2006] Sander & Bergemann [2003].
[14] Eisner et al. [1998].
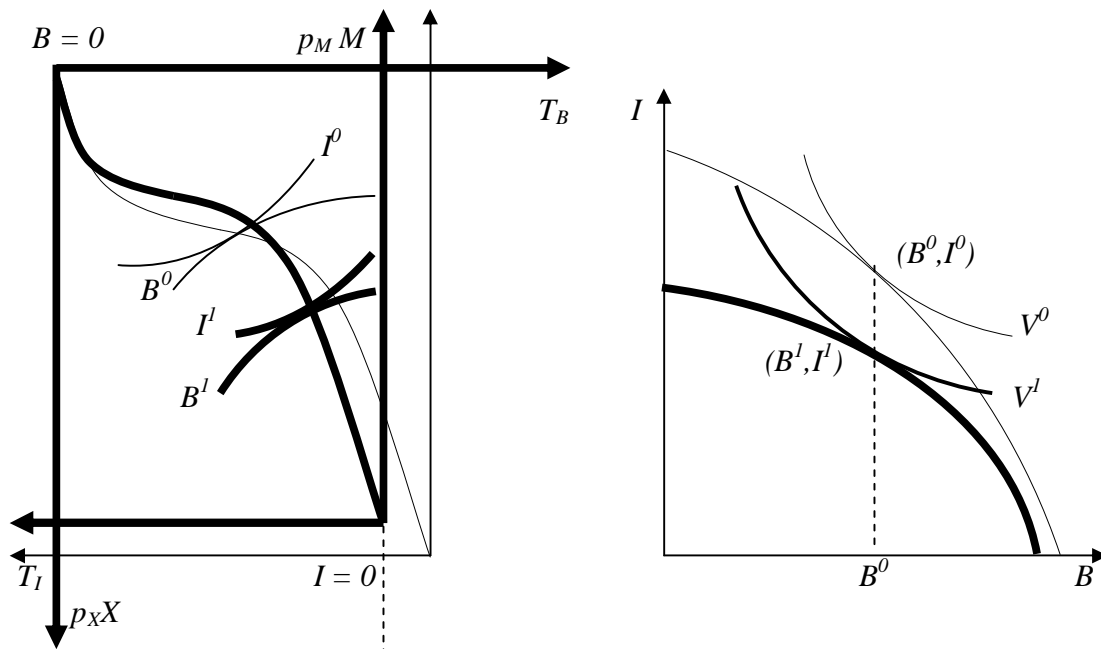
(a) Efficient reallocation of time and inputs

(b) Utility maximized over smaller individual PPF

(c) Leisure-work reallocation of time

Figure 3.8 Unhealthy exogenous change

22

of home goods at $B^1 = B^0$. Of course, this short-run adjustment may further reduce healthy time in the long run.

The analysis of an individual's healthcare decision-making has so far been conducted using the health stock model of healthcare demand. Having thus understood the individual's demand behaviour, we will briefly analyze the effects of two well-known demand management instruments used in public policy implementation.

3.4 Demand management

This section continues the theme healthcare demand in action by introducing policy instruments used for demand management. Both user fees and medical savings accounts are such tools available to policymakers.

As reimbursement insurance reduces the price of healthcare at the point of service below its cost, individuals face incentives to overuse. User fees and medical savings accounts are demand management mechanisms that can be and are used under health insurance, whether private or public. Both mechanisms can be understood within the health stock model of healthcare demand without reference to insurance.

User fees

User fees are more widely used[15] than medical savings accounts and they stand somewhere between specific per unit taxes and insurance deductibles. They differ from the first in that user fees are typically per episode in general (but per drug purchased in pharmacare) and from the second in that health insurance deductibles[16] take many forms, from per period of insurance contract to per service. They are typically used for moderating demand than a source of revenue, this demand moderation occurring through a reduction in use[17]. The demand-moderating effect of user fees is examined above in section 3.3 on demand for healthcare.

Since user fees are resisted primarily for equity reasons or being regressive taxes[18], policy-makers are hesitant to introduce them. As far as healthcare demand is concerned, introduction of user fees raises the price and, if demand is price-sensitive, use will fall. Two issues arise. First, the price-sensitivity problem is tightly related to the physician-agency problem to be analyzed later. Intuitively, once an illness episode starts, many medical services are "suggested" by the better-informed physician and typically approved by the patient. Thus, use being less decided by the patient than the physician, policies ought to target physician behaviour rather than the passive consumer. Thus, if demand is basically decided by the doctor and the user fee only affects the patient, the individual's

---

[15] See Beck & Horne [1980] for a Canadian "experiment". They are used under public insurance in countries like Sweden and France but nearly all of Western Europe for pharmacare.
[16] Newhouse et al. [1980].
[17] Creese [1997].
[18] The Economist [1997].

demand will be inelastic[19] and use will not be affected by user fees. Provider targeting policies may well be more effective. Secondly, user fees will always deter a fraction of the legitimate demand that could have acted as preventive care[20] and, consequently, cost more when the condition of the patient deteriorates. This would happen when the potential patient forgoes the first contact, typically the primary care which could decide on the seriousness of the individual's reason to access the system.

In terms of equity considerations, user fees are regressive because heavy users are normally lower income individuals. Thus, if a user fee is to be imposed as a demand management instrument, the argument must be based on efficiency. However, in terms of efficiency and, in particular, as a tax instrument, user fees would hardly qualify as a good instrument as the tax base would be very narrow if exceptions started creeping in, from children and elderly on fixed incomes to chronically ill.[21]

The more people have to cough up when they use health care, the less they will be willing to pay for it through taxes or insurance premiums. Money raised by charging might become a substitute for existing funding.

But would the fee affect the demand for health care? A forthcoming survey for the Social Market Foundation, a think-tank, and Pfizer, a drugs firm, found that 64% of British adults think people visit their doctor too often because it is free. If they are correct, a fee might help.

Most health-care systems, whether paid for by taxes or private premiums, are in effect insurance systems. Insurance cover frees people to do things they might think twice about if they were uninsured, as the insurer will pick up the tab. This can lead to inefficiencies. Having bought health cover with your premiums or taxes, why not make the most of it and visit the doctor every time you sneeze?

Sure enough, those who had to pay most used much less care than those who could, in effect, go to the doctor free. A 1% rise in out-of-pocket costs cut the amount of medical claims by around 0.3%. This suggests that there is significant scope for using fees to cut wasteful health-care consumption.

But, says Alan Maynard, an economist at York University, if charging deters sick people from going to the doctor, that might impose bigger costs on society in the long run (say, through the spread of a contagious disease) than it saves in the short run. The RAND study found no evidence that people who used less health care were less healthy than those who did, but Mr Maynard casts doubt on the reliability of that finding. He worries that charging may discourage poor people from seeking care.

---

19 Getzen [2000], Olsen [1993].
20 Laupacis [1996], Marshall [1996].
21 Birch [2004].

Medical savings accounts

Medical savings accounts (MSAs) are a more recent innovation in many countries (e.g. Singapore, China, Hong Kong, South Africa, US)[22] and they have probably become feasible due to advances in electronic individual accounts kept by fiscal systems. MSAs do indeed resemble bank accounts except that they are, by design, coupled with a catastrophic illness insurance that carries a high deductible[23]. Of course, this catastrophic illness insurance is offered by markets wherever MSAs plug into a predominantly market-based healthcare system (e.g. the US) while public insurance picks up where MSAs fall short under a predominantly public insurance system (e.g. Singapore).

The accounts are designed to pay for individual's or family's healthcare expenditures. The money put aside in the MSA is used to pay for qualifying (usually routine or minor) medical expenses while the accompanying insurance plan covers (usually major) expenses after the deductible level has been reached. Normally, the admissible range of expenditure ought to cover preventive and chronic care items. However, as the administrative expenditures avoided by small claims that do not require insurer approval is one advantage of MSAs, there is always a grey area where contacts that may prevent major future expenditures are hard to define as admissible.

Under MSA schemes, money contributed to the MSA belongs to the account holder, accumulates on a tax-free basis, and is not included in taxable income if used for admissible health care expenses. With public insurance plans (including the US Medicare and Medicaid), the insurer deposits into MSAs would come from the government whereas, under employer-mandated systems[24], deposits would come from employers. Of course, where service provision is private and unrestricted, individuals would have further incentives to make their own deposits. This property provides the individual with incentives to save for future expenses while preventing current small and avoidable claims thus becoming "activated consumers".[25] However, if this incentive is to work, individuals must perceive MSAs as hard budgets rather than just guidance.[26]

Many arguments have been made for and against MSAs. This section will cover only those relevant to individuals' decisions. First, as stated above, "activated consumers" have strong incentives to search for better prices.[27] This incentive must, however, be

---

[22] Hsiao [1995].
[23] Bloche [2006].
[24] One advantage of MSAs under employer-mandated systems would be the portability of coverage and the ensuing boost to labour market efficiency by removing the health insurance lock (Robinson [2005]).
[25] Lee & Zapert [2005].
[26] Hurley et al. [2007].
[27] Gratzer [2002].

combined with competition on the supply side to record efficiency improvements. Most public insurance systems are also characterized by capacity constraints[28] that would impair competition. Moreover, lower premia originating from high deductibles may induce individuals to take up the catastrophic illness insurance coupled with MSAs. However, a major efficiency loss and adverse selection would result from those already insured and the relatively wealthier benefiting from tax breaks.

3.5 Conclusions

This chapter covered the analysis of an individual's demand for healthcare. Healthcare is perceived as the set of purchased inputs into health investment by the individual, the other scarce input being the individual's time. The two inputs allocated to health investment, time and money, compete with the production of home-goods or, simply, all other goods. The individual's total healthy time is allocated to health investment and home-goods as well as to work that enables her earn income. The healthy time is positively correlated with the individual's health stock which, while depreciating over time, is positively correlated with health investments.

Since individuals value their future as well as their present health, they are aware of the fact that a positive net health investment means a higher health stock tomorrow and a higher health stock produces an enhanced ability to earn while the opportunity cost of such investment is the current consumption of home-goods. Thus, individuals have to allocate their resources within and across time periods. Taking into account the intertemporal tradeoff, what the individual plans to purchase and use in the production of health investment constitutes her demand for healthcare.

However, two further components are essential in a complete analysis of healthcare demand. First, the healthcare insurance need which arises from the lumpy and uncertain healthcare expenses does act on healthcare demand as market signal even if healthcare may be free at the point of purchase.[29] At first glance, health insurance may seem independent of an individual's valuation of health but, upon closer inspection, the accumulation of contingency assets in the absence of health insurance may come at the expense of current health. Thus, beyond the pure insurance motives, health insurance will interact with one's healthcare demand as developed above. Second, as discussed in the context of user fees, healthcare demand is not independent of medical information required to make health investment choices. The doctor as agent is an essential component of healthcare demand as the doctor, as the informed party (i.e. informed on medical technology), determines medical care use both in the selection of treatments and their intensity of use. While this determination takes the form of suggestion by doctor and acceptance by the patient, it does nevertheless suggest that the patient typically follows doctor's orders. Thus, the healthcare demand derived in this chapter must be construed as that of a perfectly-informed patient or principal.

---

[28] Hurley [2002], Forget, Deber & Roos [2002], Shortt [2002].
[29] Dusansky, & Koc [2006].

References

Auld, M.C. & L.M. Powell [2006], "The economics of obesity: Research and policy Implications from a Canada-US comparison", in Beach et al. [2006]

Beach, C.M., R.P. Chaykowski, S. Shortt, F. St-Hilaire & A. Sweetman (eds.), Health Services Restructuring in Canada: New Evidence and New Directions, 2006, McGill-Queen's Univ. Press

Beck, R.G. & J.M. Horne [1980], "Utilization of publicly insured public health services in Saskatchewan before, during and after copayment", Medical Care 18, 787-806

Becker, G.S. & K.M. Murphy [1988], "A theory of rational addiction", J. Pol. Econ. 96, 675-700

Birch, S. [2004], "Charging the patient to save the system? Like bailing water with a sieve", Canadian Medical Association J. 170 (120, 1812-1813

Bloche, M. G. [2006], "Perspective", New England J. Medicine 355, 1756-1759

Brainerd, E. [2005], "Autopsy on an Empire: Understanding Mortality in Russia and the Former Soviet Union" J. Economic Perspectives 19, 107-130

Chaloupka, F.J. & H. Wechsler [1997], "Price, tobacco control policies and smoking among young adults", J. Health Econ. 16, 359-373

Chang, F.-R. [1996], "Uncertainty and investment in health", J. Health Econ. 15, 369-376

Creese, A. [1997], "User fees: They don't reduce costs, and they reduce inequity", British Medical J. 315, 202-203

Dusansky, R. & C. Koc [2006], "Health care, insurance, and the contract choice effect", Economic Inquiry 44(1), 121-127

(The) Economist [1997], "Coughing up", The Economist October

Eisner, M.D. et al. [1998], "Bartenders' Respiratory Health After Establishment of Smoke-Free Bars and Taverns", JAMA 280(22), 1909-1914

European Commission [2006], "European Economy", Special Report No.1, Directorate General for Economic and Financial Affairs

Feldman, R. & B. Dowd [1993], "What does the demand curve for medical care measure", J. Health Econ. 12, 193-200

Forget, E.L., R. Deber & L.L. Roos [2002], "Medical Savings Accounts: Will they reduce costs?", CMAJ 167(2), 143-147

Getzen, T.E. [2000], "Healthcare is an individual necessity and a national luxury: Applying multilevel decision models to the analysis of healthcare expenditures", J. Health Econ. 19, 259-270

Gratzer, D. [2002], "It's time to consider Medical Savings Accounts", CMAJ 167(2), 151-152

Grossman, M. [2000], "The human capital model", in A.J. Culyer & J.P Newhouse (eds.), Handbook of Health Economics, Elsevier

Grossman, M. [1972], "On the concept of health capital and the demand for health", J. Pol. Econ. 80(2), 223-255

Grossman, M. & F.J. Chaloupka [1998], "The demand for cocaine by young adults: A rational addiction approach", J. Health Econ. 17, 427-474

Hsiao, W.C. [1995], "Medical Savings Accounts: Lessons From Singapore",
Health Affairs , 14(2), 260-266

Hurley, J. [2002], "Medical Savings Accounts will not advance Canadian health care
Objectives", CMAJ 167(2), 152-153

Hurley, J., G.E. Guindon, V. Rynard & S. Morgan [2007], "Publicly Funded
Medical Savings Accounts: Expenditures and Distributional Impacts", CHEPA
WORKING PAPER SERIES, Paper 07-01, McMaster University

Laupacis, A. [1996], "Preventive therapies: Weighing the pros and cons", CMAJ 154(10),
1510-1513

Lee, T.H. & K. Zapert [2005], "Do High-Deductible Health Plans Threaten Quality of
Care?" New England J. Medicine 353, 1202-1204

Marshall, K.G. [1996], "Prevention", Parts 1-4", CMAJ 154(10), 154(12), 155(3),
155(4)

Newhouse, J.P., J.E. Rolph, B. Mori & M. Murphy [1980], "The effect of deductibles on
the demand for medical care services", J. American Statistical Association
75(371), 525-533

Olsen, J.A. [1993], "But *health* can still be a necessity …", J. Health Econ. 12, 187-191

Pacula, R.L. [1998], "Does increasing the beer tax reduce marijuana consumption",
J. Health Econ. 17, 557-585

Rice, T. [1993], "Demand curves, economists, and desert islands: A response to Feldman
and Dowd", J. Health Econ. 12, 201-204

Ried, W. [1998], "Comparative dynamic analysis of the full Grossman model",
J. Health Econ. 17, 383-425

Robinson, J.C. [2005], "Health Savings Accounts — The Ownership Society in Health
Care" New England J. Medicine 353, 1199-1202

Sander, B. & R. Bergemann [2003], "Economic burden of obesity and its complications
in Germany", Eur. J. Health Econ. 4, 248-253

Shortt, S.E.D. [2002], "Medical Savings Accounts in publicly funded health care
systems: enthusiasm versus evidence", CMAJ 167(2), 159-162

Smith, M. [2001], "Medical savings accounts – A comparative overview", Law and
Government Division, Depository Services Program, PWGSC, Government of
Canada, http://dsp-psd.pwgsc.gc.ca/Collection-R/LoPBdP/BP/prb0120-e.htm

Discussion questions

01. If health is a stock, what flow variables affect its level and how? Classify them as
exogenous and endogenous.

02. Does the health stock model suggest a value for one's life?

03. How would one incorporate a major but non-fatal illness into the Grossman model
during the course of life rather than at the end?

04. How and whether does one plan for the end of life in the health stock model?

05. Is there a difference between an individual's responses to the health effects of newly-developed preventive measures and of treatment options?

06. How can one explain the relationships between "health and wealth" and "health and education"? Can we presume causality?

07. What major modifications would have to be brought into the Grossman model for a more realistic model of healthcare demand?

08. Indicate some demand-side and supply-side policies over health. How would they enter the Grossman model?

09. In the light of Grossman model, why would one invest in his own health?

10. Grossman's health stock model of healthcare demand abstracts from two problems inherent to healthcare. First noted in Arrow [1963], uncertainty and physician-agency make healthcare markets prime candidates to fail. Discuss why or why not.

11. Major components of demand for healthcare may be infrequent and lumpy. How does the health stock model need to be modified?

12. Given that habit models (such as in tobacco and alcohol use) reinforce the need for a sustained application of anti-addiction policy tools, can such policy measures be understood within the Grossman model?

13. Is the introduction of user fees is unavoidably inequitable?

14. The Grossman model must incorporate the incentive mechanisms needed to empower patients as consumers. Existing medical savings accounts posses such mechanisms. Discuss.

Problems

01. The amount of other goods an individual is willing to give up to improve her health falls. As a result of this change in her preferences the individual would normally choose a lower amount of health investment but the amount of home goods might well decrease too. Explain why.

02. What are the consequences of an exogenous increase in preference for smoking in the Grossman model?

03. What are the consequences of an exogenous increase in pollution for smoking in the Grossman model? Highlight, if any, the differences with question 2.

04. <u>Graphically</u> explain the efficient allocation of one's total healthy time between health-generating activities and home-goods.

05. In the health stock model of healthcare demand, a given bundle of time and purchased inputs in the Edgeworth box belongs to the contract curve. Explain the condition such bundles have to satisfy beyond the feasibility constraints that times T_B and T_I add up to the chosen total leisure time and expenditures on purchased inputs add up to the work income.

06. In the health stock model of healthcare demand, a change in preferences induces an income effect through the change in healthy time T_H. Suppose an individual's preference for I over B increases. Explain the intertemporal income effect.

07. Is there a difference between smog and second-hand smoking?

08. Obesity is not a desired state of health in that it worsens enjoyment of life and, worse, generates comorbidities. How could one design preventive fiscal policies?

09. The smoke-free workplace regulation is a property rights assignment. Does it have a different long-run effect on healthcare demand than in the short run?

10. What are the various differences between a smoke-free workplace and a workplace with a properly ventilated and insulated smoking room?

11. How would the demand for healthcare change if healthcare became more expensive but also more effective?

12. What would happen to healthcare demand if an increase in health stock boosted one's wages?

13. Briefly explain how preventive public policies might work for smoking and addictive drugs. Make sure you mention how individual incentives affect the transmission process from policy to outcome.

14. Carefully explain a policy of introducing user fees equitably into the Grossman model.

15. Briefly describe one type of medical savings account (MSA). Why does one need the intertemporal preferences and an intertemporal budget constraint to analyze MSAs under a Grossman model?

Appendix A Work-leisure choice as part of the health stock model

This appendix shows that when the individual solves her problem (3.7), she thereby splits her healthy time between leisure and work, i.e. the leisure-work tradeoff is implicitly and simultaneously solved. This decision involves equating the marginal benefits of time allocated to producing the home-good B and to the health-investment good I (in the optimal proportions as given in the Edgeworth boxes) and the opportunity cost of time used in those activities as the wage rate.

This problem is more complicated than simply trading time off between leisure as idle time and work because, in this case, leisure time consists of choosing between producing B vs. I whereas working generates the income necessary to purchase the inputs towards the production of B and I. Even the supposedly idle time of sleep, perceived more as I than B, requires the purchase of inputs such as a bed and of time as non-work.

The individual's problem (3.7) of resource allocation (illustrated in figures 3.3 and 3.5) can, upon suppressing the exogenous variable E in the health investment and home-good production functions for simplicity, be rewritten as

$$
\begin{aligned}
&\max_{\{X, M, T_B, T_I\}} \quad V\big(B(X, T_B), I(M, T_I)\big) \\
&\text{s.t.} \quad T_H = [(p_M M + p_X X)/w] + T_B + T_I.
\end{aligned}
\tag{A1}
$$

First, the constraint can simply be rewritten as

$$
T_H = T_W + T_{LE}.
\tag{A2}
$$

Second, we define consumption C as the total expenditure on inputs M and X as

$$
C = p_M M + p_X X
\tag{A3}
$$

or, noting that $p_M M + p_X X = w T_W$,

$$
C = w T_H - w T_{LE}.
\tag{A4}
$$

Third, efficiency in production requires that, for the production of each good, the individual has to equate MRTS (marginal rate of technical substitution) to the relative price of the inputs and satisfy a time-budget constraint. This process yields two equations for each of B and I. For B we obtain

$$
B_T(X, T_B)/B_X(X, T_B) = w/p_X \text{ and } T_B = T_{LE} - (p_X / w)X
\tag{A5}
$$

which, solved, yield conditional demands for the inputs

$$X = x(p_X/w, T_{LE}) \ and \ T_B = t_B(p_X/w, T_{LE}) \tag{A6}$$

and, similarly, for I,

$$M = m(p_M/w, T_{LE}) \ and \ T_I = t_I(p_M/w, T_{LE}) \tag{A7}$$

from $p_M I_T(M,T_I)/ I_M(M,T_I) = w/p_M$ and $T_I = T_{LE} – (p_M/ w)M$. $\tag{A8}$

Finally, exchange efficiency (see the contract curve in Fig.3.2) between the two goods productions requires that the marginal rates of substitution between the use of time and of expenditures on inputs be equalized across the two production lines, i.e.

$$p_X B_T(X,T_B)/B_X(X,T_B) = w = p_M I_T(M,T_I)/ I_M(M,T_I) \tag{A9}$$

Now we make the following successive substitutions into the objective function in A1. We eliminate X and T_B using (A3) and (A9). Then, using A7, eliminate M and T_I (as A6 is not independent of A7 as, given T_{LE}, once X and T_B are chosen M and T_I are automatically determined). Thus the new objective function is given as $W(T_{LE},C)$. Thence, recalling the constraint in A4, the work-leisure choice problem can be written as

$$\begin{aligned}
\max_{\{T_{LE},C\}} \quad & W(T_{LE},C) \\
\text{s.t.} \quad & C = wT_H - wT_{LE}
\end{aligned} \tag{A10}$$

We note that A10 can't be solved on its own but, rather, is derived from the original problem A1 in order to understand the implied work-leisure choice. Unlike standard work-leisure choice treatments where one chooses between idleness and consumption, leisure in the current context is the time spent in the consumption of the produced goods B and I whereas work allows the production of these latter through purchase of required inputs X and M. Finally, since the information on preferences over produced goods is suppressed in $W(T_{LE},C)$, this latter can't be used in comparing changes in an individual's situation in response to exogenous changes. It is only meant to serve as criterion in splitting the available healthy time between work and leisure.

Appendix B The intertemporal problem and medical savings accounts

The fundamental force driving the dynamics of the health stock model is the health-depreciation coefficient. While it is true that it is partially endogenous through its dependence on individual's lifestyle and hence net health investments, for a reasonable understanding of the aforementioned dynamics it would suffice to have an exogenously given time profile of the depreciation coefficient. Thus a typical individual with perfect foresight chooses its lifetime allocation of time and expenditures by solving the following problem.

$$\max_{\{X,M,T_B,T_I\}} \int_0^T e^{-\theta t} U(B,H)\,dt$$

s.t.
$$B = B(X, T_B; E) \tag{B.1}$$
$$I = I(M, T_I; E)$$
$$T_0 = T_L(H) + T_W + T_{LE}$$
$$wT_W = p_M M + p_X X$$
$$T_{LE} = T_B + T_I$$
$$\frac{dH}{dt} = I - \delta H$$
$$H(0) = H_0$$
$$H(T) = H_{min}$$
$$H(t) > H_{min},\ 0 < t < T$$
$$M(T) = 0$$

Nearly all the equations above were explained in the text. The last four are required to complete the dynamic health stock model. Several explanatory remarks are in order. First, the problem is continuously re-solved by individuals. Thus the starting time zero is not necessarily birth. The initial health stock H_0 denotes one's stock whenever the individual reevaluates her choices. Second, at the terminal time, bizarrely planned but discussed earlier in the chapter, is when one's health stock falls below H_{min}. Above in B.1, death is taken to be at H_{min}. Finally, assuming that the health stock model is solved using optimal control theory, $M(t)$ is the so-called Hamiltonian. This latter yields the instantaneous value of continuing the program in B.1 by combining the current utility derived from the consumption of produced goods B and I but also valuing the next period by taking into account the tradeoff that, beyond its consumption value, health-investment replaces the depreciated health stock as well as possibly adding to it. Thus, if the value of the Hamiltonian falls to zero, there is then no utility need to continue running the program B.1 or, in other words, the optimal terminal time T has been reached.

The omission, above, of an intertemporal budget constraint for the individual has one policy-relevant serious consequence. Modeled in the original Grossman treatments ([1972], [2000]), the intertemporal budget constraint is necessary in understanding medical savings accounts (see Forget et al. [2002], Gratzer [2002], Hsiao [1995], Hurley [2002], Hurley et al. [2007] and, especially, Smith [2001] for a comprehensive survey). As all known medical savings accounts exhibit intertemporal dynamics in the form of savings accumulating tax-free for future healthcare expenditures, the above budget constraint is hard for every period without allowing surpluses now for later expenditures. A fully dynamic budget constraint would relax the succession of hard period constraints and allow the individual more flexibility.