TAMPERE ECONOMIC WORKING PAPERS
NET SERIES

# MEDICAL DECISION CRITERIA AND POLICY FOR AN EFFICIENT ALLOCATION OF RESOURCES

Pekka Ahtiala

Working Paper 8
December 2001
http://tampub.uta.fi/econet/wp8-2001.pdf

DEPARTMENT OF ECONOMICS

FIN-33014 UNIVERSITY OF TAMPERE, FINLAND

ISSN 1458-1191
ISBN 951-44-5278-X

December 2001

# MEDICAL DECISION CRITERIA AND POLICY FOR AN EFFICIENT ALLOCATION OF RESOURCES

by

Pekka Ahtiala*
University of Tampere
Department of Economics
33014 University of Tampere
Finland
Tel +358-(0)3-364 5122

**Abstract**

This paper approaches health care criteria by maximizing society's consumption possibilities in a model where health is a special case of a good produced, consumed, and used as an input in production, and the patient chooses from alternative therapies. It complements the conventional approach, in providing conditions under which it is optimal to provide care beyond the public health standard. It is shown to be optimal to provide health care beyond the previously obtained optimum where the marginal product generated by the care equals its marginal social cost, up to the point where the sum of the marginal product and the marginal utility equals that cost - but only if the patient is willing to pay the full marginal social cost of the part that exceeds the marginal product, out of his after-tax income, the part corresponding to the marginal product being deductible from taxable income. For the decision, the social planner needs to know the costs of different therapies and the times they take to bring the patient to working condition, as well as the patient' s labor income in this condition, but not the patient's preferences. The higher the patient' s labor income, the more it is optimal to spend on more efficient therapy, and provide a "tax subsidy" by keeping the expenses for the investment part tax-deductible. An increase in hospital capacity leads to treating patients with cases medically minor to those treated before, incomes equal, if rationing is done optimally so that this finding is not necessarily a sign of demand-shifting.

Keywords: optimal medical care, incentives for optimal care

JEL classification number I1.

## 1. Introduction

Modern medicine has greatly improved our ability to keep people alive, although often at an increasing cost. Thus it is now society's resource constraint that has emerged as a key reason why it is no longer possible to sustain every person' s life by "all possible means". The new situation calls for new thinking on the decision criteria of health care.[1]

In the future, and in many countries today, we seem to be faced with a dual system. A certain amount of care is provided by the public health system. The patient may wish to purchase more, and society has an interest in ensuring that he has an incentive to purchase neither too little nor too much of it. Determining the optimal amount of this care and the policy to achieve it is the purpose of this paper, given the care provided by the public health system.

Our approach is based on the following notions. Medical decisions are special cases of resource allocation decisions on three often partly overlapping goods. If the "health" so produced is regarded as a publicly provided or financed free good, we are dealing with a "public health good". To the extent that it makes the patient more productive (e.g. fit for the labor force), we are dealing with an investment good, which will increase society's total output, and thus its consumption possibilities of both health and other goods (see Grossman (l972a,b)). To the extent that it makes the patient feel better it is a consumption good.

The above distinction is as important on a conceptual level as it is in most other fields, because the evaluation criteria of any expenditures should be determined by whether the expenditure is regarded as investment or consumption, and whether it is public or private. Moreover, since the opportunity cost of any resources on health is forgone consumption of mainly other goods, the issue of resource allocation should be addressed not only in the context of the health care sector, but also between health on the one hand, and other goods, on the other.

In the above classification, much of the literature on the subject deals with care that can be regarded as being related to the public health good. The minimum health standards secured by the public health system (or by full insurance coverage) are generally goods with a practically zero user cost, provided or financed publicly (or by the insurance company) irrespective of the ability to pay or of the marginal product generated by that health. The care to produce them is to correct for various market failures, i.e. to take care of "equity", "value of life", merit good, and externality considerations, and its extent and coverage is determined as a political decision. The issues associated with this “public health standard” are not addressed here. However, it is often consistent with the efficiency of resource allocation to provide care beyond this standard, and the focus of this paper is examining the optimal criteria for that care.

Our survey of earlier writings is brief, as this author is not aware of writings taking a similar “partial general equilibrium” approach, with health in its three roles. Arrow (1963) lays the foundations for the study of the market for medical care in general. Smallwood and Smith (1976) take an optimizing approach, analyzing the dichotomy of the life and death of the patient. They minimize the expected patient loss of the patient population, subject to a capacity constraint. Pauly (1979) suggests that the expected value of income or of the therapeutic gain of care be used as the criterion of care. There is an extensive literature where the criterion for care decisions is the number of quality-adjusted life years produced by that care for the patient, each patient counting equally. This QALY approach has been used as a criterion for the public health good.

The purpose of this paper is to devise an optimal tax and subsidy treatment of health care expenditures, as well as rationing criteria, for care in excess of the level provided by the public health system. Like a committee on such a reform, we will focus on the task on hand and take the care provided by the public health system, and the general tax and subsidy system as given. We will build a social welfare-maximizing model with standard assumptions regarding the behavior functions, where health is a consumption good and an investment good, produced in the economy. The problem consists of two parts. First, the social planner determines the socially optimal amount of health in individual cases. Secondly, he sets the tax and subsidy statuses of health care expenditures so as to induce the patient to choose the optimal amount -- without knowing the individual utility functions. The criteria derived are intended to provide a basis for more practical guidelines. The guidelines, together with the theory of physician

behavior, make it possible at least in principle to determine an incentive system to contribute to an efficient allocation of resources.
It will be shown that it is optimal to provide health care beyond the conventional optimum point where the marginal product generated by the care - the product of the marginal gain in the patient's productivity and his income in his normal state of health - equals its marginal social cost, up to the point where the sum of the marginal product and the marginal utility equals the marginal social cost.

How the patient can be induced to choose the optimum amount depends on the user cost of care and whether the patient has sick leave with pay. If the user cost equals the marginal social cost and the patient gets no sick pay, it is always optimal to provide care up to the point where its marginal social cost equals its marginal social product, and it is optimal for the patient to choose this amount if its cost is deductible from taxable income. Besides, it is optimal to provide more if the patient is willing to pay the difference between the marginal social cost and the marginal social product of the increment out of his after-tax income, while the part corresponding to the marginal product is deductible. However, if the patient has fully paid, rather than unpaid sick leave, the former, or the investment care should be free for the patient to choose the optimal amount.

If the user cost of care is kept below its marginal social cost, the optimal care is the same, but rationing is necessary for an optimal amount of care to be chosen. In rationing, it is optimal to rank the patients according to their marginal products from care, and give each patient treatment down to the point where the marginal product equals the marginal social cost of care – or its shadow price. However, if the patient requests more care and is willing to pay the part of the marginal social cost exceeding the marginal product out of his after-tax income, the part corresponding to the marginal product being deductible, it is optimal to provide him with that care. Likewise, it is optimal to admit a patient rationed on the above criterion if he is willing to pay enough to get above the line in the above manner. Of course, these results are independent of who actually delivers the care, a government or a private organization.[2]

To apply this approach on an crude level, the physician needs to know only the patient' s labor income in his normal state of health, the costs of alternative kinds of care, and the times he expects them to take to bring the patient to working condition, but not the patients' utility functions.

The above has several implications. The higher the patient' s labor income, and the greater the expected marginal therapeutic gains from care, the more it is optimal to spend on more efficient (and safer) therapy. It is also optimal to give a greater "tax subsidy" to high labor income earners by making investment health expenses, but not private consumption health expenses deductible from taxable income - contrary to conventional wisdom. The greater the marginal social rate of time preference and the more remote the patient' s entry into the labor force, the less it is optimal to spend on the patient' s care. Care to hopeless cases is not justified by investment criteria, which makes it consumption. The same applies to health insurance premia with different coverages: Costly first class care may be optimal and its investment part should be tax-deductible for high labor income earners, but not necessarily for others.

If rationing is done optimally, an increase in hospital capacity leads to treating patients with cases minor to those treated before, adjusted for incomes, and it is not necessarily a sign of demand-shifting. These and other findings suggest that much of the evidence presented in support of the supplier-induced demand (SID) doctrine is also consistent with socially optimal physician behavior.

In our model with a homogenous labor force, sick pay, payroll taxes, and substitutes' higher cost can provide the patient's employer with an incentive to pay part of the bill, which may lead to the patient getting more care than the social optimum.

While many may regard the above criteria as "unethical", we find it important to know what efficiency of resource allocation calls for beyond the public health standard so that the real resource cost of using other criteria could be determined. Namely, it is always optimal to provide this amount of care irrespective of the other objectives of health care policy, since it brings society to its production possibility frontier and maximizes the consumption possibilities of society' s members.

Of course, meeting first-best conditions (even if in a bounded rationality sense) in a subsector of the economy is no guarantee of a superior equilibrium in the whole economy of a second-best world. However, economies are reformed in a piecemeal way subject to a variety of constraints, and this paper attempts to provide conditions for the reform of one subsector. A superior solution is guaranteed if all the sectors in the economy achieve a first-best solution. It is the task of other papers to provide optimal solutions regarding the variables that we take as given.

The paper proceeds as follows. A model to determine socially optimal decision criteria is constructed and studied in Part 2. We will first determine the optimum, and thereafter see how the policy-maker can induce individuals to purchase the optimal amounts, by means of taxes and subsidies. This is followed by an illustration with a numerical example. Subsequently, we will derive the propositions enumerated above. Part 3 is the conclusion. The employer's incentives to pay part of the sick bill and cause a deviation from the social optimum are part of the puzzle, and their interplay with the patient's incentives is studied in the Appendix.

## 2. Optimal Criteria for Medical Decisions

### a. The Optimum

The social planner's objective is to maximize the welfare of society by determining the amount of (possibly subsidized) health care given to society's members beyond the public health standard, and thereafter determining the tax and subsidy status of different kinds of care to induce individuals to purchase the optimum amount. For the purposes of this problem, society's welfare is the sum of the utilities of its members, any considerations of equity, need, and externalities being taken care of by exogenous taxes and transfers and by the public health good. Each individual maximizes his utility subject to his budget constraint, given the prices he faces. The individuals' utilities are a function of their health and of other goods, where the former reflects the consumption good aspect of health. The planner does not know the individual utility functions. The individuals' health consists of their initial states of health plus the exogenous free public health good, and of the individually purchased health. In addition to being a consumption good, health is also an input in production, improving the individuals' productivity. Since individuals are paid their marginal products, their disposable income, used for the purchase of health and other goods, is a function of their actual state of health.

Production of both goods takes place on competitive markets, with labor as the variable input.

Examine the following model, with the consumer and producer sectors, and two goods, health (X) and other goods (V). We have:

$$U_T = \sum_{i=1}^{k} U_i (X_{0i} + X_{Pi} + X_{Ii}, V_i) \quad (1)$$

$$Y_{Di} = (1-t)\, Y_i = P_X^c X_{Ii} + V_i \quad (2)$$

$$Y_i = z_i \bar{Y}_i \quad (3)$$

$$z_i = z_i(X_{0i} + X_{Pi} + X_{Ii}) \quad (4)$$

Equation (1) is the social welfare function, where welfare is the sum of the utilities of society' s k members (See e.g. Ng (1975)). However, the policymaker does not know the individuals' utility functions. Each individual maximizes his own utility, which is a function of his state of health $(X_0 + X_{Pi} + X_I)$ and of other goods (V). The initial state of health $X_0$ reflects possible depreciation and losses caused by illness during the period, and $X_{Pi}$ is the health provided by the public health system, henceforth called the "public health good". The $X_0 + X_P$ can be augmented by individually purchased health $(X_I)$. The utility functions have positive first and negative second derivatives with respect to their arguments, assumed to be normal goods.

In Eq. (2), the individual' s disposable income ($Y_{Di}$) is one minus the net tax $(t > 0)$ or income transfer $(t < 0)$ rate, times his labor income $Y_i$. (Income should in this context be interpreted as including services not traded on the market, e.g. those of a housewife).[3] The $Y_{Di}$ is spent on the two goods, with $P_X^c$ the relative user price of X to the individual in terms of other goods V, the numeraire. To keep the expressions simple, we will set the user cost of public health to zero, since it is beyond the scope of this paper. Since we are interested in decision criteria for individual situations, we do not need to treat the public sector budget constraint explicitly, but taxes are collected to finance $X_P$, income transfers, and other government expenditures outside this model. This enables us to treat the tax rate as a parameter.

In Equation (3), the worker' s labor income equals his marginal product and is his productivity rate ($z_i$) times his income in his normal state of health $(\bar{Y}_i)$. Naturally, the utilities and incomes should be understood as the present values of the respective flows. An operational counterpart for z could be e.g. the number of hours worked. The $\bar{Y}$ would be the wage rate.

In Equation (4), $z_i$ is a function of the individual's state of health, consisting of $X_{0i}$, $X_{Pi}$, and $X_{Ii}$. Naturally, it also reflects any effects that the worker's health may have on his level of effort and his labor-leisure decisions. This "production function" has a positive first and negative second derivative with respect to health.[4]

The production sector consists of two industries, one for each good. Both have standard production functions (Equations (5) and (6)), with positive first and negative second derivatives with respect to effective labor (zL), the variable input, where z is the average productivity rate of labor and $L_q$ the amount of labor employed in industry q (q = X,V).

$$X = f_X(zL_X) \tag{5}$$

$$V = f_V(zL_V) \tag{6}$$

$$\pi_X = P_X X - zL_X\bar{Y} \tag{7}$$

$$\pi_V = V - zL_V\bar{Y} \tag{8}$$

$$X = \sum_{i=1}^{m}(X_{Pi} + X_{Ii}) \tag{9}$$

$$L = L_X + L_V \tag{10}$$

Equations (7) and (8) are the profit functions of the two industries, where profit $(\pi)$ is the difference between sales revenue and labor costs, with $P_X$ the relative price of health and $\bar{Y}$ the average income of labor in its normal state of health. We will refrain from complicating the model by means of explicitly introducing indirect taxes, since their effects are obvious from the expressions below. Eq. (9) is the market equation of health, stating that its total production is the sum of the $X_{Pi}$ and $X_{Ii}$. Eq. (10) is the labor market equation, stating that the exogenous labor force (L) equals the sum of $L_X$ and $L_V$.

It is worth noting that our production function of health (5) nests that of health, with health care as the input, and that of health care, with labor as the input. We have: $\partial X / \partial L = (\partial X / \partial H)(\partial H / \partial L)$, where H is health care, and qualitative statements regarding health also apply to care. Thus: $P_H / P_X = \partial X / \partial H$.

Maximizing profits with respect to ($L_X$) and ($L_V$) subject to the production functions and market equations on competitive markets yields the producer optimum:

$$(\partial V / \partial L)/(\partial X / \partial L) = P_X \tag{11}$$

The producer optimum is standard: the relative producer price of X reflects its marginal social cost of production relative to that of V.

The policy-maker is concerned with the marginal social product and the marginal social cost of health ($(\partial z / \partial X_I)\overline{Y}$, and $P_X$, respectively), rather than their private counterparts in Eq. (2). We obtain the social optimum by maximizing $U_T$ with respect to the $X_{Ii}$ and $V_i$ subject to (2) through (4), setting $P_X^c = P_X$ and t = 0:[5]

$$P_X = (\partial U / \partial X_I)_i / (\partial U / \partial V)_i + (\partial z / \partial X_I)_i \overline{Y}_i \tag{12}$$

The concavity assumptions of the utility and production functions ensure that the second-order conditions for maximum are satisfied.

Equation (12) reflects the fact that health is a joint input: it increases productivity and thereby income (Eqs. 3 and 4) in addition to generating utility directly. Hence only the difference between the marginal social cost $P_X$ and the marginal social product generated $[(\partial z / \partial X_I)_i \overline{Y}_i]$ is the cost attributable to private consumption. Thus it is normally optimal to induce the individuals to acquire private health past the point where its marginal utility equals its marginal social cost, and conversely, past the point where its marginal social product equals its marginal social cost. We have:

Proposition 1. In the social optimum, the marginal social cost of health equals the sum of the marginal social product and the marginal utility of health.

The policy-maker's optimization is illustrated in Figure 1. There, the lower declining curve depicts the marginal social product of the pure investment aspect of $X_I$ for the patient, given $X_0 + X_P$. The higher curve depicts the sum of this marginal product and the marginal utility of the private consumption aspect of $X_I$ relative to ∂U/∂V (Eq. (12)), reflecting diminishing returns and diminishing marginal utility, respectively, of health. The horizontal curve is the marginal social cost of health $P_X$, constant in an individual patient's case, so that the total cost of $X_I$ is the area under this curve.

[Figure 1 about here]

The conventional optimal amount of health is $X_{I1}$, where the marginal social cost and the marginal social product of health are equal (see e.g. Pauly (1979)). However, it leaves the net consumer surplus of ABC unutilized so that it cannot be the social optimum in a model where health generates utility. This optimum is $X_{I2}$, where the sum of the marginal product and marginal utility equals the marginal social cost.

Examining the Figure in greater detail, the social cost of health is fully covered by its investment aspect up to $X_{I1}$. The area $X_{I1}ACX_{I2}$ represents the marginal social cost of the extra health $(X_{I2} - X_{I1})$. It would generate extra output by $X_{I1}ADX_{I2}$ and extra utility by ABCD. The shaded triangle ACD is the part of the marginal social cost of the extra health that is not covered by the increase in output that it made possible. This is the private consumption health part of the marginal cost of the increment.

It is easy to see that past $X_{I2}$, the social marginal cost of health exceeds the sum of its marginal utility and marginal product. Thus if health were a free good and there were no rationing, the patient would choose $X_{I3}$, where the sum of the marginal utility and marginal product is zero. This would result in a loss represented by the area between the $P_X$ curve and the $CX_{I3}$ curve between $X_{I2}$ and $X_{I3}$ (not shown).[6]

**b. The Policy**

What should the social planner do to induce the patient to choose $X_{I2}$, given that the patient faces the private costs and prices, as specified in Eq. (2)? The patient's optimum condition is:

$$P_X^c = (\partial U / \partial X_I)_i / (\partial U / \partial V)_i + (1-t)(\partial z / \partial X_I)_i \overline{Y}_i \qquad (13)$$

This differs from the social optimum in two respects, given that we took care of externalities with $X_P$.[5] The market price of health may deviate from its marginal social cost by the taxes or subsidies on health services. (We have not modeled indirect taxes or subsidies explicitly for simplicity, as their effects are obvious.) Secondly, income taxes drive a wedge between the marginal social product of health and the marginal disposable income it generates for the individual. The social planner should now optimally determine the tax and subsidy statuses of the two kinds of health.

To determine optimal pricing, we obtain from Eqs. (12) and (13):

$$P_X - P_X^c = t(\partial z / \partial X_I)\overline{Y}, \qquad (14)$$

where $P_X$ - $P_X^c$ is the optimal subsidy per unit given to health.

This yields the following rules for the determination of the optimal pricing or tax status of health. For the private consumption part of health, or the area between the declining curves, $(\partial z / \partial X_I = 0)$ holds, which makes the right hand side of Eq. (14) zero. Therefore,

<u>Proposition 2. It is optimal to charge the private consumption health its full marginal social cost ($P_X = P_X^c$), impose on it the same consumption taxes as those on other goods, and keep its expenses non-deductible from taxable income</u>, since otherwise pricing would cause the ratio of the marginal utilities of

the two pure consumption goods to be different from their relative marginal social cost of production (Eqs. (11) and (12)).[7]

Proposition 3. It is optimal to subsidize the investment part of health ($\partial U / \partial X_I = 0$ in Eqs. (12) and (13)) by the tax on the marginal income generated by this extra health. Alternatively, the expenses for this part, but not those for the private consumption part, could be made deductible from taxable income, while pricing it at its marginal social cost. The latter alternative is more practical, since marginal income tax rates are typically different across individuals. It would produce the equivalent of t = 0 in Eq. (14) making it: $P_X - P_X^c = 0$. The investment part would now be "self-financing". If e.g. the subsidy were smaller, the marginal user cost of health would be higher, which would lead the individual to purchase less investment health than would be optimal in Eq. (12). Note, however, that this optimality holds if the patient is literally paid his marginal product and thus gets no pay when out of work. We will shortly return to the case of sick pay.

In terms of the Figure, the social planner should offer the patient $X_{I2}$ so that the cost of $X_{I1}$ is fully deductible, since it is fully covered by the investment health part. In addition, the part $X_{I1}ADX_{I2}$, of the cost of $X_{I2} - X_{I1}$ should be deductible. The remaining part of the cost ACD is not covered by additional output, and it should be non-deductible and subject to general consumption taxes as private consumption. Now, the rational patient would choose $X_{I2}$, ending up with the net consumer surplus of ABC for $X_{I2} - X_{I1}$. This is of course independent of whether the health is produced by the government or by the private sector.

In other words, it is optimal to provide $X_{I2}$ only if the patient pays the full marginal social cost of the increase in private consumption health of $X_{I2} - X_{I1}$, - but only this - out of his after-tax income in the same way as that of other private consumption goods, in addition to the deductible cost of the investment health $P_X X_{I1}$ plus $X_{I1}ADX_{I2}$.

Alternatively, the social planner could make an all-or-nothing offer of $X_{I2}$ outright, when the price should be lower than the area under the higher declining curve down to $X_{I2}$, adjusted for the fact that the area under the lower declining curve should be deductible from taxable income. The trouble with this alternative is that not only is it arbitrary, but the social planner does not know the patient' s utility function and therefore cannot determine $X_{I2}$. The social planner can even have the doctor -- his agent -- offer the patient several alternatives with assessed effects and costs, as will be demonstrated below. As will be shown, if the above optimal pricing and tax statuses are used, the social planner needs to know only the marginal costs and the marginal products, or $P_X$, $\partial z / \partial X_I$, and $\bar{Y}$, but not the patient's preferences.

In reality, there are often institutional arrangements making both the user cost of care and/or the marginal gain in disposable income zero or considerably smaller than $P_X$ or $(\partial z / \partial X_I)\bar{Y}$, respectively. While affecting the private variables, they do not affect their social counterparts and the social optimum, unless they call for more care than that optimum. A key redistributive policy instrument in this model is the public health good $X_P$ along with taxes and subsidies. If the public health good $X_P$ is provided free of charge to everyone, the above principles apply to health in excess of $X_0 + X_P$, as stated above. For instance, let $X_P$ increase to $X_{I1}$. The patient would now choose $X_{I2}$ if offered a deductible $X_{I1}ADX_{I2}$ and a non-deductible ACD. In the public health good, the patient would thus receive an additional net income transfer - or social insurance benefit - of $(1-t)P_X \Delta X_P$ (Eq. (14)). Obviously the same applies to any health with a zero user cost e.g. due to health insurance.

It is of course possible that $X_P$ exceeds $X_{I2}$. Then the motivations behind the provision of public health or health insurance call for more health than would be justified by its marginal product and private marginal utility, the efficiency criteria of this model. An example is Medicare. In either case, it is always optimal to provide at least $X_{I2}$ irrespective of the other objectives of health care policy, since this amount maximizes the welfare of society.

The case with a user cost at a fraction of $P_X$ is a straightforward extension of the above analysis, in lowering $P_X^c$ without affecting the $P_X$ and the other curves. Then rationing is necessary if the policy-maker wants the patient to choose the amount $X_{I2}$ of health.[8] We will later return to this issue.

Secondly, if the patient has a fully paid sick leave, his marginal disposable income from care would be zero. Then also the user cost of investment health would have to be zero, because he would now optimize by equating $P_X^c = (\partial U / \partial X_I)/(\partial U / \partial V)$, in Eq. (13). The patient would then choose $X_{I2}$ if offered the care exceeding $X_{I1}$ for a non-deductible ACD. Alternatively, he could be made an all-or-nothing offer: the therapy or the sick pay could be made conditional on his choosing the optimum prescribed amount, when he could be charged a non-deductible amount slightly smaller than the area between the declining curves from the point $X_0 + X_P$ down to $X_{I2}$. However, this would require knowledge of his utility function. Currently, one sees a variety of systems in health insurance. Obviously, physicians currently use completely different criteria in determining therapy and the length of the sick leave.

In determining optimal therapy, the physician needs to proceed in the same basic way as currently. He already needs to assess for the patient the expected costs of alternative kinds of care $(X_I P_X \text{ or } X_I P_X^c)$ and their clinical effectiveness (dz), given the symptoms, on the basis of which $\partial z / \partial X_I$ can be calculated. In fact, the patient performs this calculation at least implicitly in deciding whether he wants to spend more for a more effective or more pleasant treatment. The problem is made easier by the fact that there are often relatively few conceivable therapies, which makes the z-function discontinuous in many places. The physician can ask the patient about $\bar{Y}$, and in many countries the patient's answer can be verified from public sources. But since the private consumption part is determined residually, no knowledge of the utility function is needed. Yet the practical difficulties with these assessments should not be underestimated.[9] Of course, our objective function is different from that currently used, if any.

**c. A Numerical Example**

First assume that the user cost of care equals its marginal social cost ($P_X^c = P_X$) and the patient gets no pay on sick leave. Suppose the patient earns $10 000 a month and the social discount rate is 12 % per annum, or 1 % per month. Let the risks associated with different therapies be equal for simplicity. Without therapy, the patient gets to working condition (z ≈ 1) in six months. The first therapy gets the patient to working condition in four months. The present values of the patient' s marginal products in his normal state of health in the fifth and sixth months are roughly \$9 500, and \$9 400, respectively. Thus the marginal product of the therapy over no care $\left((\partial z / \partial X_I)\overline{Y}\right)$ is the present value of his income for the fifth and sixth months, or \$18 900. A therapy whose marginal social cost equals this amount on a present value basis brings the patient to point A at $X_{I1}$ in the Figure, and its cost is the fully deductible charge.

If another therapy healing the patient in four months costs \$25 000, it is optimal to provide it if the patient pays the extra, or an additional \$6 100 out of his after-tax income to get to a point between A and B in the Figure: if he is indifferent at the margin, he is at B. The marginal product remains unchanged so that the difference is private consumption.

A third therapy restores the patient to working condition in three months. Its marginal product over the second therapy is the present value of the patient's marginal product for the fourth month, or \$9 600. (If this brings the patient to the optimum, it equals $X_{I1}ADX_{I2}$ in the Figure.) If the therapy costs \$35 000, its marginal cost is \$10 000 over the second therapy so that the difference of \$400 is the increase in private consumption health (roughly ACD). This brings the total non-deductible charge to \$6 500 (= \$6 100 + \$400) which is to be paid in addition to the deductible charge of \$28 500 (= \$18 900 + \$9 600). The patient is now between points D and C; if he is indifferent at the margin, he is at C. On a crude level, it is a choice of these sums that the doctor or the social administrator should offer the patient along with a description of the therapies.

The optimal amounts of care remain the same if the user cost of the therapies is smaller than their marginal social costs. The doctor can compute the investment part of health based on the marginal social cost as before. However, if also private consumption health is subsidized , the patient has an incentive to purchase more of it than would be optimal, as seen from Eqs. (12) and (13). Therefore the doctor should be careful not to offer him alternatives that could be regarded as being beyond $X_{I2}$ by some commonly accepted standard. An obvious solution to the problem would be to make the user cost of private consumption health equal to its marginal social cost.

If the free public health good $X_P$ (or care fully covered by insurance) is the care costing \$18 900 it would of course be provided free of charge. It would be optimal to provide the $25 000 care for a non-deductible $6 100, and the $35 000 care for a non-deductible $6 500 plus a deductible $9 600.

If the patient has a sick leave with full pay, he should be offered the investment health for free: the $18 900 care for free, the $25 000 care for a non-deductible $6 100 and the $35 000 care for a non-deductible $6 500.

Correspondingly, for an otherwise identical patient with an income of $5 000 a month, the value of the marginal product of the 4-month care over no care is $9 450. So it is optimal to provide the $18 900 therapy for the deductible amount of $9 450, the remaining $9 450 being non-deductible. The deductible charge for the $25 000 care would be the same, $15 550 (= $9 450 + $6 100) being non-deductible. As seen, if private consumption health is priced at its marginal social cost, the doctor needs to know only $\partial z / \partial X_I$, $\overline{Y}$ and $P_X$, but not the patient' s utility function, and he can differentiate the prices and tax status optimally without explicitly having to bother about consumption and investment health. That is, investment health is assessed indirectly by its marginal product.

Again, if the $18 900 care is the free public health good, it is provided free of charge. The $25 000 care would be provided for a non-deductible $6 100. The same amounts apply if the patient has sick leave with full pay.

The number of possible cases seems to be large at first sight. However, many of them pertain to a regime which rules out other regimes. The number declines considerably if all the patients are under a common regime: if the public health good or health insurance coverage, and the sick pay arrangements, are the same for all citizens.

**d. Implications**

Equation (12) implies:

$$\left(\frac{\partial z}{\partial X_I}\right)_i \overline{Y}_i + \left(\frac{\partial U}{\partial X_I} / \frac{\partial U}{\partial V}\right)_i = \left(\frac{\partial z}{\partial X_I}\right)_j \overline{Y}_j + \left(\frac{\partial U}{\partial X_I} / \frac{\partial U}{\partial V}\right)_j = P_X \qquad (15)$$

Thus, abstracting from the private consumption health part for the time being, the optimal marginal gains from health in the patients' marginal products are equal across patients, and equal the marginal social cost of health. The same holds of course also for different kinds of treatment (e.g. drugs and hospital care if their effects are independent) for any patient.

Proposition 4. It is optimal to give a patient with a high expected marginal therapeutic gain from care $[(\partial z/\partial X_I)(\partial X_I/\partial H)]$ more care than to a patient whose marginal gain is lower for equal incomes. Likewise, the higher the patient' s labor income in his normal state of health, the more it is optimal to spend on more efficient (and/or safer) therapy, given the expected marginal therapeutic gains from different kinds of care. This can be read off Equation (15), given diminishing returns on health and health care (Eq. (4) and p. 7, respectively). Spending more resources to get a high labor income earner back to the labor force more quickly (and/or more safely) relinquishes more resources for society than it uses up, except at the margin.[10] In addition, if the patient is willing to pay the full marginal social cost of the private consumption part of any additional care out of his after-tax income, and that of the investment part out of his before-tax income, it is optimal to provide it for the same reason.

The result on the optimal amount of care is at variance with the critique of the "tax subsidy" to high labor income earners. (See e.g. Feldstein (1981) Enthoven (1984) and others.) A greater tax subsidy given to high labor income earners in terms of the deductibility of health expenses is consistent with efficiency as long as the expenses are for pure investment health, since if labor income is taxed, expenditures to bring this income to normal should be deductible for taxes to be neutral. The potential output so generated is greater than the resources used up by such care, except at the margin. However, unlimited deductibility amounts to a subsidy to private consumption health, whose demand is also a function of income (see Eqs. (1), (2), and (13)) and the critique is justified on this part.

Proposition 5. In rationing, it is optimal to rank the patients according to their marginal products from care, and give each patient treatment down to the point where the marginal product equals the marginal social cost of care. Requests for additional treatment are to be treated as suggested regarding $X_{I2} - X_{I1}$. This can also be read off Equation (15). Thus it is optimal to admit a patient who was rationed on the above criterion if he is willing to pay large enough a sum to get "above the line". By buying himself a hospital bed he trades other goods for rationed health with the rest of society, relinquishing more resources for society to consume than the net gain in the marginal product of the marginal patient rationed in his place. If the marginal product of a certain rationed care to a given patient is \$18 900, while the corresponding marginal product of the marginal patient is \$24 000, it is optimal to admit the patient in place of the marginal patient if he is willing to pay a non-deductible \$5 101 in addition to the deductible user cost of the \$18 900 care, which will bring the total marginal gain to \$24 001.

The above result is at variance with Smallwood and Smith (1976), who recommend the ranking of patients according to the severity of symptoms, i.e. the initial losses of health rather than the potential marginal products of care. That would lead to a ranking radically different from ours in e.g. putting hopeless cases at the head of the line. In our model, care to hopeless cases has a zero marginal product and is not justified by investment criteria but is consumption.[11]

As proposed, there is always rationing in optimum when health care is priced below its marginal social cost, since it is not optimal to provide the market-clearing amount of care at this price. However, if capacity were to fall even below its optimal level, the shadow price of health would be $P_X + \lambda_2$, where $\lambda_2$ is the Lagrangian of the capacity constraint, as the reader can verify. We have:

Corollary 1. A partial relaxation of the capacity constraint (increase in hospital capacity) leads to treating patients with marginal products from care lower than before if the ranking is done optimally, adjusted for the patients' incomes. Assuming that the amounts of private consumption health purchased are zero or randomly distributed, with optimal ranking, the new entrants are patients with a lower $\partial z / \partial X_I$, i.e. with conditions medically minor or less susceptible to care than those treated before. The empirical finding to this effect is thus also consistent with optimality and does not necessarily imply demand-shifting, which appears to be the dominating interpretation.[12]

As proposed earlier,

$$Y_i = z_i \bar{Y}_i = \sum^{\tau} z_{it} \bar{y}_{it} \, , \qquad (16)$$

t=0

where $\bar{y}_{it} \equiv E(y_{it})e^{-rt}$,

and $z_{it}$ is the expected value of $z_i$ in period t, $E(y_{it})$ that of patient i' s expected marginal product in his normal state of health in period t, $\tau$ his remaining time in the labor force, and r the marginal social rate of time preference. The interested reader can verify by obvious steps that, given the timing of the care

- the greater the marginal social rate of time preference, and
- the more remote the patient' s expected high income years, the less it is optimal to spend on the patient' s care.[13]

## 3. Conclusions and Extensions

We have derived criteria for medical care decisions beyond the public health standard, consistent with an efficient allocation of resources in a model where health is a special case of a good produced, consumed, and used as an input in production, and for the tax and subsidy statuses that the social planner has to stipulate for different kinds of care so as to induce patients to choose the socially optimal therapies.

It was shown to be optimal to provide health care beyond the conventional optimum point where the marginal product generated by the care equals its marginal social cost, to the point where the sum of the marginal product and the marginal utility equals that cost, but only if the patient is willing to pay the full marginal social cost of the part of the extra that exceeds the marginal product, out of his after-tax income. It is optimal to make the cost of all investment care deductible from taxable income. However, if the patient has fully paid sick leave, investment care should be free for the patient to choose the optimal amount of care. Thus the optimum price of investment health and sick pay are connected via the patient’s optimization.

The higher the patient' s labor income, and the greater the expected marginal therapeutic gain from care, the more it is optimal to spend on more efficient (and/or safer) therapy. Therefore it is indeed optimal to give a greater “tax subsidy” to high labor income earners by making investment health expenses, but not private consumption health expenses, deductible from taxable income. The greater the marginal social rate of time preference, or the more remote the patient' s entry in the labor force, the less it is optimal to spend on the patient' s care. Care to hopeless cases, beyond the public health standard, is not justified by investment criteria, which makes it private consumption.

If the user cost of care is kept below its marginal social cost, rationing is necessary to induce patients to choose the optimal amount of care. In rationing, it is optimal to rank the patients according to their marginal products from care, and give each patient treatment down to the point where the marginal product equals the marginal social cost of care. Requests for additional therapy should be treated as private consumption on the part that exceeds the marginal product of the additional therapy. Likewise, it is optimal to admit a patient who was rationed on the above criterion if he is willing to pay enough to get above the line.

Increase in hospital capacity leads to treating patients with cases minor to those treated before, incomes equal, if rationing is done optimally, and is not necessarily a sign of demand-shifting.

It is optimal to use these criteria for care decisions irrespective of the other objectives of health care policy, since they bring society to its production possibility frontier, maximizing its consumption possibilities. Let us return to our example from Britain, where the National Health Service restricts access to dialysis of patients over 55 years of age (see Weisbrod (1991)). Let the care up to this age be the public health good. Our approach would suggest administering care thereafter up to the point where the marginal product of the care equals the marginal social cost, the user cost, if any, being deductible from taxable income. It is optimal to provide more if the patient is willing to pay for the private consumption part of the extra care out of his after-tax income, and the investment part out of his taxable income. On retirement, rather than at 55, the entire cost would thus become non-deductible.

It is optimal for the employer to pay the more of the worker’s health care bill, the lower the user cost of health, the more efficient care is in shortening the sick leave, the greater the substitute’s wage cost and the patient’s unit sick pay costs, and the smaller the patient’s wage cost to the employer, all including payroll taxes. Depending on these factors the employer may be disinterested in spending anything on the patient's care, or bring the patient's health past the social optimum.

The above propositions have several implications. Take, for instance, the observed correlation between the number of operations and the number of surgeons, both per capita, which has been interpreted as evidence in support of demand shifting. (See e.g. Fuchs, (1978), p. 46, whereas Pauly, (1980), p. 107, fails to get a significant parameter). The above decision criteria suggest first that since health is a normal good, the consumption demand for health care is positively correlated with (total household) income (Equations 1 and 2). Second, it is optimal to give more expensive therapy to patients with high labor incomes if it brings them more quickly (and more safely) to working condition (Proposition 1): surgery is often quicker and would dominate in marginal cases, were it not for cost considerations. Third, it is rational for surgeons - and economists - to locate in areas where the demand for their services is high - i.e. high income areas. And fourth, since the $\partial X/\partial H$ of physicians are a function of their training and experience, it is optimal for a surgeon to exercise his comparative advantage by using surgery in marginal cases where others would not have used it. Therefore the correlation between surgery rates and the supply of surgeons also finds an explanation in their causal link with patient income.[14] Since most would agree that this mechanism is there, any test for SID needs to control for it. The qualitative results stand in both cases also if doctors use the user cost, rather than the marginal social cost as the shadow price in the care decision, as long as the two costs are positively related.[15]

We believe that it is important to examine health care issues in a broader context of general resource allocation, explicitly allowing for the three aspects of health. It has applications for a number of problems in health care policy. We also believe that as society's ability to offer ever more costly new therapies as a free service becomes more and more limited, decision-makers would benefit from the present approach when making the hard choices of the future.

**Appendix**

Could the patient' s employer have an incentive to pay part of the bill to induce the patient to choose more effective and more costly care and return to work sooner? The employer maximizes profits, as in Eqs. (5) and (7), which implies minimization of the cost of the sick leave (W). Let us assume for simplicity that the employer does not find it optimal to adjust output because of a worker' s sick leave e.g. because of adjustment costs. Each period the cost (W) consists of the substitute' s (or alternative arrangement's) productivity-adjusted wages (b) plus the patient' s sick pay (s), minus the wages that would have been paid to the patient (w) had he been at work producing the output now produced by the substitute. The W is the sum of the present values of the period costs for the length of the sick leave (T), plus the amount $\left(X_{IE} P_X^c\right)$ that the employer pays for more intensive care to shorten the length of the sick leave:

$$W = (b + s + w)\int_0^T e^{-rt}\,dt + X_{IE} P_X^c \tag{17}$$

$$T = T\left(\underset{-}{X}_{IE}\right), \tag{18}$$

where $X_{IE}$ is the amount of health purchased by the employer and r the interest rate. In Eq. (18), the length of the sick leave is a declining function of $X_{IE}$. Diminishing returns make $\partial^2 T / \partial X_{IE}^2$ positive.

<u>Proposition 6. It is optimal for the employer to purchase the more health care to get the patient back to work sooner, the more effective the care and the lower its user cost, the greater the substitute's cost and the patient's unit sick pay cost, the smaller the patient's unit wage cost and the interest rate, and the shorter the sick leave before the employer's intervention.</u> Minimizing Eq. (17) with respect to $X_{IE}$, subject to Eq. (18) yields:

$$P_X^c = -(\partial T / \partial X_{IE})(b + s - w)e^{-rT} \tag{19}$$

The second-order condition for minimum is satisfied by the assumption of diminishing returns on $X_{IE}$ with respect to T. Thus it is optimal for the employer to purchase more health care to get the patient back to work, the lower $P_X^c$, the more effective care is in shortening the sick leave, and the greater the substitute's wage cost and the patient's unit sick pay costs and the smaller the patient's unit wage cost, all including payroll taxes, and the smaller the interest rate, and the shorter the sick leave before the employer's intervention.

In our model with a homogeneous labor force the cost and productivity of the substitute are equal to those of the patient in his normal state of health, i.e. b = w. If the patient gets no sick pay, s = 0. The right-hand side of Eq. (19) is thus zero, and the employer hires a substitute from the market for the duration of the patient's sick leave at the same cost as the patient's wage cost. Thus the employer has no incentive to pay anything to accelerate the patient's recovery.

Secondly, if the patient has a sick leave with full pay, s = w, and Eq.(19) reads $P_X^c = -(\partial T / \partial X_{IE}) b / e^{-rT}$, which is equal to $(\partial z / \partial X_I)\overline{Y}$ in the absence of payroll taxes, with z defined with respect to time: quicker recovery saves the employer the substitute's wage cost. Thus it is optimal for the employer to pay an amount making the value of the marginal product of health equal to the marginal user cost of the health. If $P_X^c = P_X$, this would bring the patient to point $X_{I1}$ in the Figure in the absence of payroll taxes, and beyond with them. If $P_X^c < P_X$, it is optimal for the employer to pay for more care. In our model, sick pay and the employer's payroll taxes are the reasons why it can be optimal for the employer to pay part of the cost to induce the patient to buy more health.

Next combine the employer's incentives with those of the employee, assuming that health is priced and taxed optimally. We obtain from Eqs. (13) and (19):

$$-\frac{\partial T}{\partial X_{IE}}(b+s-w)e^{-rT} = \frac{\partial U / \partial X_I}{\partial U / \partial V} + (1-t)\frac{\partial z}{\partial X_I}\overline{Y} = P_X^c . \qquad (20)$$

The left-hand side of Eq. (20) equals $(\partial z / \partial X_I)\overline{Y}$ if e.g. (b = s = w) and there are no payroll taxes, as shown above. With full sick pay, the patient does not have an incentive to pay for investment health. Since health care is priced optimally, $P_X^c = 0$ for investment health. Returning to the numerical example, the first patient gets the $18 900 care for free, and would get the $35 000 care for a non-deductible \$6 500. If the patient initially chose the first care, the employer would have to pay up to this amount to induce the patient to choose the third therapy, as the patient is presumably willing to pay something to be well for an extra month $(\partial U / \partial X_I > 0)$. Since the present value of the patient's income (and sick pay) for the fourth month is \$9 600, it would be optimal for the employer to pay the \$6 500 even in the absence of payroll taxes to get the worker back to work sooner. It is of interest to note that in this situation the employer pays for private consumption health in order to get the benefits from investment health.

Next examine the second patient's situation (income \$5 000 per month). The present value of the marginal product of the four-month care over no care is \$9 450. With full sick pay this is optimally free of charge so that it is optimal to provide the \$18 900 therapy for a non-deductible \$9 450. The employer would be indifferent between paying this amount to get the patient to work in four months on the one hand, and hiring a substitute, on the other, in the absence of payroll taxes, since the employer would be saving an equal amount in sick pay. If there are payroll taxes, it would be optimal for it to choose the former alternative.

The marginal cost of the \$35 000 care over the first care is \$16 100(= \$35 000 - \$18 900), while the present value of the fourth month's income is \$4 800, leaving the non-deductible net marginal cost of \$11 300. If the patient is willing to pay nothing for the more effective care, the employer would have to pay this amount extra to get the worker back to work after three months. The payroll taxes would have to be over 135 per cent of the extra wages of \$4 800 to make the employer interested in paying for this extra care. Obviously, the game-theoretic issues opened by this situation are beyond the scope of this paper.

The above examples were cases of the employer paying conditional on the patient choosing a particular therapy. However, considerations of equity between employees may cause the employer to opt for a standard rule, which may make the optimal value of T a function of the terms of the employer's offer if it affects the patient's marginal conditions. An unconditional lumpsum contribution fails to affect the marginal conditions, which makes dT zero: the patient chooses the same amount of health as originally and keeps the contribution. An offer to pay a given percentage of the bill is equivalent to a subsidy in Eqs. (13) and (20), which lowers the patient's $P_X^c$. Applying the Figure to the patient's situation, the price curve shifts down by the amount of the subsidy, which increases the quantity of health purchased by the patient. The employer could also offer to pay part of the extra health in excess of a given level, where the value of investment health is calculated as in the numerical example, or make an all-or-nothing offer conditional on the patient choosing a given therapy. Of course, the terms of sick pay are generally determined by a standard rule.

There are numerous combinations of partial sick pay, partial tax deductibility of the patient’s investment health expenses and payroll taxes that can make the employer interested in paying part of the bill. However, these examples suffice to show how the employer’s incentives are determined and how they can affect the patient.

In reality, substitute labor is often less productive than encumbent workers, which increases b and makes greater spending on the patient' s health optimal. Substitutes may also need training. The training costs are normally a sunk cost, which do not affect the marginal conditions. They may, however, affect the global optimum by making overtime work by the encumbent labor force the dominating alternative - the more so the shorter the sick leave. In that case the unit cost (b) is overtime pay. It may thus be optimal for the employer to bring the patient past the social optimum $X_{I2}$. On the other hand, it is also possible that b < w, as when the patient is an old unproductive senior worker in a simple job. If, besides, s = 0, the employer can actually make money on the worker' s sick leave and has no incentive to accelerate his recovery.

In summary, in our model with a homogenous labor force, sick pay, payroll taxes, and substitutes’ higher costs can provide the patient’s employer with an incentive to pay part of the bill, which may lead to the patient getting more care than the social optimum.

Would the government have an interest in sharing in the patient' s hospital bill? If society' s objective is maximization of societal welfare as postulated in Eq.(1), it would be optimal for society to make sure, one way or another, that the patient gets $X_{I2}$ in the ways discussed above. In reality, an impediment to this may be that some patients are liquidity-constrained. Then the availability of financing for medical costs is an essential condition for the optimum of these patients, as argued by Feldstein (1981). The possibility of making $X_P = X_{I2}$ and the implications it has for the optimal pricing and tax status of care has been discussed above. It is becoming increasingly costly in addition to being politically questionable on equity grounds. Of course, discussion of the implications of other policy objectives is beyond the scope of this paper.

**Notes**

*This paper was written while the author was a Visitor at Northwestern University. He wishes to thank Bernard Friedman, Mark Pauly, Uwe Reinhardt, and Mark Satterthwaite for useful comments. Financial support from the Center for Health Services and Policy Research at Northwestern University, the Yrjö Jahnsson Foundation, and the Academy of Finland is gratefully acknowledged.

1. Accordingly, in Britain, the National Health Service restricts access to dialysis of patients over 55 years of age (see Weisbrod (1991)). In the United States, discussion on the Patient's Bill of Rights continues, a quarter of the population being without health insurance protection. Yet the systematization effort of the Oregon plan has widely had a hostile reception.

2. Optimality requires that the public health system should reimburse a patient receiving care belonging to the public health standard from a private provider by its cost at a public hospital.

3. Of course, the consumption of a household being a function of total household income, also the demand for consumption care of its individual members is a function of that income.

4. Of course, only values above the minimum level where the patient is fit for the labor force count in the calculation. Thus the z function intercepts the X axis at this minimum.

5. Naturally, the marginal social product and marginal social utility include in principle any externalities that $X_{Ii}$ may have, positive externalities increasing the values of x and U above their private counterparts. As is well known, they constitute a case for subsidizing $X_I$ by the value of the externality per unit of $X_I$, leading to the same optimum as with the private variables. We have refrained from complicating the model at this point and let externalities be explicitly taken care of by $X_P$. There is of course a one-to-one mapping between the social and private marginal product and marginal utility of $X_I$.

6. For a discussion on whether it is meaningful to even have a demand curve for health care, see e.g. Feldman et al. (1993) and Dranove (1995).

7. Eq. (13) implies

$$\left(\frac{\partial U}{\partial X_I}\right)_i\left(\frac{\partial X_I}{\partial H}\right)_i/\left(\frac{\partial U}{\partial V}\right)_i = P_H - (1-t)\left(\frac{\partial z}{\partial X_I}\right)_i\left(\frac{\partial X_I}{\partial H}\right)_i\overline{Y_i}$$

8. The system that is being introduced in Finland is consistent with the above principles regarding the status of private consumption health, but not the criteria for investment health. In terms of our terminology, the public health good is determined by clinical criteria and is provided by health clinics and public hospitals. The patient is given a voucher for referral care not provided by the clinic. It entitles to essentially free care at a public hospital, where queuing is part of the package for non-acute cases. Alternatively, it is good for (often partial) payment at a private hospital at the price that the public hospital charges the health clinic. The non-deductible extra at a private hospital is thus regarded as private consumption health, involving such consumption goods as prompt and individual attention, choice of the physician, and a cognac after surgery, but also the option of therapy more costly than that prescribed.

A corresponding practice could be extended to prescription drugs: The prescription could specify a generic product qualifying for health insurance benefits. If the patient wants a branded product, he could be responsible for the difference in price on a non-deductible basis. The national health insurance system also provides full catastrophic insurance for annual drug and hospital costs in excess of $500.

9. One may have to base these calculations on averages in the experimental population. However, physicians already give patients cost and duration forecasts of alternative therapies. The moral hazard problems of doctors can be alleviated by monitoring physician assessments of the lengths of care against their realizations. Analogous standards are already used in hospital and even physician performance measurement.

10. An implication is that it is optimal to have different kinds of health insurance, and the premia of first class coverage should be deductible for high labor income earners on the part regarded as investment health. This is problematic from several points of view, which may be a reason why many companies have purchased health insurance for their top managements.

11. It follows that in cases where $\partial X_I / \partial H$ is practically zero, as in advanced cases of terminal cancer, AIDS, or "obviously" fatal injuries, care is not justified by investment criteria, which makes even the alleviation of pain consumption. If it is called for by legislation, court verdicts, or other societal norms, a case could be made for regarding it as free publicly provided consumption, when it would be financed mainly from public sources. The rest is private consumption. This has important implications for when "the plug should be pulled", who is to decide on it at different stages, and who should pay. This is a corollary of the above optimum condition. Of course, care to retirees is likewise consumption.

12. Physician surveys in the United States have suggested that there has seldom been rationing of hospital care to patients. However, the hospitals that have added capacity are likely to be in this minority: investments are made in response to demand so that hospitals expand only if they have actual or expected capacity shortages. It is also rational to tighten criteria for hospital care and use substitutes when there is a shortage of capacity (i.e. prior to expansion), and relax them when the shortage is over (i.e. after expansion).

13. It is seen that preventive care is a special case of investment (and consumption) care, to which the present analysis applies. It increases the expected present value of the worker' s productivity in the future by affecting his probability of getting sick, given his exposure to agents causing illness.

14. To get around the identification problem, Fuchs explains surgeon supply with per capita hotel receipts of the area and finds that physicians tend to settle in attractive areas. He then explains demand with the predicted surgeon supply and per capita income. However, attractive areas are generally also high income areas. Thus both surgeon supply and the demand for surgery are highly correlated with income, apparently rendering the demand-shifting correlation spurious. This is also supported by another symptom to that effect: income

becomes significant only when surgeon supply is omitted as an explanatory variable.

15. What would constitute evidence of demand-shifting? If the physician uses more costly care that is widely recognized as being clinically weakly dominated by less costly (or no) care (i.e. care having no higher $\Sigma\Delta E(z_t)/(1+r)^t$) for equal risk for any specialty of doctors, given the set of symptoms, he is either incompetent or shifting demand (c.f. Pauly (1994)). Of course, even here the evaluation must be based on the information available at the time of the care decision and not on the information that care produces, as has often been done (c.f. Reinhardt (1983, 1985)). This makes demand shifting very hard to prove. It becomes even harder when risk is taken into account: as long as the degree of societal risk aversion is not explicitly stipulated, almost any care that reduces risk can be justified.

**References**

Arrow, Kenneth J. (1963). "Uncertainty and the Welfare Economics of Medical Care." American Economic Review, December.

Dranove, D. (1995). “A Problem with Consumer Surplus Measures of the Cost of Practice Variations.” Journal of Health Economics, Vol. 14, No. 2, June, 243-251.

Enthoven, Alain. (1984). Health Plan: The Only Practical Solution to the Soaring Cost of Medical Care. Reading, MA: Addison-Wesley.

Feldman, R., and Dowd, B. (1993). “What Does the Demand Curve for Medical Care Measure?.” Journal of Health Economics, Vol 12, No. 2, July, 193-200.

Feldstein, Martin. (1981). Hospital Costs and Health Insurance. Cambridge, MA: Harvard University Press.

Friedman, Bernard. (1978). "On the Rationing of Health Services and Resource Availability." The Journal of Human Resources, Vol. XIII, Supplement, 57-76.

Fuchs, Victor R. (1978). "The Supply of Surgeons and the Demand for Operations." The Journal of Human Resources, Vol. XIII, Supplement, 35-56.

Grossman, Michael. (1972a). "The Demand for Health: A Theoretical and Empirical Investigation." New York: The National Bureau of Economic Research.

-------- . (1972b). "On the Concept of Health Capital and the Demand for Health." Journal of Political Economy, March.

Ng, Yew-Kwang. (1975). "Bentham or Bergson? Finite Sensibility, Utility Functions, and Social Welfare Functions." Review of Economic Studies, 42(4), October, 545-569.

-------- . (1984). "Quasi-Pareto Social Improvements." American Economic Review, December, 74,1033-1050.

Pauly, Mark V. (1979). "What is Unnecessary Surgery?." Millbank Memorial Quarterly/Health and Society, Vol. 57, No. 1, 95-117.

-------- . (1980). Doctors and Their Workshops: Economic Models of Physician Behavior. Chicago: University of Chicago Press.

-------- . (1994). "Editorial: A Re-Examination of the Meaning and Importance of Supplier-Induced Demand." Journal of Health Economics 13, 369-372.

Reinhardt, Uwe. (1983). "The Theory of Physician-Induced Demand and its Implication for Public Policy." Beiträge zur Gesundheitsökonomie, Band 4.

-------- . (1895). "The Theory of Physician-Induced Demand: Reflections after a Decade." Journal of Health Economics, June.

Smallwood, Dennis E. and Kenneth R. Smith. (1976). "Optimal Treatment Decisions and the Allocation of Medical Resources." Manuscript, Evanston, Illinois.

Weisbrod, Burton. (1991). “The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment.” Journal of Economic Literature, June, 523-552.

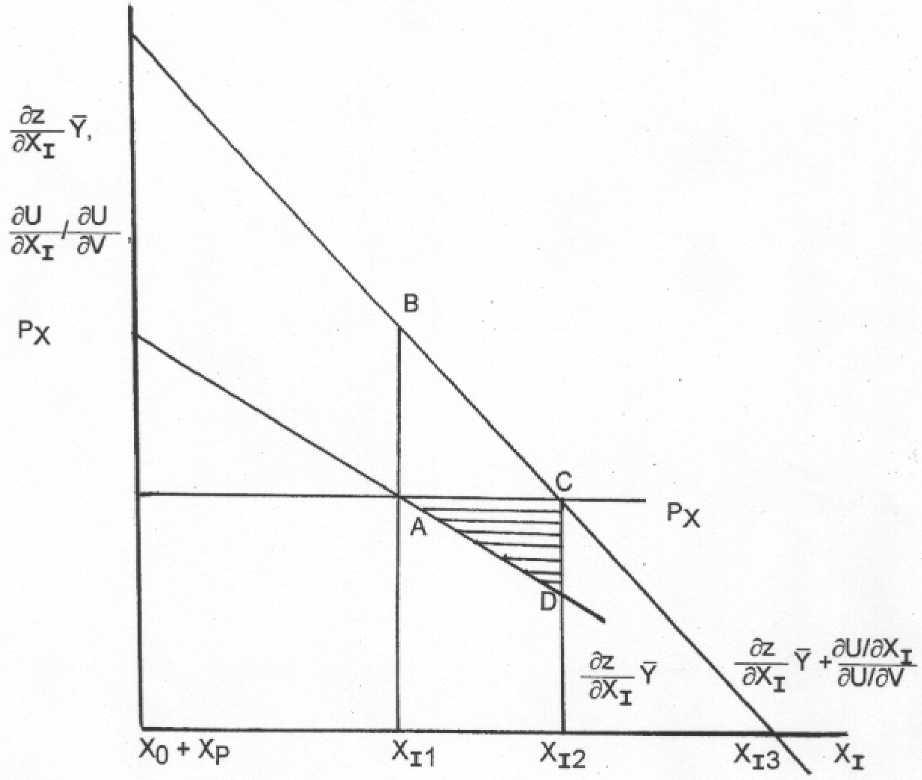


Figure 1. Determination of the Optimum Amount of Health