UNIVERSITY
OF OSLO
HEALTH ECONOMICS
RESEARCH PROGRAMME

Redistribution at the hospital

Anne Wenche Emblem

Faculty of Economics and Social Science
Agder University College

Working Paper 2002: 4

Redistribution at the hospital

By Anne Wenche Emblem*

February 2002

Health Economics Research programme at the University of Oslo
HERO 2002

* *Author's address:*
Faculty of Economics and Social Science,
Agder University College, Servicebox 422
N-4604 Kristiansand, Norway

E-mail : anne.w.emblem@hia.no

Health Economics Research programme at the University of Oslo
Financial support from The Research Council of Norway is acknowledged.
ISSN 1501-9071, ISBN 82-7756-085-0

Redistribution at the hospital

Anne Wenche Emblem anne.w.emblem@hia.no
Faculty of Economics and Social Science
Agder University College, Servicebox 422
N-4604 Kristiansand, Norway.

February 26, 2002

Abstract

This paper studies redistribution by means of a public supply of medical treatment. We show that the government can redistribute income towards low-ability individuals in a world of asymmetric information by offering bundles of medical treatment and redistributive payment. If self-selection is a problem, then the separating scheme offers high-ability individuals complete treatment against a high payment, and low-ability individuals partial treatment against a low payment. In particular, the level of treatment offered low-ability individuals is distorted downwards.

Keywords: health, medical treatment, insurance, redistribution, self-selection

JEL Code: I18, H42, D81

1 Introduction

Medical treatment is but one example of a private good which in a number of countries is publicly supplied. Redistribution, along with paternalism, has been, and probably still is, a major reason why this is so. Public supply of treatment is usually financed by individuals contributing according to their (gross) income, either through general taxation or through earmarked contributions. The supply of treatment, on the other hand, is often provided according to individuals' need for treatment, rather than their contributions. Traditionally, this way of organising the financing and supply of medical treatment is thought to facilitate some kind of income redistribution towards low-income individuals. Underlying this way of reasoning is the assumption that individuals with identical medical needs prefer identical levels of medical treatment. We will, however, show that individuals with identical medical need for treatment may indeed prefer different levels of treatment and, consequently, different levels of recovery. In particular, individuals' preferred level of treatment is shown to be higher the higher their level of innate ability. This suggests that those contributing the most to the public health sector (i.e. high-income individuals) are also those utilising the services provided the most. The extent of redistribution may, therefore, not be as large as one would think.

The purpose of this paper is to identify Pareto-efficient bundles of medical treatment and payment that facilitate income redistribution towards low-ability individuals when information about individuals' level of ability (income) is private to the individuals. To simplify the analysis, we assume that this is the government's only means of redistribution, moreover, there is no private supply of medical treatment. We postulate that the government does *not* pursue a particular distribution of medical treatment (health) *per se*.[1]

There exists a fairly extensive literature on public provision of private goods as a means of redistribution in a second best world, some of which are: Blackorby and Donaldson (1988), Ireland (1990), Besley and Coate (1991), Epple and Romano (1996), Blomquist and Christiansen (1995, 1998), and Boadway, Marchand and Sato (1998).[2] The present paper is most closely related to that of Blackorby and Donaldson (1988, hereafter B-D) in that we use bundles of transfers and medical treatment as a means of redistribution. Whereas B-D study redistribution between ill and healthy individuals, i.e. redistribution *ex post*, we study redistribution between income groups *ex ante*, i.e. prior to knowing whether individuals are ill or not. Consequently, B-D study adverse selection problem in health types (i.e. whether ill or not), whereas we study adverse selection problem in ability types (i.e. whether high or low ability). B-D show that in a second-best world, publicly provided medical treatment to the ill is 'overprovided' in the sense that their marginal willingness to pay for treatment is less than marginal cost of treatment. Here, we will show that subsidised treatment to the ill is 'underprovided', i.e.

[1] One may argue that health is a particular commodity for which the society has distinct egalitarian ambitions, i.e. that health is subject to what Tobin (1970) calls specific egalitarianism. The society may, for instance, aspire to achieve equality in health (e.g. as measured as quality adjusted life years over individuals' lifetime), or somewhat less ambitious: to reduce inequalities in health.

[2] Balestrino (1999) provides a survey of the literature on in-kind transfers in the presence of distortionary taxes.

their marginal willingness to pay is higher than marginal cost of treatment, if their level of innate ability is low. Our findings thus run counter to those of B-D, due to different informational assumptions.

The substance of our model is outlined as follows. Individuals may fall ill thus suffering a loss both in utility directly, and in ability to earn income. They can, however, buy medical treatment that restores health, and thus also ability, with certainty. Medical treatment at a given level is assumed to restore ability in the same proportion for all individuals. Benefits from treatment, and subsequently willingness to pay for treatment, is consequently higher the higher the level of innate ability. Since medical treatment is more valuable to high-ability individuals than to low-ability individuals, the government can separate the two types of individuals by offering two bundles, each specifying level of treatment and level of payment: one containing complete treatment and a high payment; type (i), and one containing partial treatment and a low payment; type (ii). If redistribution leads to a binding self-selection constraint, bundle (i) will not be distortionary, while bundle (ii) will be. In particular, the level of treatment provided in bundle (ii) will be distorted downwards, as this will be more costly to high-ability individuals than to low-ability individuals. The subsidised medical treatment allows low-ability individuals to have a higher level of consumption in the two possible states of the world (healthy or ill), and are, consequently, better off relatively to a situation without redistribution.

This paper is organised as follows. In Section 2 we derive the model and undertake a preliminary analysis, and in Section 3 we derive the government's Pareto-efficient menu of contracts facilitating redistribution towards low-ability individuals. Finally, we discuss our findings in Section 4.

2 Preliminary analysis: Allocation of income across states of the world.[3]

In the subsequent analysis, we describe a representative individual's *ex ante* optimisation problem. The individual has preferences over consumption (c) and health (h): $u(c, h)$. There are two possible states of health: she may with probability $(1 - \pi)$ be in good health: state 1, or with probability π $(0 < \pi \leq 0,5)$ be ill: state 2.[4] The two states are jointly exhaustive and

[3] The analysis in this section is based on a somewhat modified version of a model developed in Asheim, Emblem and Nilssen (2000).

[4] We assume that the individual can neither influence the probability of falling ill, nor the costs associated with the illness, i.e. no moral hazard.

verifiable. Information on risk is symmetrically distributed. Health in state 1 is normalised to 1: $h_1 = 1$, while health in state 2 is assumed to be zero: $h_2 = 0$. Health if ill can (with certainty) be partly or fully restored if medical treatment t ($0 \leq t \leq 1$) is utilised. Treatment leading to complete recovery, i.e. $t = 1$, has a cost of production equal to C, while treatment leading to partial recovery has a cost tC. Health if ill is henceforth represented by t: $h_2 = t$. Consumption in state 1 and 2 is denoted c_1 and c_2, respectively.

The individual is an expected utility maximiser. Hence, her preferences are represented by the following von Neumann-Morgenstern utility function:

$$(1 - \pi)u(c_1, 1) + \pi u(c_2, t). \tag{1}$$

We assume that u is twice continuously differentiable, strictly increasing and strictly concave. Also, $u_{ch} \geq 0$, where the partial derivative is denoted by subscript. Health is thus not only an important factor of well-being in its own right, but may also affect the individual's ability to enjoy consumption. It follows that c and h are normal goods. Furthermore, $u_c(c, h) \to \infty$ as $c \downarrow 0$ whenever $h > 0$ and $u_h(c, h) \to \infty$ as $h \downarrow 0$ whenever $c > 0$. Moreover, $u_c(c, h) \to \infty$ or $u_h(c, h) \to \infty$ as $c \downarrow 0$ and $h \downarrow 0$. Hence, she strictly desires a positive level of consumption and health.

The individual's level of innate ability (productivity) is given by A. Information about innate ability is private to the individual. If in good health, her level of ability is equal to A, while if ill, her level of ability is equal to tA. Earnings are assumed to be proportional to ability.

By the properties of u (strict concavity) it follows that the individual is risk averse. There exists a perfectly competitive private insurance market offering cash compensation if illness occurs. The compensation can be used to cover medical expenditures and partly compensate for (permanent) loss in income due to reduced ability, e.g. in the form of a disability payment. Insurance is offered at an actuarially fair premium πI, so that πI must be paid in both states in order to have coverage equal to I if ill. To insure is the only way the individual can transfer income across the two possible states of the world. Let P represent payment for treatment. Then her budget constraint in state 1 and 2 is: $c_1 + \pi I = A$ and $c_2 + \pi I + P = tA + I$, respectively. Her ex ante choice of c_1 and c_2 when the budget constraint is binding in both states of the world, is thus constrained by:

$$(1 - (1 - t)\pi) A = (1 - \pi)c_1 + \pi(c_2 + P). \tag{2}$$

The individual's expected utility maximising choice of consumption in the two states of the world is derived by maximising eq.(1) subject to eq.(2).

The Lagrangian is given by:

$$\begin{aligned}\mathcal{L}(c_1, c_2, \lambda; t, P, A) &= (1-\pi)u(c_1, 1) + \pi u(c_2, t) \\ &\quad +\lambda\left[(1-(1-t)\pi)A - (1-\pi)c_1 - \pi(c_2 + P)\right].\end{aligned}$$

It follows from the first-order conditions that:

$$u_c(c_1(t, P, A), 1) = u_c(c_2(t, P, A), t) = \lambda, \tag{3}$$

that is; marginal utility of consumption is equal across states. The consumption demand function in each state is given by: $c_1(t, P, A)$ and $c_2(t, P, A)$. The private insurance market consequently allows her to attain her optimal distribution of consumption across states. In the subsequent analysis, we will therefore concentrate on characterising her preferences if ill; in particular, her marginal willingness to pay for medical treatment at the expense of consumption if ill.

The individual's indirect utility function is given by:

$$V(t, P, A) = (1-\pi)u(c_1(t, P, A), 1) + \pi u(c_2(t, P, A), t). \tag{4}$$

V is strictly increasing in t, strictly decreasing in P, and strictly increasing in A. We can therefore define a curve, call it $\mathcal{P}(t, A; \bar{t}, \overline{P})$, going through $(\bar{t}, \overline{P})$ in (t, P)-space and showing combinations of t and P yielding a constant level of utility. Accordingly, the utility, $V(t, P, A)$, of an individual with ability A facing (t, P) is equal to $V(\bar{t}, \overline{P}, A)$ if and only if $P = \mathcal{P}(t, A; \bar{t}, \overline{P})$. The indifference curve is upward sloping both in t and A. Moreover,

$$\begin{aligned}\frac{\partial \mathcal{P}(t, A; \bar{t}, \overline{P})}{\partial t} &= -\frac{\frac{\partial V}{\partial t}}{\frac{\partial V}{\partial P}} = -\frac{\frac{\partial L}{\partial t}}{\frac{\partial L}{\partial P}} = \frac{\pi(u_h(c_2, t) + \lambda A)}{\pi\lambda} \\ &= \frac{u_h(c_2, t)}{u_c(c_2, t)} + A,\end{aligned} \tag{5}$$

where the second equality follows from the envelope theorem, and the fourth equality is implied by the first-order condition in eq.(3). This means that the marginal willingness to pay for treatment equals the sum of consumption and production value[5] of health.

The indifference curve $\mathcal{P}(t, A; \bar{t}, \overline{P})$ can be shown to be a strictly concave function of t. If follows that the marginal willingness to pay for treatment

[5] As measured by the additional earnings capacity generated by treatment.

is positive and decreasing in t. Moreover, since $u_{cc} < 0$, $u_{ch} \geq 0$ and $\partial c_2/\partial A > 0$, then eq. (5) implies that:

$$\frac{\partial \mathcal{P}(\bar{t}, A; \bar{t}, \overline{P})}{\partial t \partial A} = \frac{\partial}{\partial A}\left[\frac{u_h(c_2(\bar{t}, \overline{P}, A), \bar{t})}{u_c(c_2(\bar{t}, \overline{P}, A), \bar{t})} + A\right] > 1, \tag{6}$$

i.e. marginal willingness to pay for treatment is **increasing** in her level of innate **ability**. Graphically, this implies that the slope of an indifference curve through any point $(\bar{t}, \overline{P})$ is increasing in A. Indifference curves are consequently steeper the higher the level of innate ability and they cross only once, i.e. **single-crossing**.[6] The single-crossing property is illustrated in Figure 1 for two different values of ability: $A_L < A_H$, where L and H denotes low and high ability, respectively.

Holding t and A constant, then a higher payment for treatment implies that she will have to reduce her level of consumption in state 2. Marginal willingness to pay for treatment is consequently **decreasing** in P:

$$\frac{\partial \mathcal{P}(\bar{t}, A; \bar{t}, \overline{P})}{\partial t \partial \overline{P}} = \frac{\partial}{\partial \overline{P}}\left[\frac{u_h(c_2(\bar{t}, \overline{P}, A), \bar{t})}{u_c(c_2(\bar{t}, \overline{P}, A), \bar{t})} + A\right] < 0, \tag{7}$$

hence, the income effect of a higher P is negative, as would be expected. It follows that the slope of the indifference curve through any point $(\bar{t}, \overline{P})$ is decreasing in payment.

The following result is useful for the main analysis. It shows that if innate ability is sufficiently high, then marginal willingness to pay for treatment exceeds cost of treatment, independently of payment. Moreover, if innate ability is sufficiently low, and payment equals cost of treatment, then marginal willingness to pay for treatment equals marginal cost only if treatment is partial.

Lemma If $A \geq C$, then $\partial \mathcal{P}(t, A; t, P)/\partial t > C$ for all t and P. If $A \leq \pi C$, then $\partial \mathcal{P}(t, A; t, tC)/\partial t = C$only if $0 < t < 1$.

Proof. Recalling that C is the production cost of complete treatment, then if $A \geq C$, it follows from eq.(5) that the individual's marginal willingness to pay for treatment: $\partial \mathcal{P}(t, A; t, P)/\partial t$, is greater than C. Assume now that $A \leq \pi C$. If $t = 1$ and $P = C$, then we see from eq.(2) and (3) that $c_1 = c_2 = A - \pi C$, implying that $t = 1$ is not feasible when $A \leq$

[6] The single-crossing property corresponds to the 'Agent Monotonicity condition' in the literature on income taxation (Seade, 1982) and the 'Spence-Mirrlees condition' in the literature on screening (Macho-Stadler & Pérez-Castrillo, 1997).

πC. Moreover, if $A \downarrow \pi C$ and $t = 1$, then it follows from eq. (2) that: $c_2 = (A - \pi C + c_1(\pi - 1))\frac{1}{\pi} \leq (A - \pi C)\frac{1}{\pi} \downarrow 0$. By the properties of u[7], $\partial \mathcal{P}(1, A; 1, tC)/\partial t \to \pi C < C$ when $A \downarrow \pi C$. Consequently, if $A \leq \pi C$, then $\partial \mathcal{P}(t, A; t, tC)/\partial t = C$ only if $0 < t < 1$. ■

3 Main analysis: Health vs subsidy

We now expand the analysis to include two types of individuals who are identical in all respects save their individual level of innate ability: they may either have a high ability: $A_H \geq C$, or a low ability: $A_L \leq \pi C$.[8] The number of individuals of each type is given by N_i, $i = H, L$. The government knows the proportion of each type of individuals, but cannot observe their identity. The individuals' indirect utility function is given by eq.(4).

The government designs a menu of Pareto-efficient contracts specifying bundles of payment and treatment: $\{(t_L, P_L), (t_H, P_H)\}$, where (t_L, P_L) and (t_H, P_H) denotes the contract intended for low- and high-ability individuals, respectively. Contracts cannot be traded once they have been signed, moreover, medical treatment can not be supplemented. Contracts specifying combinations of treatment and payment are derived by:

$$\max_{(t_L, P_L),(t_H, P_H)} V(t_L, P_L, A_L)$$

subject to:

$$\begin{aligned}
\underline{V} &\leq V(t_H, P_H, A_H) \\
V(t_H, P_H, A_L) &\leq V(t_L, P_L, A_L) \\
V(t_L, P_L, A_H) &\leq V(t_H, P_H, A_H) \\
N_L(P_L - t_L C) + N_H(P_H - t_H C) &= 0 \\
0 &\leq t_i \leq 1,\ i = H, L.
\end{aligned}$$

The first constraint ensures high ability individuals a certain level of utility. The second and the third constraints are the self-selection constraints. The fourth constraint is the government's balanced budget constraint, while the fifth constraint is the restriction that individuals cannot obtain more

[7] $u_c(c, h) \to \infty$ as $c \downarrow 0$ whenever $h > 0$.

[8] By assuming that they face identical risk of falling ill, we disregard questions regarding the relationship between ability and health (e.g. whether the likelihood of falling ill is correlated with the individuals' ability). Obviously, there is a relationship between health and socio-economic status, a fact that is of importance to the discussion of redistribution, yet the direction (and the strength) of causation is not straightforward.

than complete treatment (i.e. $t \leq 1$), nor 'sell' treatment (i.e. $t \geq 0$). As can be checked, the self-selection constraint and the constraint on the level of treatment are both satisfied for low-ability individuals in the subsequent analysis. It also holds that the non-negativity constraint on treatment is satisfied for high-ability individuals. Forming the Lagrangian:

$$\begin{aligned} \mathcal{L} &= V(t_L, P_L, A_L) + \lambda [V(t_H, P_H, A_H) - \underline{V}] + \mu [V(t_H, P_H, A_H) - V(t_L, P_L, A_H)] \\ &\quad + \gamma [N_L(P_L - t_L C) + N_H(P_H - t_H C)] + \phi (1 - t_H). \end{aligned}$$

The efficient t_i and P_i satisfy the conditions:

$$\frac{\partial \mathcal{L}}{\partial t_L} = \frac{\partial V(t_L, P_L, A_L)}{\partial t} - \mu \frac{\partial V(t_L, P_L, A_H)}{\partial t} - \gamma N_L C = 0 \quad (8)$$

$$\frac{\partial \mathcal{L}}{\partial P_L} = \frac{\partial V(t_L, P_L, A_L)}{\partial P} - \mu \frac{\partial V(t_L, P_L, A_H)}{\partial P} + \gamma N_L = 0 \quad (9)$$

$$\frac{\partial \mathcal{L}}{\partial t_H} = \lambda \frac{\partial V(t_H, P_H, A_H)}{\partial t} + \mu \frac{\partial V(t_H, P_H, A_H)}{\partial t} - \gamma N_H C - \phi = 0 \quad (10)$$

$$\frac{\partial \mathcal{L}}{\partial P_H} = \lambda \frac{\partial V(t_H, P_H, A_H)}{\partial P} + \mu \frac{\partial V(t_H, P_H, A_H)}{\partial P} + \gamma N_H = 0 \quad (11)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} \geq 0, \frac{\partial \mathcal{L}}{\partial \mu} \geq 0, \frac{\partial \mathcal{L}}{\partial \gamma} \geq 0, \frac{\partial \mathcal{L}}{\partial \phi} \geq 0. \quad (12)$$

For the moment, we assume that the government has no redistribution ambitions, and that payment reflects cost of production. In addition, we assume that the self-selection constraint on high-ability individuals is not binding (i.e. $\mu = 0$). Since, by assumption, the government has only one means of redistribution, the following provides a benchmark against which redistribution can be compared. Dividing eq.(8) by eq.(9) and using eq.(5) we find:

$$\partial \mathcal{P}(t_L, A_L; t_L, t_L C)/\partial t = C,$$

hence, low-ability individuals' marginal willingness to pay for treatment equals cost of treatment. Since $A_L \leq \pi C$, it follows from Lemma 1 that $0 < t_L < 1$, thus: $P_L = t_L C$. Dividing eq.(10) by eq.(11) and using eq.(5) we get:

$$\partial \mathcal{P}(t_H, A_H; t_H, t_H C)/\partial t = C + \phi/\gamma N_H.$$

The government's zero-revenue constraint is assumed to hold, thus $\gamma > 0$. The marginal imputed cost incurred in restraining the individuals' level of treatment (i.e. $t_H \leq 1$) is given by ϕ. From Lemma 1 we know that $\phi > 0$, and thus $t_H = 1$, implying that the treatment constraint is binding. Moreover, $P_H = C$. Consequently, $\underline{V} = V(1, C, A_H)$.

The efficient bundles of treatment and payment when there is no redistribution: $\{(t_L, t_L C), (1, C)\}$, are illustrated in Figure 2. As can be seen, self-selection will not be a problem since high-ability individuals would suffer a loss in utility if choosing low-ability individuals' bundle (and since low-ability individuals can not afford high-ability individuals' bundle).

Redistribution

We now assume that the government has **redistribution ambitions**, in particular, it seeks to maximise the sum of utilities, i.e. a **utilitarian welfare function**. This corresponds to $\lambda = N_H/N_L$, entailing that the weight of high-ability individuals relative to low-ability individuals corresponds solely to the numbers of individuals in each group. Utilitarianism leads to redistribution from high- to low-ability individuals if $\lambda > N_H/N_L$ in the situation without redistribution. In the subsequent, we will show that λ is indeed greater than N_H/N_L when there is no redistribution.

Since $\mu = 0$ when the government does not redistribute, it follows from eq.s (9) and (11) that $\lambda > N_H/N_L$ corresponds to: $-\partial V(t_L, t_L C, A_L)/\partial P > -\partial V(1, C, A_H)/\partial P$. Since, by eq.(3), $-\partial V(t, P, A)/\partial P = -\partial \mathcal{L}/\partial P = \pi u_c(c_1, 1)$, it follows, as $u_{cc} < 0$, that λ in a situation without redistribution exceeds N_H/N_L if and only if $c_1(1, C, A_H) > c_1(t_L, t_L C, A_L)$. To show that this is the case, note that it follows from the constraint in eq.(2) and the fact that consumption in state 2 is non-negative, that: $(1 - (1 - t_L)\pi) A \geq (1 - \pi) c_1(t_L, t_L C, A_L) + \pi(t_L C)$. Hence, since $A_L \leq \pi C$, we obtain $c_1(t_L, t_L C, A_L) \leq (1 - t)\pi C$. Moreover, it follows from eq.(3) and the constraint in eq.(2) that high-ability individuals' level of consumption when $t_H = 1$ is given by $c_1(1, C, A_H) = c_2(1, C, A_H) = A_H - \pi C$. Recalling that $\pi \leq 0.5$, we thus see that $c_1(1, C, A_H) > c_1(t_L, t_L C, A_L)$. The level of consumption in state 1 is, in other words, higher for high-ability individuals than for low-ability individuals. Consequently, we have established that the government under utilitarianism wants to redistribute towards low-ability individuals.

Proposition 1 If the government has a utilitarian welfare function, then income is redistributed from high-ability individuals to low-ability individuals.

However, even if utilitarianism leads to redistribution from high- to low-ability individuals, we cannot determine without further assumptions whether redistribution is carried out to the extent that the self-selection constraint on high-ability individuals is binding. Self-selection may in fact not be a problem in the utilitarian optimum even if marginal utility of consumption is equalised across individuals and states. This is because the low-ability individuals' level of consumption and treatment in state 2 may be sufficiently low to prevent high-ability individuals from wanting to masquerade. In the following, we will study the situation where the self-selection constraint binds, i.e. $\mu > 0$, using the same approach to the self-selection problem as suggested by Stiglitz (1987).

The bundle intended for the low-ability individuals is found by dividing eq.(8) by eq. (9):

$$\frac{\mu\frac{\partial V(t_L,P_L,A_H)}{\partial t} + \gamma N_L C.}{-\mu\frac{\partial V(t_L,P_L,A_H)}{\partial P} + \gamma N_L} = -\frac{\frac{\partial V(t_L,P_L,A_L)}{\partial t}}{\frac{\partial V(t_L,P_L,A_L)}{\partial P}}.$$

Defining $v \equiv \mu\frac{\partial V(t_L,P_L,A_H)}{\partial P}/\gamma N_L$ and using eq.(5), we can rewrite the condition as:

$$\frac{\partial \mathcal{P}(t_L,A_L;t_L,P_L)}{\partial t} = C + \left(\frac{\partial \mathcal{P}(t_L,A_H;t_L,P_L)}{\partial t} - C\right)\frac{v}{v-1}.$$

From Section 2 we know that high-ability individuals' marginal willingness to pay for treatment is higher than that of low-ability individuals at any treatment-payment combination, so also for (t_L, P_L): $\partial\mathcal{P}(t_L, A_H; t_L, P_L)/\partial t > \partial\mathcal{P}(t_L, A_L; t_L, P_L)/\partial t$. Since $v < 0$[9], it follows that $C < \partial\mathcal{P}(t_L, A_L; t_L, P_L)/\partial t < \partial\mathcal{P}(t_L, A_H; t_L, P_L)/\partial t$ in the self-selection equilibrium. Low-ability individuals are, in other words, offered a level of treatment t_L at which their marginal willingness to pay exceeds the marginal cost of production. The level of treatment is consequently distorted downwards. Graphically, the slope of both high- and low-ability individuals' indifference curve through the point (t_L, P_L) exceeds the slope of the isocost line.

The bundle intended for high-ability individuals is derived by diving eq. (10) by eq. (11), and using eq.(5):

$$\frac{\partial \mathcal{P}(t_H,A_H;t_H,P_H)}{\partial t} = C + \frac{\phi}{\gamma N_H}.$$

By assumption, $\gamma > 0$. From Lemma 1 it follows that $\phi > 0$ and thus $t_H = 1$. Hence, the marginal willingness to pay for treatment exceeds the marginal

[9] Because $\mu \geq 0, \gamma \geq 0$ and $\partial V_H/\partial P < 0$.

cost of treatment for the non-distortionary reason that treatment cannot restore health beyond its original level. Therefore, the contract intended for the high-ability individuals is not distorted at the margin.[10]

Proposition 2 If the self-selection constraint binds, then the optimal separating scheme is such that:

(i) High-ability individuals' bundle of treatment and payment is not distortionary.

(ii) Low-ability individuals' bundle of treatment and payment is distortionary. In particular, the level of medical treatment is distorted downwards since this is more costly to high-ability mimickers than to low-ability individuals.

Consequently, if the self-selection constraint binds, the government induces individuals to reveal information by offering two bundles: $\{(t_L, P_L), (1, P_H)\}$, where $0 < t_L < 1$.[11] High-ability individuals are discouraged from masquerading as low-ability individuals by restricting the level of treatment available to low-ability individuals. The attained self-selection equilibrium is illustrated in Figure 3.

Relating our findings to those of B-D; we have shown that when information on ability is asymmetric (and self-selection is a problem), then treatment is 'underprovided' to ill low-ability individuals, whereas B-D show that when information on health status (ill/healthy) is asymmetric, then treatment is 'overprovided' to ill individuals.

Extent of insurance coverage

We apply the following terminology: by **full insurance**, we mean that $u(c_1, h_1) = u(c_2, h_2)$, i.e. that utility is constant across the two states, and by **partial insurance**, we mean that $u(c_1, h_1) > u(c_2, h_2)$, i.e. that utility is lower if ill than if healthy.

High-ability individuals are defined by $A_H \geq C$. From Lemma 1 we know that $t_H = 1$, hence, high-ability individuals' level of ability will be equal across states: $h_1 = h_2 = 1$. Moreover, their level of consumption in the

[10] This is analogous to the optimal taxation problem where the marginal tax rate faced by high-ability individuals is zero, while the marginal tax rate faced by low-ability individuals is positive (Stiglitz, 1987). This is often called a 'non-distortion at the top' property, where 'top' refers to individuals that no one would choose to masquerade as.

[11] Payment may in fact be negative if the subsidy is large, that is; individuals may not only receive medical treatment, but also a cash transfer from the hospital if ill (the cash transfer then being contingent on the individual undergoing treatment).

two states of the world will be identical: $c_1(1, P_H, A_H) = c_2(1, P_H, A_H) = A_H - \pi P_H$. It follows that $u(c_1, h_1) = u(c_2, h_2)$ and, consequently, they are *fully insured*.

Low-ability individuals are defined by $A_L \leq \pi C$. From the above analysis we know that $0 < t_L < 1$, hence, $h_1 = 1$ and $h_2 = t_L < 1$. It follows from eq. (3) and the properties of u that $c_1(t_L, P_L, A_L) \geq c_2(t_L, P_L, A_L)$. Consequently, $u(c_1, h_1) > u(c_2, h_2)$ and the individuals are *partly insured*. When the government redistributes and self-selection is a problem, then low-ability individuals' level of medical treatment is distorted downwards. Their level of insurance coverage is hence reduced relatively to a situation without redistribution. They will, however, receive a subsidy which enables them to have a level of consumption in excess of their earnings minus the cost of treatment, and which makes them better off.

4 Discussion

So far, we have studied the individuals' *ex ante* decision regarding their optimal level of consumption in the two states of the world, as well as their optimal level of medical treatment given the two bundles. If individuals are rational and have perfect foresight, then their *ex post* preferred level of medical treatment will also be preferred *ex ante*. Prior to knowing which state of the world has occurred, they are therefore willing to sign a contract with the public supplier, i.e. hospital, specifying both payment and level of treatment that is to be made available if ill. Indeed, one would expect that the individuals would prefer to do so as this would prevent them from potential transaction costs associated with having to 'shop around' for the appropriate contract when ill. Such a scheme would in fact also be more in accordance with what may be observed empirically: Public supply of medical treatment financed by individuals contributing through (earmarked) contributions and supplied free at the point of delivery. The public supply of treatment may thus be thought of as an *insurance* where individuals are compensated in the form of medical treatment directly if illness occurs, i.e. *indemnity in kind*. The preceding analysis will still be valid in such a setting, the only difference is that individuals *ex ante* determine their optimal level of private and public insurance coverage.[12]

[12] Such a setting would be more in line with the analysis in Asheim *et al.*, (2000) where individuals' demand for insurance against medical expenditures (i.e. in kind) and/or loss in income due to disability (i.e. in cash) are integrated. Their analysis takes place in a perfectly competitive private market with no public interference.

The preceding analysis is based on a highly stylised model. We assume that individuals with certainty can recover completely by consuming the appropriate level of medical treatment. Moreover, illness is presumed to be observable, hence, the private insurers do not face problems of adverse selection. We have also ignored differences in the risk of falling ill which indeed is an important reason for having a (mandatory) public health insurance (Breyer and Haufler, 2000). Possible commitment problems associated with a re-optimising government aspiring to increase treatment above the 'announced' level, is not discussed. The perhaps most striking assumption underlying this analysis is, however, the assumption that the government cannot implement income taxation. Consequently, we cannot infer whether redistribution through public pricing of contracts is more efficient than other means of redistribution. Extension of the analysis to include also individuals' labour supply and distortionary income taxation, will be the subject of future research. We think still that our analysis provides interesting results, results that run counter to what often seems to be implicitly underlying many studies of public provision of health care; namely that individuals wish to fully recover from an illness.

Lastly, our analysis follows the tradition of neo-classical economics in that individuals' utility (or preferences) provide the foundation of the analysis. We study distribution of welfare, that is; well-being assessed in utility terms, thus, we assume that the value of medical treatment to an individual is represented by her willingness to pay for it.[13] An alternative approach would be the extra-welfarist framework in which health, and not utility, is the primary outcome of interest (Hurley, 1998). According to this approach, distributional equity (in the egalitarian concept) implies that medical treatment should be provided according to need, and not according to ability or willingness to pay.

Acknowledgment

We are grateful to Geir B. Asheim, Robin Boadway, Tore Nilssen, Agnar Sandmo and an anonymous referee of this journal for their valuable comments and suggestions. We are also grateful for helpful comments from the participants at the Oslo workshop on Health Economics and the 57th Congress of the IIPF in Linz. The usual disclaimer applies. Financial support from the SiS-programme at Agder University College is gratefully acknowledged.

[13] Moreover, we disregard any merit good arguments which may justify a separation of ability to pay and treatment provided, an argument which is prevailing in the redistributional objectives in many countries.

References

[1] Asheim, G.B., A.W. Emblem and T. Nilssen. (2000). "Health insurance: Treatment vs. compensation." Available at: http://www.oekonomi.uio.no/memo/memopdf/memo4000.pdf

[2] Balestrino, A. (1999). "The desirability of in-kind transfers in the presence of distortionary taxes." Journal of Economic Surveys, 13(4), 333-354.

[3] Besley, T. and S. Coate. (1991). "Public provision of private goods and the redistribution of income." The American Economic Review, 81(4), 979-984.

[4] Blackorby C. and D. Donaldson. (1988). "Cash versus kind, self-selection, and efficient transfers." The American Economic Review, 78(4), 691-700.

[5] Blomquist S. and V. Christiansen. (1995). "Public provision of private goods as a redistributive device in an optimum income tax model." Scandinavian Journal of Economics, 97(4), 547-567.

[6] Blomquist S. and V. Christiansen. (1998). "Price subsidies versus public provision." International Tax and Public Finance, 5 (3), 283-306.

[7] Boadway, R., M. Marchand and M. Sato. (1998). "Subsidies versus public provision of private goods as instruments for redistribution." Scandinavian Journal of Economics, 100(3), 545-564.

[8] Breyer, F. and A. Haufler. (2000). "Health care reform: Separating insurance from income redistribution." International Tax and Public Finance, 7(4-5), 445-461.

[9] Epple, D. and R.E. Romano. (1996). "Public provision of private goods." Journal of Political Economy, 104 (1), 57-84.

[10] Hurley, J. (1998). "Welfarism, extra-welfarism and evaluative economic analysis in the health sector." In: M.L. Barer, T.E. Getzen and G.L. Stoddart (eds.), Health, Health Care and Health Economics, 373-395. John Wiley & Sons Ltd.

[11] Ireland, N.J. (1990). "The mix of social and private provision of goods and services." Journal of Public Economics, 43(2), 201-219.

[12] Macho-Stadler, I. and J.D. Pérez-Castrillo. (1997). *An introduction to the economics of information. Incentives & contracts.* Oxford University Press.

[13] Seade, J. (1982). "On the sign of the optimum marginal income tax." *Review of Economic Studies*, 49(4), 637-643.

[14] Stiglitz, J.E. (1987). "Pareto efficient and optimal taxation and the new new welfare economics." In: A.J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, 2, 991-1041. North Holland.

[15] Tobin, J. (1970). "On limiting the domain of inequality." *Journal of Law and Economics*, 13(2), 263-277.

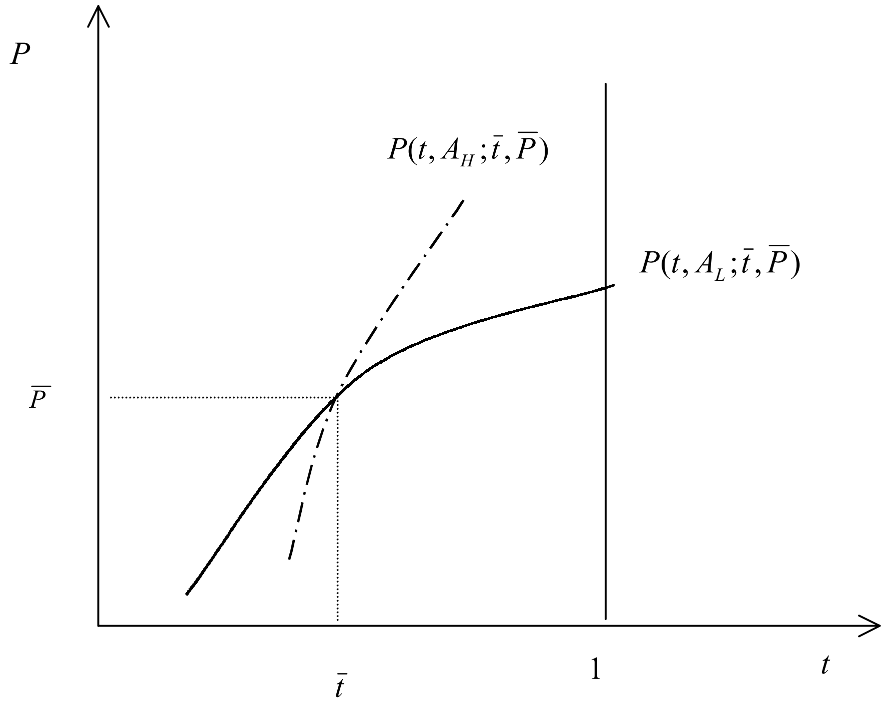

Figure 1. The single-crossing property.

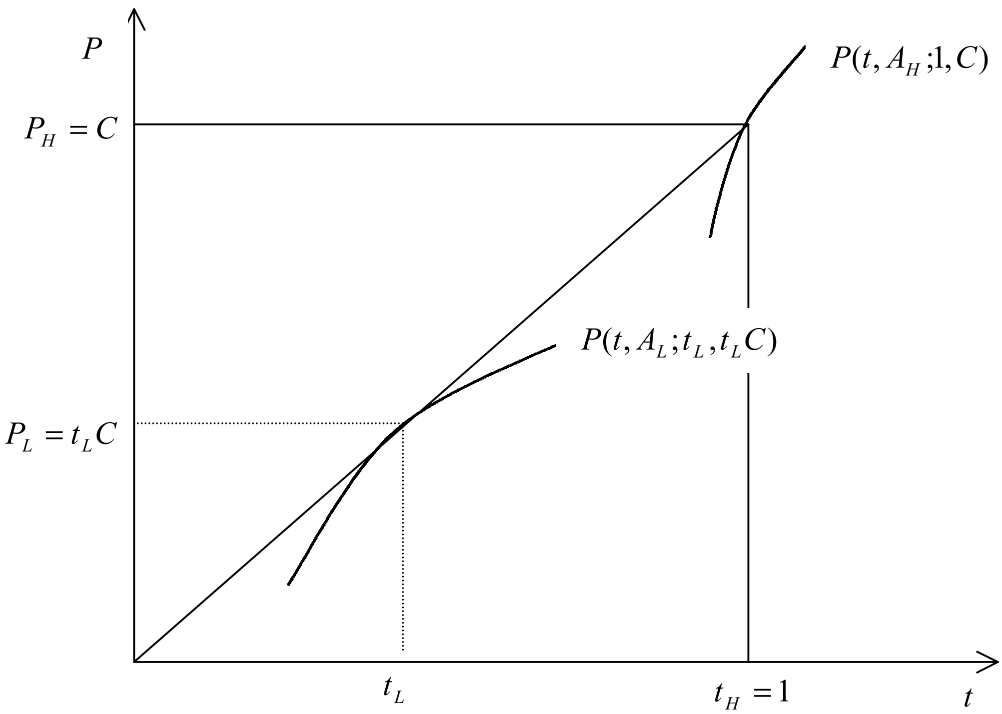

Figure 2. Efficient bundles of treatment and payment when there is no redistribution.

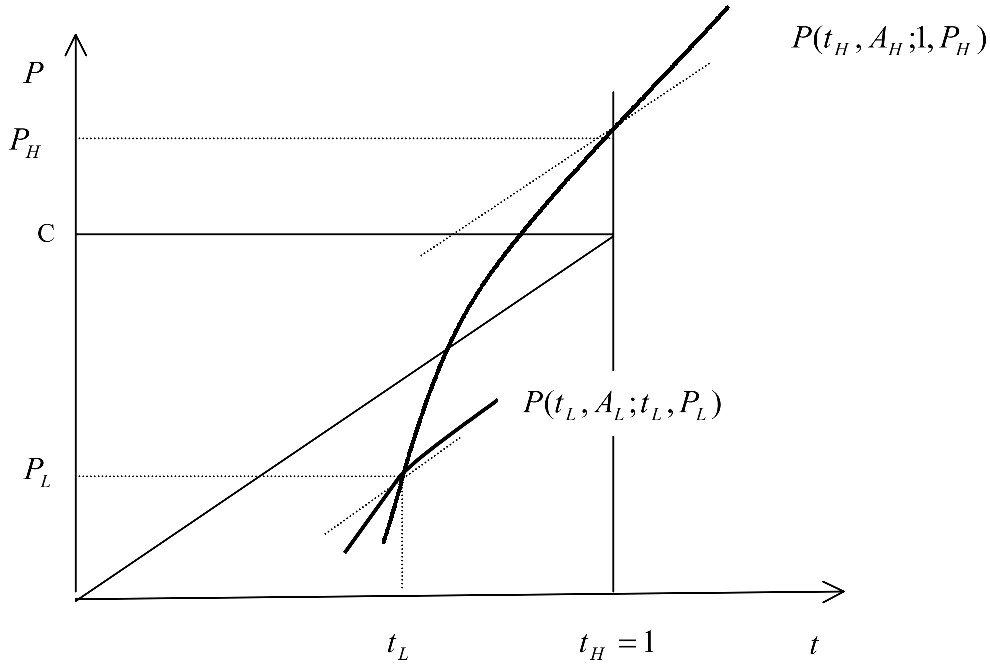

Figure 3. Efficient bundles of treatment and payment when self-selection is a problem.